BIOCHEMICAL INVESTIGATION OF THE ENZYMES INVOLVED IN TAURINE METABOLISM

by

JESSICA RAE O'BRIEN

(Under the Direction of William N. Lanzilotta)

ABSTRACT

Structural and biochemical work on the enzymes taurine α-ketoglutarate dependent dioxygenase TauD (*Esherichia coli*), cysteine dioxygenase CDO (*Homo sapiens*) and cysteine sulfinic acid decarboxylase CSAD (*Homo sapiens*) have provided the necessary groundwork from which biomedical studies can now proceed in studying disease states associated with the taurine metabolic pathway and the respective enzymatic family members. TauD is the enzyme responsible for the conversion of taurine (2-aminoethanesulfonate) to sulfite and aminoacetaldehyde while concurrently decomposing α-ketoglutarate to succinate and carbon dioxide. Expressed during conditions of sulfur starvation, TauD enables the cell to utilize taurine, and other similar sulfonates as an alternative sulfur source. The crystal structure of TauD solved in the apo (1.9 Å resolution, R_{cryst}=21.2%, R_{free}=24.9%) and holo (2.5 Å resolution, R_{cryst}=22.5%, R_{free}=24.9%) forms revealed two possible dimeric arrangements; the one supported in this work has a significant amount of evidence to suggest the functional importance of dimerization. The 2-his 1-carboxylate facial triad was identified in this structure as well as the residues involved in substrate binding for both taurine and α-ketoglutarate. From this structure a feasible reaction mechanism is proposed and is currently being tested through mutagenic and

crystallographic studies. A significant area of interest is the method of oxygen activation, which remains unclear. Studies with another dioxygenase, CDO, hope to provide insight into the method of oxygen activation. CDO and CSAD are the two enzymes involved in the taurine biosynthetic pathway primarily located in the human brain, heart, liver, and central nervous system. Since taurine is a known osmoregulator, neurotransmitter, radical scavenger, and is central to the cholesterol excretion pathway, any information gathered on these enzymes provides the base for future disease research. The recombinant expression of human CDO and CSAD is new research. Two forms of CDO were observed, a 23 kDa form and a 25 kDa form. Here we report the expression, purification, and kinetic information for CDO.

INDEX WORDS: taurine, dioxygenase, α-ketoglutarate dependent dioxygenase, cysteine metabolism, taurine α-ketoglutarate dependent dioxygenase, TauD, cysteine dioxygenase, CDO, cysteine sulfinic acid decarboxylase, CSAD

BIOCHEMICAL INVESTIGATION OF THE ENZYMES INVOLVED IN TAURINE METABOLISM

by

JESSICA RAE O'BRIEN

Bachelor of Science, Shepherd University, 2001

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2006

BIOCHEMICAL INVESTIGATION OF THE ENZYMES INVOLVED IN TAURINE METABOLISM

by

JESSICA RAE O'BRIEN

Major Professor: William N. Lanzilotta

Committee: Harry Dailey
B.C. Wang

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
August 2006

DEDICATION

I dedicate this dissertation to my family. Without the collective help of my family this degree would never have been completed and my life would have been plagued with a “what if”.

ACKNOWLEDGEMENTS

Science is in your soul. It must be recognized, bridled, broken, exercised, and then harnessed. Otherwise the natural curiosity goes off in every direction never creating a lasting effect and never making a difference. An odd way to begin an acknowledgements section, but it is necessary to show how the following individuals have contributed to focusing my natural curiosity into a scientific career that will make an impact.

My first acknowledgement is to my heavenly father that has given me the driving curiosity and the desire to seek truth in the foundations of my soul. Without this desire for truth, a casual observer of nature becomes an artist, equally important, but a completely different path.

My parents, Bobby and Nora Martin, deserve a standing ovation for dealing with an unbelievably explorative child. From cleaning up the messes from my random dissections of earthworms, to wondering where the laundry detergent kept disappearing to in my early experimental days, and finally agreeing to pay for my tutoring when I wanted a solid foundation in math during high school. My sister, Jackie Mitchem, was forced to live in a household with a sister that was not satisfied with "because I said so" and often ended up with her clocks, bicycles, and tapes taken apart so I could see how they worked.

My family has helped me fuel the flames that became this career early on when my aunts, Frances, Pat & Myrtle Bair bought me a chemistry set, a pumping heart, and my first microscope. They were always willing to go along with whatever experiment I had planned for the weekends I had with them.

My husband, Jay O'Brien, has been my rock through this adventure. In all of the times I wanted to quit and all of the times I came home knowing that I had failed tremendously throughout my day, he picked me up and helped me move on. He gave me my daughter, Taylor O'Brien, who is the light of my life. Without his unwavering support, help at home, and his shoulder to cry on I would not have completed my degree.

These people were all involved with providing a strong foundation for which my taming process could begin. My parade of tamers are next...

Mrs. Newcomb, my anatomy and physiology teacher in high school first recognized my potential and encouraged me to do my best. She recognized my zeal for science and turned that into an insatiable appetite for information. Additionally, she began my career path in the pursuit of a Bachelor of Science degree in college through much encouragement and letters of recommendation.

In college I met Dr. Burt Lidgerding, a man that would ultimately change my life. He recognized my potential and began the bridling process. As he introduced me to the world of cellular biology and virology I began to realize that things were much more complicated than I ever imagined and that there may be a realm where I could intensely learn and investigate without reaching an endpoint. The thought was exhilarating. He introduced me to Dr. Robert Warburton, my first biochemistry professor, who then in turn introduced me to the world of biochemistry. The rest is history.

Dr. Lidgerding challenged me to do the impossible, supported me when I needed help, and told me he was proud of me when I accomplished it. In all of my scientific endeavors, I

know that there is someone out there that has invested a lot in me and who will always believe in me.

My committee has taken on the role of breaking or taming me. Dr. Harry Dailey and Dr. Mike Johnson have done this through challenging me in my written and oral exams. Dr. Dailey has provided me with much guidance both academically and scientifically. Not to mention the amount of editing time he sacrificed for this dissertation. Dr. B.C. Wang has been an encourager to me and an ever ready scientific guide. My committee has done a wonderful job of raising me up as a scientist.

Tammy Dailey deserves a medal as far as my degree is concerned. If it can be cloned and expressed, she is the one that can make it happen. She has proofread for me, given me experimental guidance, and has worked side-by-side with me on several projects. For this I am incredibly grateful.

My major professor, Dr. William Lanzilotta, is the last, but not least on my list. This is because his job was to focus my energy and build me up to be a productive scientist. I was not necessarily what he pictured as his first graduate student and I'd like to think that we have managed to break each other in along the journey. Dr. Lanzilotta has seen me at my best, chatting away to my experiments, and at my worst, hanging my head in the glove box crying. I don't think it has been an easy road for either of us, but getting a degree is not about the degree itself, it is about the journey. I have grown tremendously as a person and as a scientist by being Dr. Lanzilotta's first graduate student. I thank him for the opportunity to work under his direction and for the opportunity to learn from his wealth of knowledge.

TABLE OF CONTENTS

Page

CHAPTER 1

INTRODUCTION

In the United States cardiovascular disease is the leading cause of death in both men and women over the age of thirty-five. According to the American Heart Association, one million Americans die each year due to cardiovascular disease thus constituting 42% of all deaths (*1*). Cardiovascular disease is traditionally thought of as a disease primarily affecting men. The US Food and Drug Administration reports that, while men will suffer from three times as many heart attacks and strokes as women, the death rate among women is almost twice that among men (*1*). In women, more deaths result from cardiovascular disease than the next six leading causes of death combined (*2*).

Since the prediction of sudden heart attack and stroke is complicated and often impossible, prevention is the best method of lowering the cardiovascular death toll. While diet and exercise remain important in prevention, some patients are still prone to cardiovascular disease regardless of their lifestyle (*3*). In an effort to aid these patients, the pharmaceutical industry has introduced a plethora of cholesterol lowering medications (*4*). Family medical history is the most accurate tool for predicting the onset of cardiovascular disease (*5*). Unfortunately, research finds that physicians are unable to predict the sudden onset of heart attack even when family medical history and cholesterol levels are taken into consideration (*6*). In this study the medical records of 326 patients were collected within a six month period prior to their death by sudden heart attack (*6*). Physicians were unable to predict a heart attack in all of the cases in the study under study (*6*). Fore example, pediatric patients suffering from

cardiovascular disease cannot be attributed to lifestyle, and often these children have no notable family medical history of early onset cardiovascular disease (*7*). Therefore, screening methods are needed to determine which patients are at risk when family medical history, physical exams and cholesterol screens reveal nothing unusual.

The development of biological indicators, other than cholesterol levels and family medical history, for cardiovascular disease has been underway since the 1960's. In the early 1970's Ryan Huxtable and Rubin Bressler investigated cardiac pathology and noted elevated levels of 2-aminoethanesulfonic acid (taurine) in patients with cardiovascular disease (*8, 9*). This discovery sparked a new field of research in cardiovascular disease. Stemming from this research, taurine has been found to be the most abundant free amino acid in the human body. A normal 70kg human contains 7g of taurine, primarily concentrated in the brain, heart, and liver (*10*). Taurine cannot be utilized for the construction of proteins because it is a sulfonic acid. However, it plays an important role in various other biological processes including processes directly related to cardiovascular disease.

The most well known function of taurine is in the cholesterol excretion pathway. The conjugation of taurine and cholate to form taurocholic acid is the body's primary pathway for excreting cholesterol. From cholesterol, the liver synthesizes two primary bile acid salts, cholate and chenodeoxycholate, which are then excreted as taurine or glycine conjugates into the small intestine. There the taurine and glycine conjugates aid in the digestion and absorption of fats and fat-soluble vitamins. Much of the bile salts will be recycled and passed through the liver several more times each day. The bile salts that are not recycled are metabolized by microorganisms in the large intestine and are excreted, thus providing the only means by which cholesterol is released from the body.

Huxtable's and Bressler's (*9*) work indicates a correlation between the concentration of taurine in cardiac tissue and cardiovascular disease. We have expanded upon their work by focusing our study on taurine metabolism, specifically on how this molecule is catabolized by bacteria and biosynthesized in humans. The loss of taurine in humans is due to excretion rather than catabolism (*11*). Bacteria, on the other hand, have the ability to break down taurine and use it as an alternative sulfur source (*12*). An enzyme known as taurine α-ketoglutarate dependent dioxygenase (TauD) catalyzes the hydroxylation of taurine to aminoacetaldehyde and sulfite in the presence of α-ketoglutarate (α-KG), Fe(II) and oxygen (*13*). As TauD is a mononuclear non-heme iron dioxygenase, the bacterial system can be used in biochemical studies to understand the hydroxylation of taurine and the method of oxygen activation in mononuclear non-heme iron dioxygenases. By studying how taurine is catabolized in bacteria, we can also begin to understand how it is constructed in humans.

The taurine biosynthetic pathway in humans involves two enzymes, cysteine dioxygenase (CDO) and cysteine sulfinic acid decarboxylase (CSAD). This pathway was chosen for our study because, like the taurine catabolic pathway, this particular anabolic pathway also utilizes a mononuclear non-heme iron dioxygenase and can contribute to the field of mononuclear non-heme iron dioxygenase mechanisms and oxygen activation in addition to understanding the taurine biosynthetic process.

Taurine Catabolism by Taurine α-Ketoglutarate Dependent Dioxygenase (TauD). TauD is a sulfur starvation induced (ssi) protein that is found in the tauABCD gene cluster. This cluster is involved in the uptake of alkanesulfonates during times of sulfur starvation (*14*) of the organism *Escherichia coli*. TauD catalyzes the conversion of taurine to sulfite and aminoacetaldehyde coupled to the degradation of α-ketoglutarate (α-KG) to succinate and carbon

dioxide (Figure 1.1) (*13*). The enzyme is promiscuous, accepting 3-N-morpholino propansulfonic acid (MOPS), 1,3-dioxo-2-isoindolineethaesulfonic acid, butanesulfonic acid, pentanesulfonic acid and hexanesulfonic acid as substrates in addition to taurine (*13*).

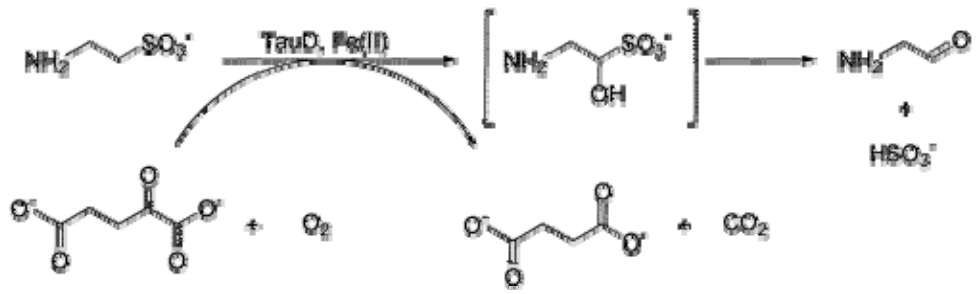

Figure 1.1 : The reaction catalyzed by TauD. The reaction scheme is from Elkins et al (*15*).

The crystal structure of TauD was initially reported by Elkins et al. to a resolution of 3.0Å (R_{cryst} = 28.1%, R_{free}=32.0%) (*15*). The structure revealed a conserved jellyroll motif, and confirmed that H99, D101, and H255 form the 2-his 1-carboxylate facial triad iron coordination motif. The orientation of α-KG could also be ascertained from the observed electron density. α-KG occupies two coordinating positions on the iron, and binds in a bidentate fashion with the C-1 carboxylate and the 2-oxo-group coordinated to the iron. The C-5 carboxylate forms a salt bridge with R266 and is hydrogen bonded to T126. Due to the poor resolution of the structure, the binding of taurine was not clear in the electron density maps. A putative assignment of taurine binding was achieved by empirically positioning the substrate in such a way that it created a logical hydrogen bonding network with TauD.

Stopped-flow kinetic analysis by Ryle et al. (*16*) suggests that α-KG binding is the initial step in the TauD mechanism. It has been proposed that taurine binding occurs rapidly after the binding of α-KG, producing a lilac-colored taurine-α-KG-Fe(II)TauD chromophore that rapidly disappears (*17*). A similar phenomena has been observed with another α-KG-dependent dioxygenase, 2,4-dicholorphenoxyacetic acid dioxygenase (TfdA), where addition of α-KG and Fe(II) produces a weak, pink-colored chromophore that is indicative of the Fe(II)-to- α-KG charge transfer transitions (*18, 19*). In both cases, addition of the substrate and exposure to oxygen results in the rapid decay of the chromophore (*20*). Additional support for this assignment and the working model of the TauD reaction mechanism, was provided by Ho et al. through a resonance Raman study on the chromophores of TauD (*21*).

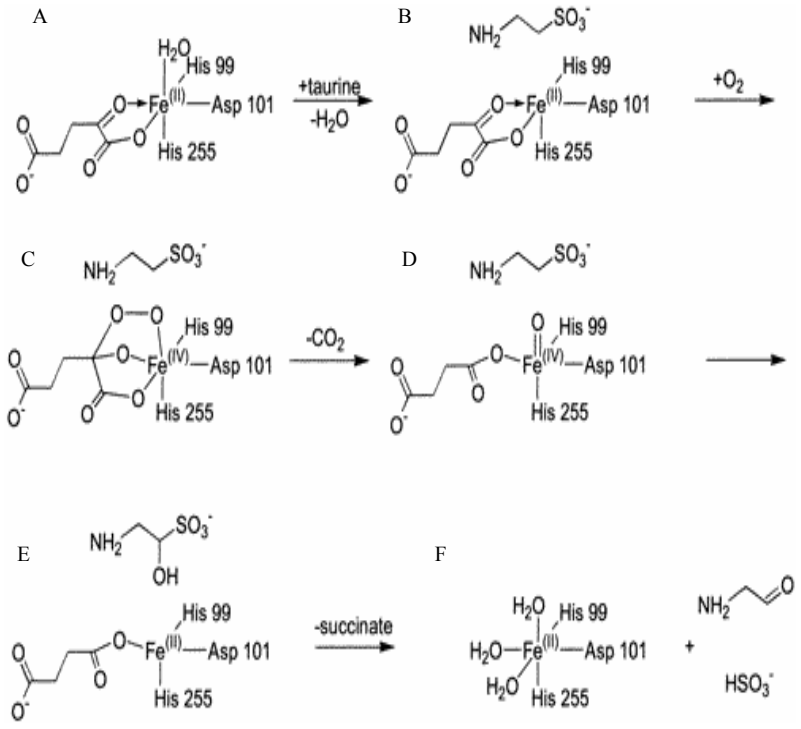

Figure 1.2: Proposed reaction mechanism for TauD. Figure is from Elkins et al. (*15*).

Unfortunately, the published structure does not provide any information on the initial binding of oxygen to TauD. In the heme oxygenase systems, due to steric hindrance, there are fewer modes of binding molecular oxygen due to steric hindrances than in the mononuclear non-heme iron oxygenase systems, where there is more flexibility around the iron coordination site. This allows oxygen to bind in numerous positions in the mononuclear non-heme oxygenase systems. As a result, there are two proposed mechanisms of oxygen activation. The first is the end-on binding mode of oxygen in which the iron directly coordinates one oxygen atom while the other atom is coordinated either by water molecules or side chains in the binding pocket (*22*). However, in this binding mode the oxygen atoms are activated unequally and would produce an oxygen species that is less reactive when compared with the side-on binding mode.

The second binding mode is the side-on binding mode in which both oxygen atoms are coordinated to the iron thus both being equally activated. This appears to be the more probable of the two binding modes. While not a member of the α-KG dependent dioxygenase family, the crystal structure of naphthalene 1,2-dioxygenase (*23*) bound to molecular oxygen bound has shed some light on the mode of oxygen activation in the αKG dependent dioxygenase family. The authors used cryogenic time resolved crystallography to trap molecular oxygen in the active site and were able to solve the resultant crystal structure. The molecular oxygen was found to be bound in a side-on fashion as was originally proposed for this enzyme (*23*) (Figure 1.3).

Figure 1.3: Binding of molecular oxygen to the iron center of naphthalene dioxygenase. Oxygen is shown in red and iron in purple. Figure is from Karlsson et al. (*23*).

The mechanism of oxygen activation by TauD is proposed to occur through a debated Fe(IV)=O intermediate (Figure 1.2 C & D). In particular, due to the reactivity of such a species, evidence for this intermediate has only recently been provided by the work of Price et al. (*24*). Mössbauer and EPR spectroscopy showed that a new intermediate in the TauD reaction had a low isomer shift (0.31 mm/s) and a S=2 ground state and that subsequent cryoreduction yields a high spin Fe(III) (*vide infra*). This information, coupled with the observation that the quadrupole splitting parameters were similar to the high-spin Fe(IV) intermediate of ribonucleotide reductase (*25*), provides strong evidence for the Fe(IV) intermediate (Figure 1.2, Step D) in the α-KG dependent dioxygenase family.

The presence of such a reactive oxygen intermediate is also supported by the long standing observation that TauD undergoes self-inactivation (*17, 20*). If oxygen is added in the absence of any substrate, the lilac chromophore, formed by the addition of taurine and α-KG to TauD, converts to a yellow species (λ_{max} = 408nm) which slowly transforms to a green/brown

chromophore (λ_{max} = 550nm) (*17*). Resonance Raman spectroscopy of the green/brown chromophore revealed a set of vibrations not observed in the lilac chromophore (λ_{max} = 530nm) (*17*). When this data was compared with data available for other iron-binding proteins (*26*) it was clear that these vibrations were not arising from a Fe(III)-O-Tyr or Fe(III)-O-Trp interaction. This suggests that a tyrosine residue has been hydroxylated to 3,4-dihydroxyphenylalanine (DOPA) during the formation of the green/brown chromophore and has inactivated the enzyme. How this inactivation occurs is still unknown, but may hold clues to the enzyme mechanism of TauD.

Ascrobic acid has an important effect on the TauD reaction mechanism. In the absence of ascorbic acid the enzyme will function properly, however the rate of the reaction falls off dramatically with time as the enzyme undergoes self-inactivation (*27*). The precise function of ascorbic acid in the reaction mechanism is unknown although it is apparent that is has a protective effect. It is thought that ascorbate may act to keep the iron in the ferrous state (*16*). However, the rate of ascorbic acid consumption is not stoichiometric with the amount of substrate consumed (*16*). Therefore, the nature of the protective mechanism must be more complicated.

Studies of the TauD Dimerization Site. TauD is known to be active in the dimeric form (*13*) and presently there are two proposed dimerization sites for TauD in the literature. The first was proposed by Elkins et. al based on the crystal packing and interactions observed in their structure (*15*). The authors noted eight hydrogen bonds between the two monomers and a buried surface area of 1770Å. Similar size proteins in their correct dimeric form normally have a buried surface area of 2000Å. The temperature factors (B factors) in the proposed interface region were as expected for residues in the interior of a multimer.

Interestingly, a second dimeric arrangement also exists within this data suggesting that there are two possible sites for dimerization. In the second published structure of TauD the fourth monomer in the asymmetric unit was in a different conformation forcing the data into a different spacegroup (*28*). Because of this shift, two different dimers were observed. Interestingly, when two 6-fold symmetry mates are generated in the Elkins data, the presence of a tetramer becomes apparent. The first is the dimeric arrangement described above, and the second is a dimeric arrangement not detailed in the Elkins publication but observed in the O'Brien structure (*28*). This second dimeric arrangement contains secondary structural elements, a conserved set of amphipathic helices, that have been observed in the literature for other members of the α-ketoglutarate dependent dioxygenase family (*28-30*). The residues at the center of these helices are highly conserved and hydrophobic (*15*). Two residues at the center of these helices, L145 and L221, present a two fold symmetry axis between the two molecules (*28*). The buried surface area that has been calculated for the second dimeric arrangement is 400 $Å^2$ as compared to 550 $Å^2$ calculated in the dimeric arrangement proposed by Elkins et al. (*15, 28*). However, the data suggests that water is completely excluded from the second dimerization site which, according to work by Fernandez et al., provides strong evidence for inter-protein interaction (*28, 31, 32*).

Taurine Biosynthesis by Cysteine Dioxygenase (CDO) and Cysteine Sulfinic Acid Decarboxylase (CSAD). The taurine biosynthetic pathway has been identified and shown to involve two enzymes (*33*). The first-pass metabolism of sulfur amino acids is accomplished in the liver where an increase in dietary sulfur amino acids results in the upregulation of hepatic CDO (*34*). CDO catalyzes the conversion of cysteine to cysteine sulphinic acid (CSA). CSA is then converted to hypotaurine by CSAD (Figure 1.4). The decomposition of hypotaurine to

taurine occurs non-enzymatically (Figure 1.4). It is clear that hepatic taurine synthesis is largely restricted by the low levels of CSA as a substrate for CSAD and, therefore, CDO is presumed to be the rate-limiting step in taurine biosynthesis (*35*).

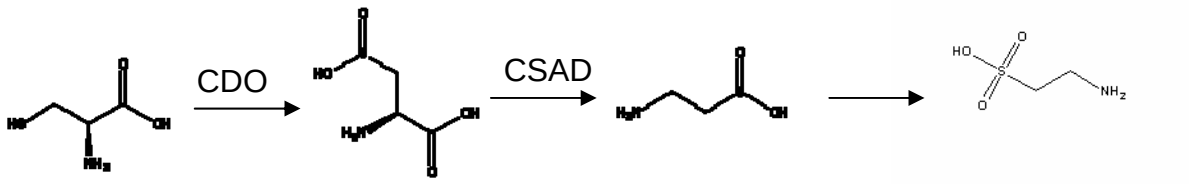

Figure 1.4: Taurine biosynthetic pathway.

CDO Regulation, Structural Studies, and Reaction Mechanism. CDO catalyzes the first reaction of the taurine biosynthetic pathway by performing the oxidation of L-cysteine to cysteine sulfinic acid. The upregulation of cysteine dioxygenase protein levels occur when the cellular cysteine pool begins to rise. The upregulation of CDO occurs on the post-transcriptional level as the levels of mRNA have been shown to remain constant (*35*).

Down-regulation of CDO activity occurs on the post-traslational level. Two isoforms of the enzyme have been reported in the literature, a 25kDa and a 23kDa form (*35-37*). The degradation of this protein into an inactive form suggests the possibility of ubiquitination as a form of regulation, and Dominey et al. have recently published the first evidence of this regulation (*38*). They found that CDO could be stabilized by adding proteosome inhibitor 1 to inhibit ubiquitin-26 S proteosome (*38*) demonstrating that ubiquitinylation is involved in CDO destabilization.

There are several inconsistencies in the literature regarding the biochemistry of CDO. The gene has been successfully cloned from *Homo sapiens* (ACC# D85782.2) encoding what would appear to be a 200 amino acid protein suggesting a molecular weight of 23 kDa. Yamaguch et al. first suggested a molecular mass of 22.5 kDa in the initial characterization of this enzyme (*39*). Later it was suggested that there was also a 68 kDa isoform of the enzyme (*40-45*). The matter was addressed by using antibodies providing evidence for a 23 kDa and 25 kDa form of CDO (*36*).

An additional controversy involves the measured kinetic parameters for CDO. The enzyme's Michaelis constant (K_m) for substrate L-cysteine was first determined to be 5 mM (*39*), whereas it was later reported to be 0.45 mM (*39*). Recently Chai et al. have reported the K_m for L-cysteine to be 2.5 mM (*46*). Simmons et al. report kinetic parameters that agree closely with the original measurement of 0.45 mM (*47*). The large variation in the reported kinetic parameters can be attributed to differences in the respective assay procedures. Two assays exist for CDO, one in which ophthaldialdehyde is used to derivatize the product which is then analyzed by HPLC and the other in which UV absorbance at 215nm is used to measure the product. Table 1.1 outlines the assay procedures and reported data on CDO.

Table 1.1: Comparison of reported kinetics for CDO.

Author	Optimum pH	Cysteine Km	Fe Kd	Assay Procedure	Additives
Yamaguchi (*39*)	9	0.45mM	ND	[35S] cysteine; Dowex 50W with HEMA IEC BioQ DEAE guard cartridge	None
Sakakibara (*42*)	8.5 – 9.5	1.3mM	ND	[35S] cysteine; Dowex 50W with HEMA IEC BioQ DEAE guard cartridge	None
Bagley (*48*)	6.1	1.2mM	ND	[35S] cysteine; Dowex 50W with HEMA IEC BioQ DEAE guard cartridge	hydroxylamine
Chai (*46*)	7.5	2.5mM	ND	215nm detection of CSA	None
Simmons (*47*)	6.1	0.45mM	.09mM	Derivitazation by ophthaldialdehyde (OPA)	Bathocuprione disulfate (BCS)
McCoy (*49*)	7.5	3.4mM – 2.1mM	ND	Reaction quenched by freeze/thaw cycle 215nm detection of CSA	None

The preliminary biochemical characterization of rat CDO clearly suggests that the enzyme is a non-heme iron oxidase (*42*). CDO is thought to contain a mono-nuclear iron site, in the ferrous form, that is required for activity though no spectroscopic or biophysical characterization of the non-heme iron center has confirmed this hypothesis (*50*). Unlike their heme counterparts, the iron site of the non-heme dioxygenases/oxidases does not exhibit the intense spectral features that are characteristic of the porphyrin ligation. While some alternative spectroscopic approaches can be applied, the lack of intense spectral features also emphasizes the importance of crystallographic analysis in uncovering the atomic details responsible for the enzymatic diversity among the non-heme iron oxidases.

The structure of the *Mus musculus* CDO reported by McCoy has been solved to 1.75 Å resolution (*49*). This structure, however, is not physiologically relevant as it contains a metal ion refined as Ni(II). The overall fold of the protein is a small β-barrel structure that is typical of the cupin superfamily (*49*). The metal binding site is reminiscent of the 2-His 1-carboxylate facial triad as observed in TauD with one important distinction. Instead of utilizing a carboxylate CDO utilizes a histidine in its stead producing a 3-H binding motif (H86, H88 & H140) that is accompanied by three water molecules lending to a distorted octahedral nickel coordination site.

A Y157-C-93 adduct was described in the structure and has only been seen once before in the structure of galactose oxidase, a member of the cupin superfamily (*51*). In the proposed reaction mechanism for galactose oxdiase, the adduct allows the reaction to proceed without the formaction of Cu(III). The hydroxyl group of Y157 is located 4.4Å away from the metal in CDO and may serve a similar purpose in its reaction mechanism (*51*).

The reaction mechanism proposed by McCoy et al. (*49*) follows that of aromatic ring extradiol dioxygenases (*52*) and since sulfur is able to access cation-radical intermediates (*53, 54*) the aid of the C-Y adduct may not be needed. A reaction mechanism for CDO has been proposed and proceeds as follows: L-cysteine displaces two water molecules upon coordinating to the iron, oxygen then binds the iron in an end-on fashion (Figure 1.5 C). This produces a suitable environment for an oxygen attack on the sulfate, resulting in a sulfoxy cation. Transfer of the metal-bound oxygen follows to form the product.

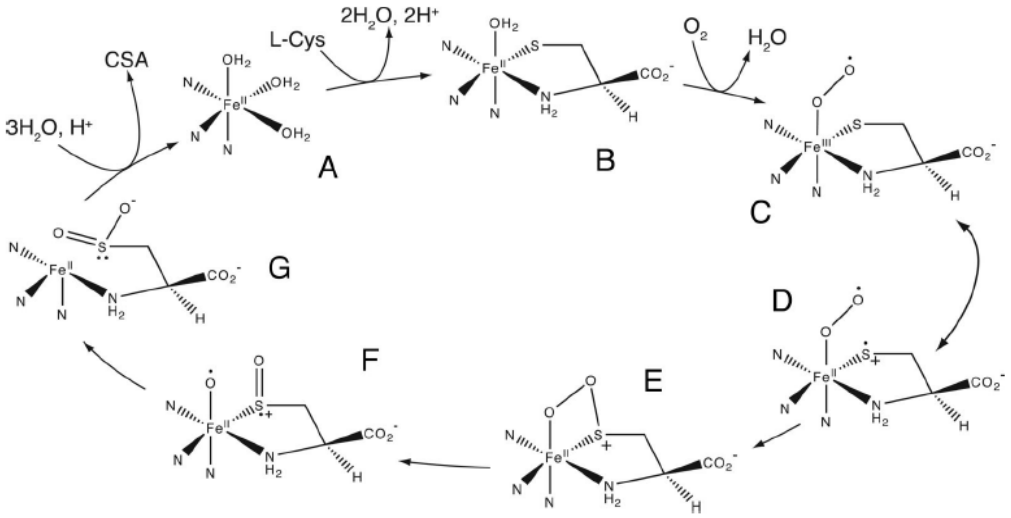

Figure 1.5: McCoy et al. Proposed Reaction Mechanism for Cysteine Dioxygenase. This figure is from McCoy et al. (*49*)

The recently reported structure of the *Rattus norvegicus* CDO solved by Simmons et al. (*55*) agrees with the overall fold of the protein reported by McCoy et al. (*49*). However, as Simmons' structure contains iron rather than the previously reported nickel, the metal coordination geometry is no longer the distorted octahedran reported previously. Instead the iron coordinates tetrahedrally with three histidine residues (H86, H88, and H140) and a highly conserved water molecule (Figure 1.6) (*55*). The iron binding site is well defined, with B factors of 12 $Å^2$, 10 $Å^2$ and 11 – 14 $Å^2$ for the water molecule and coordinating histidines respectively. The binding site defined by this geometry is restrictive, thus the authors propose that oxygen activation occurs in an end-on fashion. The structure suggests that the highly conserved Y157 is involved in coordinating the oxygen in this end-on binding mode.

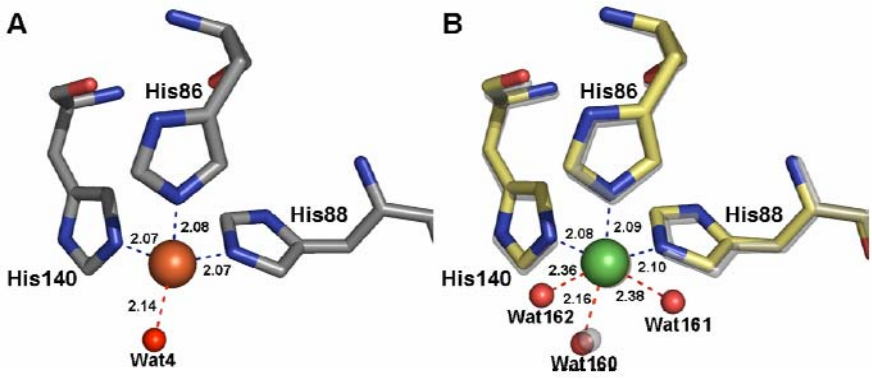

Figure 1.6: Comparison of CDO metal binding site in the rat (A) and mouse (B) crystal structures. This figure is from Simmons et al. (*55*).

Attempts to co-crystallize the protein with substrate or substrate analogs have been unsuccessful (*55*). Protein crystals readily grew however, but when data was collected there was not enough evidence to support substrate binding. The authors (*55*) did observe a tunnel from the iron binding site to the bulk solvent. This tunnel leads to a pocket that is lined with non-polar residues (L75, W77, V142, F161, C164, V177, and M179) which would be suitable for the binding of a single L-cysteine molecule. A mixed disulfide between C164 and substrate molecules disrupted the solvent structure within the binding pocket and changed the conformation of M160 and R60. Due to the inability to bind the substrate the mixed disulfide had to be modeled as a methyl-sulfide and the binding pocket for L-cysteine could not be defined.

CSAD Biocehmical Studies. CSAD catalyzes the second step in the taurine biosynthetic pathway, the decarboxylation of cysteine sulfinic acid, produced by CDO, to hypotaurine. Regulation of CSAD is believed to occur on the transcriptional level. Bella et al. (*35*) studied the

effect of cysteine metabolism in rats fed on varying diets and found that the levels of CSAD mRNA were lower in rats given a high protein diet (*35*). Purification of CSAD has been achieved from bovine brain and rat liver (*56-60*). The brain and the liver forms of the enzyme are encoded by mRNA that differ in the 5' untranslated regions. This suggests that the CSAD mRNAs are differentially regulated in these two tissues.

Substrates for CSAD include cysteine sulfinic acid, pyrodoxal phosphate, and, to a much lesser degree, cysteic acid (*61*). Other sulfonyl containing amino acids are not substrates. The protein has a NPHK sequence, a putative pyridoxal phosphate binding site that is also found in several other amino acid decarboxylases. The reaction mechanism of CSAD is still unknown.

Research Goals. Due to the importance of taurine to normal cardiovascular health, we have undertaken a biochemical and biophysical investigation of the enzymes involved in taurine metabolism (TauD, CDO, and CSAD). The goals of this study are:

1 Provide additional structural and kinetic data for TauD.

2 Recombinantly express human CDO and address the matters of molecular mass and optimal pH.

CDO and TauD are both mononuclear non-heme iron dioxygenase, therefore this study also contributes to the fields of oxygen activation and reaction mechanisms for members of the non-heme iron dioxygenase family. An understanding of the kinetics and functional biological structure of these enzymes provides medical researchers with a baseline upon which a search for mutations in CDO and CSAD that may be useful in identifying and treating cardiovascular disease may be performed.

CHAPTER 2

SUBSTRATE-INDUCED CONFORMATIONAL CHANGES IN *ESCHERICHIA COLI* TAURINE/α-KETOGLUTARATE DIOXYGENASE AND INSIGHT INTO THE OLIGOMERIC STRUCTURE.[1]

[1] O'Brien, J, R, Schuller, DJ, Yang, VS, Dillard, BD, and Lanzilotta, WN. 2003 Biochemistry. 42:5547-5554. Reprinted here with permission of publisher.

Abstract

The enzymes in the α-ketoglutarate (αKG) dependent dioxygenase superfamily represent the largest class of non-heme iron oxidases and have important medical, ecological, and biotechnological roles. One such enzyme, taurine/α-ketoglutarate dioxygenase (TauD), catalyzes the conversion of 2-aminoethanesulfonate (taurine) to sulfite and aminoacetaldehyde while decomposing αKG to succinate and CO_2. This αKG dependent dioxygenase is expressed in *Escherichia coli* under sulfur starvation conditions and allows the cell to utilize taurine, and other similar sulfonates in the environment, as an alternative sulfur source. In this work, we report the structures of the apo and holo forms of TauD to 1.9 Å resolution (R_{cryst} = 21.2%, R_{free} = 24.9%) and 2.5 Å resolution (R_{cryst} = 22.5%, R_{free} = 27.8 %) respectively. The models reported herein provide significant new insight into the substrate orientations at the active site and the conformational changes that are induced upon taurine binding. Furthermore, analysis of our crystallographic data coupled with re-analysis of the crystallographic model (resolution = 3.0 Å, R_{cryst} = 28.1, R_{free} = 32.0) presented by Elkins et al. (*Biochemistry* 1999, 38, 15278-15286) reveals an alternative oligomeric arrangement for the enzyme that is consistent with the conserved primary and secondary structure elements of other αKG dependent dioxygenases.

Introduction

Oxygenases and oxidases that contain iron as the sole cofactor constitute a large family of redox enzymes that utilize either di-iron centers or mono-iron centers for catalysis. Mono-nuclear iron enzymes can be iron(III)-dependent, such as lipoxygenase or the intradiol cleaving catechol dioxygenase, and iron(II)-dependent, such as the extradiol cleaving catechol dioxygenases and the αKG dependent oxygenases. The members of the latter family, the αKG dependent oxygenases, represent the largest class of non-heme iron oxidases. These enzymes

require ferrous iron and αKG as a co-substrate to catalyze biological oxidations of otherwise inactivated carbon-hydrogen bonds (*52, 62, 63*). Like their heme counterparts, these oxygenases perform a variety of essential reactions, including hydroxylations, desaturations, and oxidative ring closures in the biosynthesis of antibiotics, modified amino acids/peptides, carnitine, and signaling molecules. Notably, in higher organisms these enzymes catalyze the post-translational hydroxylation of proline residues during collagen biosynthesis and in the hypoxia inducible factor (HIF)-mediated pathway for dioxygen sensing (*64-66*).

Sequence analysis of many members within αKG dependent superfamily has revealed a conserved His-X-Asp-$X_{(\sim 55)}$-His sequence (*67*) that forms a 2-His-carboylate facial triad at the active site of the enzyme (*68-70*). The three amino acids that form the 2-His-carboylate facial triad represent the only protein ligands to a bound iron atom. Previous studies have shown that in the absence of αKG and substrate, solvent molecules occupy the remaining three coordination sites of the six coordinate iron (*68, 71*). Two of these solvent ligands are displaced when αKG binds to the ferrous center through its C-1 carboxylate oxygen and C-2 carbonyl oxygen (*68, 71*).

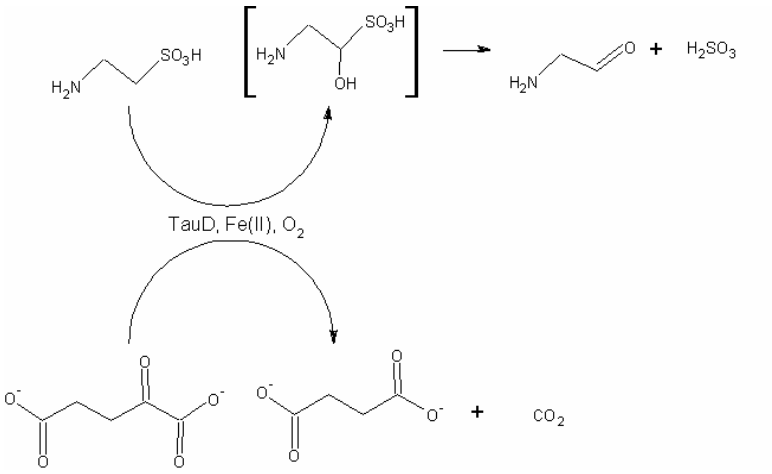

Figure 2.1: Overall reaction catalyzed by Escherichia coli taurine//α-ketoglutarate dioxygenase (TauD).

The overall reaction catalyzed by *Escherichia coli* taurine/α-ketoglutarate dioxygenase (TauD) is summarized in figure 2.1. TauD is a member of the group II αKG dependent dioxygenases and a crystal structure of TauD with Fe(III), taurine, and αKG bound at the active site has been reported to 3.0 Å resolution (R_{cryst} = 28.1%, R_{free} = 32.0%)(*15*). Although the exact orientations of the substrates could not be determined, it was clear from this work that the 2-His-carboylate facial triad architecture was preserved in TauD. Similar to what has been proposed for other αKG dependent dioxygenases, this structural work combined with spectroscopic measurements suggested that the binding of taurine displaces a solvent molecule resulting in a shift from 6-coordinate to 5-coordinate geometry at the active site and creation of an oxygen binding site (*15, 16, 72-74*). In the absence of taurine, oxygen can still react with the αKG- and Fe(II)-bound form of TauD and leads to hydroxylation at tryptophan residues 128, 240 and 248

(*20*). At the present time, the mechanism by which these modifications are made and their physiological role remains unclear.

Biochemical analysis suggests that TauD functions as a dimer (*13*) and a model for the physiological dimer was proposed based on the crystallographic data (*15*). Interestingly, the proposed oligomeric structure was inconsistent with some of the conserved primary and secondary structural elements found throughout the group II αKG-dependent dioxygenase family of iron dioxygenases. Specifically, all group II αKG-dependent dioxygenase, that have been structurally characterized thus far, contain a "jelly roll" motif around the active site as well as a pair of structurally conserved α-helices (*15*). For TauD and clavaminate synthase (CAS) these are helices 4 and 6. The amino acids that make up these helices are composed of primarily conserved hydrophobic residues and when the jelly roll motifs are aligned, these two helices also align. A similar pair of structurally conserved helices has also been observed in the structure of proline 3-hydroxylase and has been proposed to be the site of dimer formation (*29*).

In order to provide further insight into the substrate interactions and the oligomeric nature of TauD, we have determined the crystal structure of apo and holo TauD to 1.9 Å and 2.5 Å resolutions, respectively. In both cases, the substrate taurine was an absolute requirement for crystallization. The apo structure therefore represents crystals prepared in the presence of taurine, while the holo structure contains taurine, αKG, and ferrous iron bound at the active site. Due to the crystallographic symmetry, four copies of the TauD monomer are present in the asymmetric unit. In both the holo and the apo structures, three of the monomers contain taurine bound at the active site, while the fourth does not. Significant conformational changes are observed in the monomer without taurine bound. In both the apo and the holo data the exact orientations of the substrate molecule can be observed in the monomers with taurine bound.

Taken together, the data presented here provides an improved understanding of the substrate interactions within the active site and insight into the conformational changes associated with substrate binding. In addition, evidence for an alternative dimeric arrangement for TauD is presented that is consistent with the conserved sequence and structural elements found in other αKG-dependent enzymes.

Experimental Procedures

Enzyme Expression, Purification, and Assay.- The *tauD* gene was cloned from the *E. coli* genome by using traditional PCR methods. The oligos used for amplification of the *tauD* gene were designed such that the correct flanking sequences required for subcloning into the pET-TOPO™ vectors (Invitrogen Corporation, Carlsbad CA) were incorporated in the PCR product. Using these vectors, gene expression is under the control of the T7 RNA polymerase/promoter system (*75*). One advantage to using this approach is that a 6x His tail can be attached at either the N- or C-terminus for rapid isolation. For optimal expression levels using this system, 5 mL of cells (0.6-1.0 OD_{600nm}) harboring the expression plasmid were inoculated into 1 L of TB media and grown for 12 hours at 34°C. The cells were then induced by addition of IPTG (1 mM final) and grown an additional 4 hours before being harvested. Purification of the tagged TauD was performed in a single step using chelating sepharose (Amersham Pharmacia, Piscataway NJ) as previously described (*76*), except that salt was omitted from all elution buffers. Typical yields were between 40-60 mg of pure protein from 1 L of cell material. To ensure that no cationic metals were contaminating the purified protein, purified enzyme was exchanged from 10 mM TRIS buffer pH 8.0 into 100 mM EDTA pH 8.0 and back into the original buffer. The reconstituted enzyme had a specific activity of 14.0 ± 1.5 μmoles min^{-1} (mg TauD)$^{-1}$ at pH 7.0 when assayed by the method described by Eichorn et al (*13*).

Crystallization. Crystals were grown by batch method in an anaerobic environment of 95% nitrogen and 5% hydrogen gas. The temperature of the glove box was 22°C and the oxygen level was maintained below 1 ppm at all times. All solutions for crystallization were made anaerobic by several rounds of vacuum degassing and flushing with argon using a vacuum manifold. The protein to precipitant ratio was 2:3 in the batch experiment and the initial protein concentration was 30 mg/mL. Due to protein requirement for crystallization and the ease of purification, we decided to use TauD that was tagged at the N-terminus. Diffraction quality crystals typically appeared within two weeks after the conditions were set up. Crystals were grown using the TauD protein as isolated and a precipitating solution consisting of 0.2 M di-sodium tartrate dihydrate, 5 mM taurine and 20 % PEG 3350. These crystals represent the apo crystals used in this study. Interestingly, once these crystals were formed, they could be transferred to mother liquor containing 0.1 M TRIS-HCL pH 7.5 in place of di-sodium tartrate dehydrate. This could only be accomplished if iron ($FeSO_4$), ascorbate, and αKG were also present at 10 mM. These crystals represent the holo crystals used in this study. In order to prepare both the apo and holo crystals for freezing, the PEG 3350 was replaced with PEG 4000 and the concentration was increased to 30% in 2.5 % increments.

Data collection and structure determination. Data for the apo-crystals was collected at the Cornell High Energy Sychrotron Source (CHESS) on beam line A1. Data were collected for the iron-treated crystals at the University of Georgia on a Rigaku RU-200 rotating anode equipped with Osmic focusing mirrors and a R-axisIIc image plate detector. Data were processed using MOSFLM (*77*) and the CCP4 suite of programs (*78*). Our initial approach was to treat data collected from the apo crystal as native data and data from the iron-treated crystal as a derivative. Anomalous difference Patterson methods indicated four weak signals in the

asymmetric unit. Prior to applying for beam time to collect MAD data, a 3.0 Å structure (R_{cryst} = 28.1 % R_{free} = 32.0 %) of TauD was reported by Elkins et al. (*15*). Using a poly-alanine version of the momomer reported by Elkins et al., without any substrate or cofactors, molecular replacement was performed using CNS (*79*) and gave a good solution that was consistent with four monomers in the asymmetric unit. Subsequent rounds of model building and refinement was performed using the programs O (*80*) and CNS (*79*). Protein atoms were placed prior to modeling the water molecules and no atoms or molecules were treated as rigid groups during refinement. All Figures were generated using the programs Molscript (*81*), Xtalview (*82*), and Raster3D (*83*).

Results and Discussion

Crystal growth and Structure Refinement. Consistent with Elkins et al (*15*), we observed that the exposure of iron-treated crystals to oxygen had a negative impact on diffraction (exposure during crystallization or post-crystallization). Therefore, prior to freezing the crystals for data collection, sodium dithionite was added to a final concentration of 2 mM. A crystal, in mother liquor containing dithionite, was then removed from the glove box and immediately frozen. We found that taurine, αKG, and molecules with structure similar to αKG (e.g. tartrate or citrate) could also be used in crystal growth. In addition, once crystals were formed, they could be transferred to mother liquor in which the original buffer had been replaced with a variety of buffers including 50 mM imidazole (pH 6.5-7.0), 50 mM HEPES pH 7.5, or 50 mM TRIS (pH 7.5-8.5). The only essential requirement for preserving the crystal quality appeared to be taurine and the precipitant (Either PEG 3350 or PEG 4000). Taken together, these data indicate that conformational changes are occurring in TauD upon substrate binding or during turnover.

The structure of TauD with 5 mM taurine bound (apo TauD) is described to 1.9 Å resolution. In addition, the structure of TauD with αKG, Fe(II), and taurine bound (holo TauD) is described to 2.5 Å resolution. In both cases, the crystals belong to the space group $P2_12_12_1$ with similar unit cell dimensions (See Table 2.1). There are four monomers in the asymmetric unit. In all models, the main chain atoms were clearly visible in their respective $2F_o$-F_c composite omit maps for residues 3 through 282. The 6x poly-histidine tail was not observed in any of the models. All side chain atoms could be placed with the exception of residues 168-170 in the D monomer. In this case, the residues were modeled as alanine.

The overall fold of the TauD monomer forms an α-β structure with a "jelly-roll" motif at the core of the structure. Despite the low sequence identity (7.7%) the TauD monomer shows a high degree of structural similarity to clavaminic synthase (CAS) (*84*). Forty-eight percent of the backbone atoms align with a rms deviation of 1.64 Å (*15*). Not too surprisingly, the region of greatest alignment is centered on the jelly roll motif, where the active site typically resides in this superfamily. Interestingly, the other region of structural alignment is found around helices 3, 4, 6.

Table 2.1 Data and Refinement Statistics. [a]$R_{sym} = \sum_{hkl}[\sum_{I}(|I_{hkl,I} - \langle I_{hkl}\rangle|)] / \sum_{hkl,I}\langle I_{hkl}\rangle$, where I_{hkl} is the intensity of an individual measurement of the reflection with indices hkl and $\langle I_{hkl}\rangle$ is the mean intensity of that reflection. [b]The numbers in parenthesis refer to the outer resolution bin used in data processing. [c]The numbers refer to cell edges a, b, and c respectively. The space group is $P2_12_12_1$ so all angles are 90°.

Data Statistics				
Data Set	λ (Å)	Resolution Range	Completeness (%)	R_{sym} [a]
Apo	0.97	50.0-1.90 (1.97-1.90)[b]	96.9(78.9)	0.049(0.334)
Holo	1.54	38.0-2.50 (2.59-2.50)	99.9(98.4)	0.059(0.431)

Refinement Statistics		
Model	Apo	Holo
Unit Cell[c]	95.8 117.8 118.2	92.5 118.8 118.8
Protein/Cofactor Nonhydrogen Atoms	6810	6850
Solvent Molecules	706	353
Resolution Range (Å)	50.0-1.9	36.5-2.5
Total Refections	99459	46047
Total Reflections Used in R_{free}	5034	2321
R_{cryst}	0.212	0.225
R_{free}	0.254	0.278
R.m.s. Deviations		
Bond Distances (Å)	0.005	0.006
Bond Angles (°)	1.34	1.35
B-Factors (Å^2)		
Average	34.5	47.9
Minimum	13.3	10.8
Maximum	76.9	98.0

In both the apo and the holo models, substrate is bound in only three of the four monomers. If taurine was modeled in the active site of the fourth monomer, both R_{cryst} and R_{free} would increase by 5.0 %. Due to this observation and the high B-factors of residues covering the active site, we suspected that the fourth monomer might be in multiple conformations due to

partial occupancy of taurine in the active site. To test this we modeled monomer D with taurine having an occupancy of 0.5 and the residues that cover the active site in both an "open" and a "closed" conformation. Like taurine, each of the conformations was assigned an occupancy of 0.5. This approach resulted in an increase in both R_{cryst} and R_{free} by ~2.5 %. These results indicate that the structure that best fits the data is a crystallographic tetramer with three monomers in a closed conformation while the fourth monomer is in an open conformation. The nature of the open and closed conformations is discussed in the sections that follow.

Interestingly, if all four monomers were in the closed conformation, then the crystallographic tetramer would exhibit 222 symmetry. Considering that two of the unit cell edges are very similar (See Table 2.1) and the observation that all angles are 90°, this explains why our crystals were initially indexed in a tetragonal space group. However, integration and scaling of the data using *P4* resulted in unreasonable values for R_{sym}. Subsequently, all data reduction and refinement was performed with the orthorhombic space group *P222*. Further analysis of the systematic absences suggested that the correct space group was $P2_12_12_1$. This space group was used for all subsequent refinement for both the apo and the holo data and the same set of reflections were used in both cases for calculating R_{cryst} and R_{free}. The structure reported by Elkins et al. described crystals belonging to the hexagonal space group $P6_222$ with unit cell dimensions a = b =117, c = 202 Å. This was interesting due to the similar cell dimensions observed for our crystals (See Table 2.1) and would suggest a small movement of the unique monomer as well as the tetramer itself could generate two six fold symmetry axes. Consistent with this proposal is the observation of a tetramer in the data presented by Elkins et al. that is similar to the tetramer found in our asymmetric unit (RMSD for α-carbon atoms < 2.0 Å). This is shown in Figure 2.2. Figure 2.2 was generated using the coordinates deposited by

Elkins et al. (PDB ID 1GQW) and shows the position of this tetramer relative to their asymmetric unit. The observation of this tetramer in the hexagonal unit cell as well as in our asymmetric unit is significant because it suggests two possible arrangements for the biological dimer.

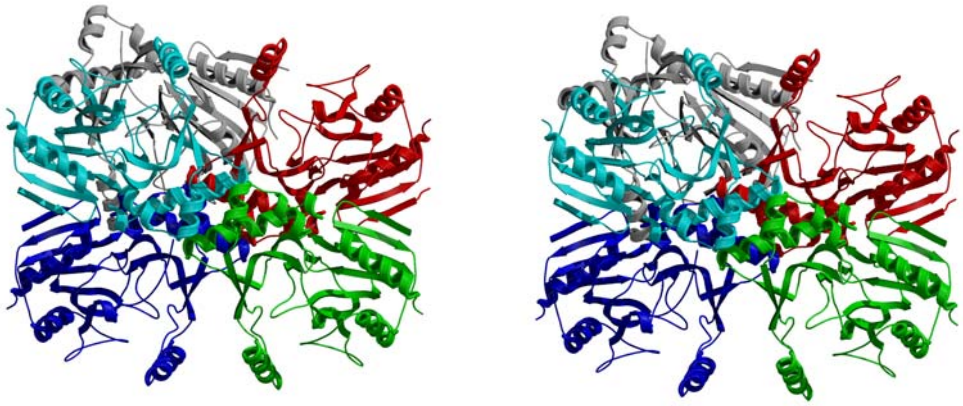

Figure 2.2: Cross-eyed stereo view showing the identification of the asymmetric unit for space group *$P2_12_12_1$* relative to the asymmetric unit (space group *$P6_222$*) presented by Elkins et al. (*13*). The monomers colored in grey and red represent the asymmetric unit for the space group *$P6_222$* and the monomers colored in red, green, cyan, and blue represent the asymmetric unit found in the space group *$P2_12_12_1$*. The grey and red monomers are taken from the coordinates submitted by Elkins et al. (PDB ID=1GQW) and the green, cyan, and blue monomers were generated using the appropriate symmetry operators for the space group *$P6_222$*. The biological dimer proposed by Elkins et al. is represented by the red and green monomers.

Dimerization Interface. The proposed oligomeric arrangement that we feel best explains the dimeric nature of the enzyme (*13*) is shown in Figure 2.2 and is in contrast to what was previously proposed (*15*). The reason we favor this dimer arrangement stems from several lines of evidence. First, the two helices that form the dimer interface are conserved elements of secondary structure in other αKG-dependent iron dioxygenases (*29, 30*). Moreover, many of the residues that are found in the two helices at the dimer interface are highly conserved and hydrophobic in nature (*15*). Specifically, four highly amphipathic helices (two from each subunit) form a hydrophobic core that extends back into the hydrophobic center of each monomer. The tertiary structure that comprises our proposed dimer interface can therefore be described as an anti-parallel four helical bundle. At the center of the dimer interface are residues L145 and L221 (two from each subunit). An axis of symmetry can be drawn between the two equivalent pairs and a two fold rotation of one monomer around this axis will generate the other. Using the CCP4 program AREAMOL, the calculated buried surface for this dimer is ~400 $Å^2$, compared with ~550 $Å^2$ for the alternative dimer arrangement (*15*). The hydrophobic core at the dimer interface is then covered by the helical structure itself and by a number of hydrogen bonds between the monomers at either end of the helices. It is very clear, especially in the 1.9 Å data, that water is completely excluded from this interface. Extensive work by Fernandez et al. (*31*) indicates that there is significant hydrogen bond desolvation observed in the interior of proteins. These authors have extended their analysis to show that insufficiently dehydrated hydrogen bonds are good determinants of protein interactions (*32*). Consistent with their findings, is the observation that formation of our dimer interface allows these hydrophobic helices to dehydrate the hydrogen bonds that are central to the stability of the secondary structure. There are no such interactions at the dimer interface proposed by Elkins et al. (*15*). Additional support for this

dimer arrangement comes from the observed B-factors for atoms at our proposed dimer interface. The B-factor can be described as a measurement of thermal motion. In Figure 2.3 it is clear that the region of the model with the lowest thermal motion is centered on our proposed dimer interface. Finally, a dimeric arrangement that is similar to our proposed dimer has been observed in the high resolution structures of other αKG-dependent dioxygenases (*29*).

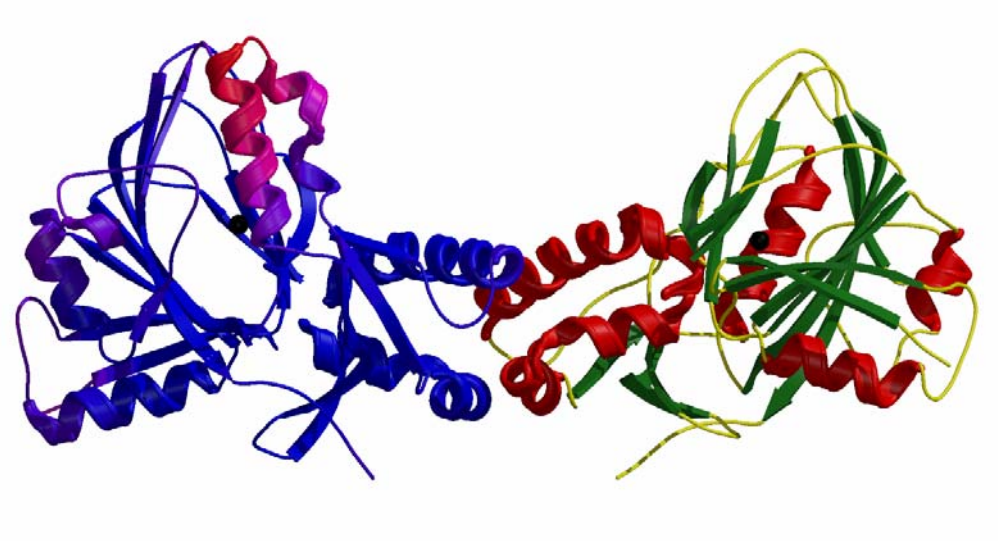

Figure 2.3: Diagram of the proposed biological dimer for TauD. Side view of the proposed biological dimer for TauD. The monomer on the left is colored by B-factor with the color ranging from blue to red and corresponds to a B-factor range from 10 to 95 $Å^2$. The monomer on the right is colored by secondary structure with helices shown in red, strands in green, and coils are colored yellow. The Figure was generated using the coordinates from monomer A (right side) and monomer D (left side) of the Holo model.

Taurine Binding in the Apo Enzyme. The binding mode of taurine, prior to the addition of σKG and iron, is presented in Figure 2.4. The focus of Figure 2.4 is on the amine group of taurine. This is primarily because the taurine sulfonate interactions are similar to what was reported by Elkins et al. Briefly, this includes residues R270, H70, and the backbone N-H of V102. Interactions of the taurine sulfonate clearly support the hypothesis presented by these authors to explain the selection of substrates with tetrahedral sulfonate moieties by TauD.

The taurine amine group is bound in a distorted tetrahedral environment with hydrogen bonds to two water molecules and the amide oxygen of N97. The phenolic oxygen of Y73 extends across the substrate from a loop region to come within 3.0 Å of the carboxylate oxygens of D94. While the phenolic oxygen of Y73 is 3.7 Å away from amine group of the taurine, the interaction of Y73 with D94 positions the phenolic ring within 3.5 Å of the C2 carbon of taurine. This interaction essentially locks the substrate in place, with two additional phenyl side chains (F159 and F206) approaching the C1 and C2 carbon atoms of the taurine molecule from the other side.

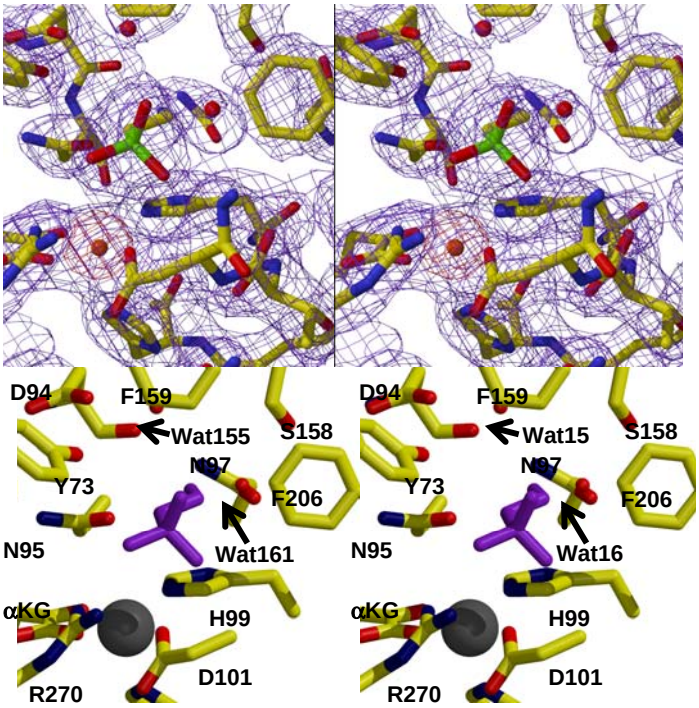

Figure 2.4: Cross-eyed stereo view of the 2Fo-Fc composite omit map contoured at 1σ and the corresponding model of the active site region for apo TauD with taurine bound.

Top Panel: Stereoview of the substrate taurine and key residues in the active site of the apo model and the corresponding 2Fo-Fc composite omit map (Purple Cage). The composite omit map was generated using the simulated annealing protocol with 5 % of the asymmetric unit being omitted at a time. *Bottom Panel:* Stereoview taken from the same perspective as the top panel with taurine colored purple. For additional clarity, the amino acid D94 and the side chain atoms of F159, Y73, N97, N95, S158, H99, F206, as well as D101 are labeled. Five water atoms atoms are also shown, and unless stated otherwise, all atoms are shown in stick format using cpk colors.

Taurine-Induced Conformational Changes. For both the apo and the holo data sets, the crystallographic tetramer was best modeled with three monomers in a closed conformation with substrate bound. As previously mentioned, the position of taurine was not well defined in the electron density of the fourth monomer. Given that the position of taurine was not well defined in the forth monomer (Monomer D) for both the apo and holo data sets, as well as the lack of any strong interactions between this region of monomer D and other molecules in the unit cell, some conclusions can be made regarding conformational changes that may be due to taurine binding alone. The overall secondary structure for the open conformation is compared with that of the closed conformation in Figure 2.5. Two regions of secondary structure are seen to undergo dramatic movement. One possible explanation for the conformational change from the closed to the open state is the loss of several key interactions with the substrate taurine. In the region of random coil (See Figure 2.5) both H70 and Y73 have key interactions with taurine. On the opposite side of taurine, is a hydrophobic cluster of side chains including F159, Y164, and W174. In the absence of taurine, the hydrogen bonding network in the active site is lost. This action releases the residues 60-80 (random coil region) and exposes a significant portion of these hydrophobic side chains in the region of the helix (See highlighted in helix in Figure 2.5). Conformation adjustments must be made to shelter these hydrophobic side chains from solvent. This is accomplished in part by a shift from α-helical structure to that of a $^{3\text{-}10}$helix in the region from T167 to W174.

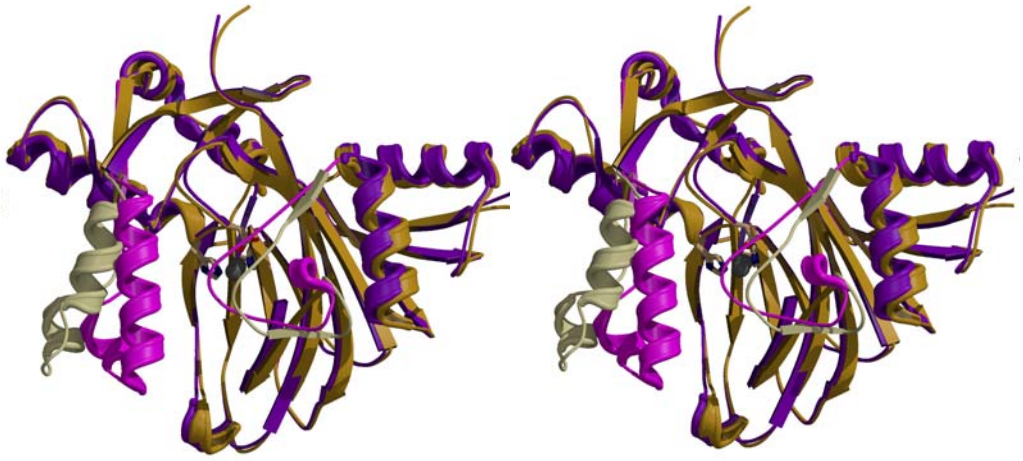

Figure 2.5: Cross-eyed stereo view depicting the difference between the "open" and the "closed" conformation of TauD. Overlay of the C and D monomers from the apo model. Monomer C is shown in purple while monomer D is shown in brown. Regions of movement are highlighted by a lighter shade. The difference in the relative position of helix 5 is further highlighted by the black arrow.

Iron and αKG Binding. The binding of iron and αKG has minimal impact on the orientation of the substrate taurine (Figure 2.6). However, the hydrogen bonding network around the taurine amine group is in contrast to what was reported by Elkins et al. (*15*). In our model the hydroxyl oxygen of S158 is positioned 2.6 Å from the backbone carbonyl oxygen of G205 and 2.5 Å from a bound water molecule that is interacting with the amine group of taurine. This water molecule is 2.5 Å away from the amide oxygen of N97. It is clear in both apo and holo data sets that water bound near the amino group of taurine provides a short solvent channel to the bulk solvent. Residues in the active site that have hydrogen bond interactions with the ordered water include D94, N97, and S158. Consistent with what was proposed by Elkins et al., we

clearly see a pentacoordinate ferrous iron bound to the protein through a 2-His-carboxylate facial triad consisting of residues H99, D101, and H255. The αKG molecule is bound to the iron in a bidentate manner with the opposite carboxylate end being locked into place by the strictly conserved residues R266 (salt bridge) and T126 (hydrogen bond). In the holo structures, all four monomers in the asymmetric unit show good density for the iron and αKG molecules. As was observed in the apo data, the density in the active site of the fourth monomer (D monomer) was consistent with no taurine being bound. In addition, the B-factors for atoms in the active site region of the D monomer were also higher than average suggesting greater thermal motion in the absence of substrate.

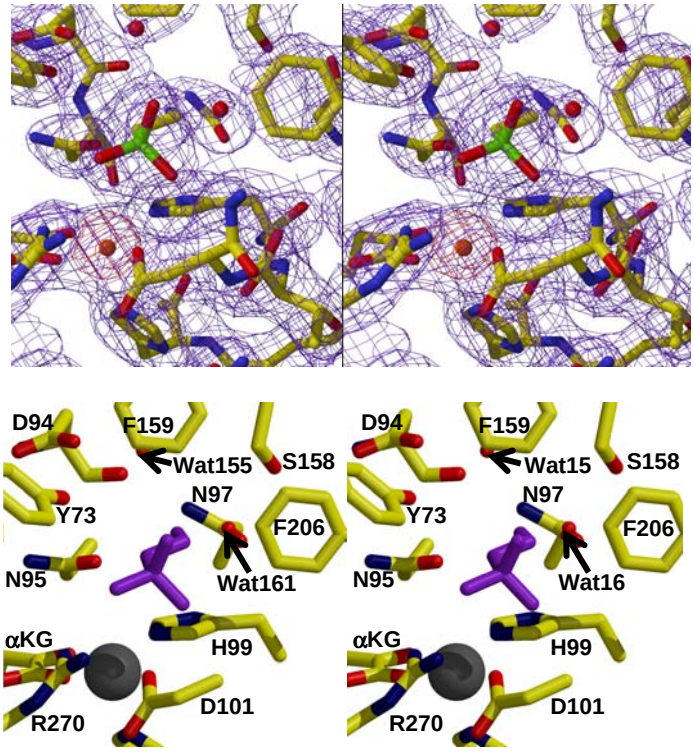

Figure 2.6: Cross-eyed stereo view of the 2Fo-Fc composite omit map contour at 1σ and corresponding model of the active site region for holo TauD with taurine bound. *Top*

Panel: Stereoview of the substrate taurine and key residues in the active site of the holo model and the corresponding 2Fo-Fc composite omit map (purple cage). The 2Fo-Fc composite omit map was generated using the simulated annealing protocol with 5 % of the asymmetric unit being omitted at a time. The anomalous difference map is also shown contoured at 12 σ around the iron position (red cage). *Bottom Panel:* Stereoview taken from a similar perspective with taurine colored purple. For additional clarity, the amino acid D94 and the side chain atoms of F159, S158, Y73, N97, F206, N95, H99, R270 as well as D101 are labeled. The iron atom is colored black, αKG is labeled, and two water atoms are also shown. Unless stated otherwise, all atoms are shown in stick format using cpk colors.

Implications for the Mechanism of TauD. All αKG dependent dioxygenase family members demonstrate a common jellyroll fold with three iron ligands found in a common 2-His-carboxylate facial triad motif. While the source of the carboxylate ligand can be either an aspartate or glutamate residue, the relative positions of the histidine and carboxylate side chain further divide these enzymes into three distinct groups. Both TauD and CAS belong to the group II αKG-dioxygenase subfamily due to the additional 138-207 residues that separate the carboxylate ligand from the second histidate ligand (*85*). While TauD and CAS demonstrate low sequence identity (<10 %), a remarkable 48 % of the protein backbone atoms align with an RMS deviation of 1.64 Å (*15*). In fact, if the residues that form the 2-His-carboxylate facial triad in both enzymes are explicitly aligned (residues 144, 145, and 146 in the CAS model, and residues 99, 100, and 101 in the TauD model), then two regions of high homology are immediately clear. The first region of homology is the β-strand section, or jelly-roll motif, that is important for the active site structure. The second region of alignment is centered on the two helices that we propose are involved in dimer formation.

Given the complex regulation of metal ions in the cytoplasm (*86*), the relative binding constants for substrates (*16*), and the absolute requirement for taurine and molecules with structures similar to αKG for crystal growth, a reasonable hypothesis is that *in vivo* activity of TauD may ultimately be limited by the availability of iron. Furthermore, given the similar positions of taurine αKG and the 2-His-carboxylate facial triad of TauD, compared to the equivalent positions in CAS, it is reasonable to predict that the oxygen binding mode between these two group II αKG-dependent dioxygenases is similar (*84*). In this mode, oxygen binds opposite the second histidine residue and is positioned for insertion into the appropriate C-H bond of αKG. However, recent work using nitric oxide (NO) as an oxygen analog has shown the potential for Fe-O rearrangements (*30*) at the active site in CAS. At the present time this has not been tested in TauD. In any case, decomposition of αKG has been proposed to occur prior to hydroxylation of taurine (*15, 16*). It has further been suggested that the hydroxylation reaction will proceed through an Fe(IV)=O intermediate in this family of enzymes (*84, 87, 88*). At the present time there is no crystallographic information for these intermediate states or the product-bound form of TauD.

One final point regarding the structural information presented here must be incorporated into the current model for the mechanism of TauD and the observed properties of the enzyme. In addition to the conformation changes that are revealed in this work, we have also observed significant differences in the positions of side chains near the amino group of the substrate taurine. In contrast to what was reported previously, the hydroxyl oxygen of S158 and the phenolic oxygen of Y73 are not interacting directly with taurine. Alternatively, the hydroxyl side chain of S158 is hydrogen bonded to the carboxylate oxygen of a back residue and the phenolic oxygen of Y73 is hydrogen bonding with the carboxylate side chain of D94. It is

interesting to note that, despite poor sequence and structural alignment in the variable loop regions covering the active site, a similar interaction observed in CAS between Y149 and D202. While the functional role of these interactions remains to be investigated, the observation of a short solvent channel to the bulk solvent is significant. Specifically, expulsion of the water atoms allows additional room to accommodate larger sufonates, consistant with the activity of TauD (*13*).

CHAPTER 3

CLONING, EXPRESSION, PURIFICATION, AND CHARACTERIZATION OF RECOMBINANT HUMAN CYSTEINE DIOXYGENASE.

In the early 1970's Ryan Huxtable and Rubin Bressler published an investigation of cardiac pathology reporting elevated levels of 2-aminoethanesulfonic acid (taurine) (*8, 9*). This discovery sparked a new field of research in cardiovascular disease. Taurine is now known to be the most abundant free amino acid in the human body. Lack of taurine during development and growth can lead to cardiomyopathy, retinal degeneration, and growth retardation (*11*).

Humans express taurine biosynthetic enzymes in the brain and liver (*10, 89*). Cysteine dioxygenase (CDO) catalyzes the oxidation of L-cysteine to cysteine sulfinic acid. Cysteine sulfinic acid decarboxylase (CSAD) decarboxylates cysteine sulfinic acid to hypotaurine which is subsequently auto-oxidized to taurine (Figure 3.1).

CDO has been successfully cloned from human (ACC# D85782.2) encoding a 200 amino acid protein suggesting a molecular weight of 23 kDa if the entire gene is translated (*44*). Several isoforms of the enzyme were thought to exist (*40, 48*) however, analysis using antibodies provides evidence for a 23 kDa form and a 25 kDa form (*36*). The isolation of the distinct forms of the enzyme suggests degredation as a form of regulation. Dominey et al. have recently published the first evidence of ubiquitination in the role of regulating cysteine dioxygenase and its effects on the cysteine pool (*38*). CDO could be stabilized by adding proteosome inhibitor 1 to inhibit ubiquitin-26 S proteosome.

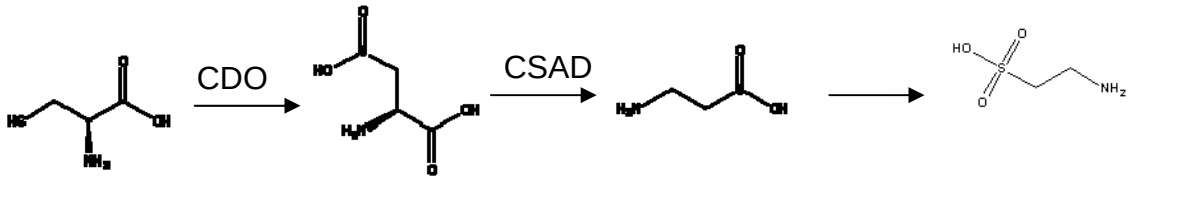

FIGURE 3.1: The taurine biosynthetic pathway.

Two crystal structures of CDO have been solved, the structure of mouse CDO by McCoy et al. (*49*) and rat CDO by Simmons et al. (*55*). Both structures agree on an overall fold that is similar to the cupin family of proteins. The differences in these structures occur in the metal binding motif. In the mouse structure the metal refined as nickel (II) and the metal coordination site consisted of three conserved histidine residues (H86, H88, H140) and three water molecules in a distorted octahedral environment (*49*). In the rat structure the metal refined as an iron (II) and while the same three histidines were involved in metal coordination, only one water molecule was observed conferring a more universally accepted tetrahedral coordination geometry (*55*). Neither structure was able to reveal the binding mode of L-cysteine, however a putative binding pocket was identified in the Simmons et al. structure (*55*).

Biochemical analysis has also failed to provide insight into substrate binding due to the inability to reproduce the data reported in the literature. Specifically, the measured kinetic parameters for the rat and mouse forms of CDO have been debated in the literature. Yamaguchi et al. initially measured a K_m for L-cysteine of 5 mM (*39*) whereas it was reported later as 0.45 mM (*39*). Recently Chai et al. have reported the K_m for L-cysteine to be 2.5 mM (*46*).

Simmons et al. report kinetic parameters that agree closely with the original measurement of 0.45mM (*47*).

Due to the inconsistencies in the literature a careful biochemical study of human cysteine dioxygenase was necessary. We present the cloning, expression, purification, and properties of recombinant human cysteine dioxygenase.

Experimental Procedures

PCR and Cloning. The human CDO gene was PCR amplified from the human liver genome library using the following primers:

5' GTT CTT GAA CAG ACC GAA GTG CTG AAG CC 3'

3' GTT CTT CAC CTT AGT TGT TCT CCA GCG AGC CCG 5'

NheI and HindIII sites were stitched onto the gene for cloning into pTrcHisA (Invitrogen). The gene and vector were parallel double digested with NheI and HindIII for two hours at 37ºC. The cut vector was then electrophoresed on a 1% agarose gel. The digested vector and insert were purified together using a Qiagen gel extraction kit. The ligation reaction was performed by adding 5µL of the cleaned insert/vector into a mixture of 12 µL H_20, 2 µL 10X T4 buffer, 1 µL T4 DNA ligase and incubating overnight at 16ºC. 5 µL of the ligation reaction was transformed into the chemically competent TOP10 (Invitrogen) *Esherichia coli* cell line. After a one hour recovery the culture was plated onto a carbenicillin resistant plate and incubated at 34ºC overnight. Individual colonies were grown in 5 mL LB cultures for plasmid purification (Qiagen) and sent off for sequencing.

Expression and Purification. 1 µL of miniprepped DNA is transformed into the 10 µL JM109 (Invitrogen) *Esherichia coli* cell line and, after the addition of 250 µL LB media, is allowed to recover for one hour shaking at 250 rpm at 34ºC. The culture is plated onto a

carbenicillin resistant plate and incubated overnight at 34ºC. A 25 mL culture containing carbenicillin is incubated at 34ºC shaking at 250 rpm overnight. 1 L TB cultures containing carbenicillin are inoculated with 1 mL of the overnight culture and are incubated at 34ºC shaking at 200 rpm for 20 hours. The cells are harvested by centrifugation and resuspended in 50 mM $NaPO_4$, 100 mM NaCl pH 7.5 buffer before being frozen at -80ºC.

The cell suspension is thawed at 34ºC while shaking at 250 rpm for 30 minutes. DNAse, lysozyme, and PMSF are added and the culture is incubated for a further 15 minutes at 34ºC shaking at 250 rpm. The cells are broken by French pressure cell at 1,200 psi. The cell suspension is cleared of debris a by ultracentrifugation (16,000 rpm for 45 minutes). The supernatant is then applied to a 5 mL bed volume of a Talon (Clonetech) resin gravity flow column that has been pre-equilibrated with 50 mM $NaPO_4$, 100 mM NaCl pH 7.5. The resin is washed with 50 mL 50 mM $NaPO_4$, 100 mM NaCl pH 7.5 before eluting with 15 mL 50 mM $NaPO_4$, 100 mM NaCl, 150 mM imidazole pH 7.5.

The protein is concentrated by Millipore ultracell centrifugal concentrator (5,000 Da MWCO) at 3,000 rcf. The protein is then buffer exchanged into 100 mM $NaPO_4$ pH 7.5. Buffer exchange is accomplished by concentrating the sample down to 1 mL and refilling the concentrator to 15 mL with buffer. This step is repeated three times. The samples are then analyzed by SDS PAGE using the Laemmli method (*90*). Protein concentrations are calculated using the Bradford assay (Bio-Rad) (*91*). Protein not immediately used in experimentation is flash frozen in liquid nitrogen and stored at -80ºC.

Determination of Bound Iron. Inductively coupled plasma-emission spectrometry (ICP) was used to analyze bound iron in the purified CDO. Protein was purified from two sources, an iron-spiked growth and a normal TB growth. The iron spiked growth contained 1 mL 10 mg mL^{-}

[1] ferrous sulfate in the 1 L 20 hour growth and expression phase. Protein was purified as reported, concentrated, and buffer exchanged into chelex (Bio-Rad Analytical Grade Chelex Resin 100-200 mesh, sodium form) treated buffer. 1 ppm, 5 ppm, and 10 ppm samples were made assuming a 100% bound 1:1 iron:protein binding ration. Control samples were also made with ferrous sulfate to correspond to the 1 ppm, 5 ppm, and 10 ppm experimental samples. Samples were treated with 2 M HCl, to precipitate the protein, and spun in a tabletop centrifuge at 14,000 rpm. The samples were transferred to clean tubes and sent to the University of Georgia chemical analysis facility to be analyzed by ICP.

Enzymatic Assay. The enzymatic assay is adapted from Chai et al. (*46*). Due to the controversy in the literature about the optimal pH for the enzyme, it was first necessary to conduct trials to find the optimal pH. The reported optimal pH for cysteine dioxygenase ranges from 6.5 to 9.5. Therefore a buffer must be found within each of these ranges that have the least absorption at 208 nm.

The assay mixture includes 1 mM L-cysteine, 1 mM ferrous sulfate, 100 mM $NaPO_4$ pH 7.5. The reaction is initialized by the addition of 3 μM cysteine dioxygenase. The reaction is incubated at 30ºC shaking at 250 rpm for 10 minutes. The reaction is terminated by the addition of 1 mL 1 M heptaflurobutyric acid (HFBA) followed by a 3 minute boil. The mixtures are then refrigerated for 3 minutes to bring them back down to room temperature before being centrifuged at 14,000 rpm. The supernatant is then transferred to a clean microfuge tube.

Separation of the reaction product from the substrate and oxidized degredative produces was achieved by ion paired reversed phase HPLC using a particil ODS-3 C18 column (Alltech). Isocratic analysis of a 100 μL injected sample was achieved using a mobile phase consisting of 58:42 acetonitrile:deionized water (v/v) with 0.3% HFBA at a flow rate of 0.5 mL per minute

and detection at 208 nm. The peak eluting at six minutes was confirmed to be cysteine sulfinic acid (CSA) by retention time comparison with standards.

Results and Discussion

Expression and Purification. Recombinant cysteine dioxygenase was expressed at 20 mg L^{-1} and contained only minor larger molecular weight contaminants that could be polished away using an S200 size exclusion column. The purified protein required 100 mM NaCl in the buffer for stabilization. CDO is stable in solution at 4ºC for several weeks, a notable difference from the unstable rat and mouse CDO (*39, 40, 49*). While the mouse and rat CDO share 100% sequence similarity, they share 92% sequence similarity with human CDO. This is significant similarity however; the difference in the stability of these proteins emphasizes the importance of studying the human CDO for future biomedical hypotheses.

The SDS PAGE analysis of the purified recombinant CDO showed a repeatable double banding pattern (figure 3.2). The two bands coincide with the 23 kDa and 25 kDa forms that have been previously reported in the literature (*36*). Bacteria do not have the ubiquitin pathway to which this degradation has been attributed. Therefore, another form of degradation is occurring. It is probable that the degradation is occurring at the carboxy terminal end of the protein since the metal affinity column purifies the protein based on the amino terminal 6X his tag. A look into textbook biology reveals a possible explanation of this observation. The stability of proteins in mammals is dependent on the amino terminal sequence whereas the stability of proteins in bacteria is dependent on the carboxy terminal sequence. In bacteria, the presence of histidines on the carboxy terminus is a signal for degradation and one such histidine (H182) can be found on the carboxy terminus of CDO.

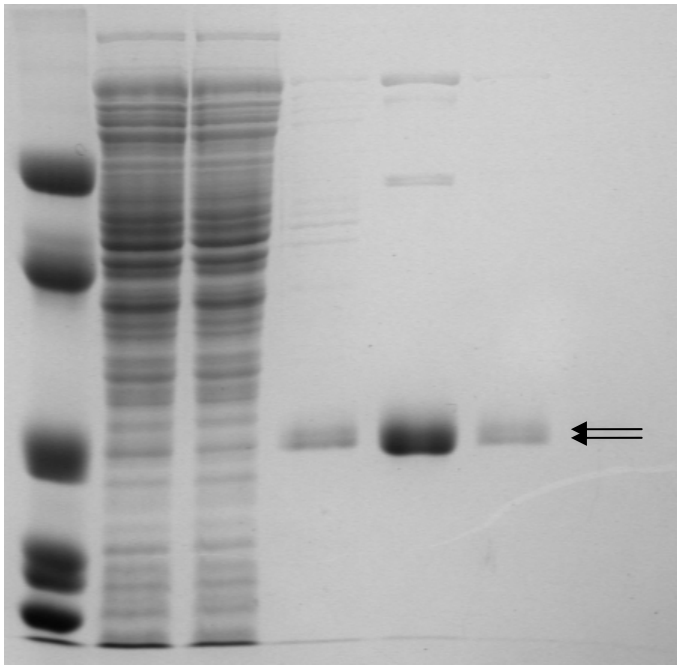

Figure 3.2 SDS PAGE analysis of recombinant human CDO purification.

The hypothesis we have put forth is completed by considering the metabolic consequences of overexpressing CDO in *E. coli.* In this case, CDO is overexpressed providing an excess of cysteine sulfinic acid within the cell. In the absence of overexpressed cysteine sulfinic acid decarboxylase, the excess cysteine sulfinic acid is transaminated to β-sulfinylpyruvate which is then rapidly degraded to pyruvate and sulfite. The pyruvate enters the TCA cycle providing an excess of ATP. The ATP dependent proteases are activated and presumably providing the degradation that we see in our expression of recombinant CDO.

Determination of Bound Iron. ICP analysis data suggests that recombinant human CDO is purified with 13% iron bound (data not shown). This percentage could not be increased by

adding iron to the growth media, changing temperature, or changing the expression line. The only reported dissociation constant (K_d) for iron is 90 μM (*47*). The binding site, according to published structures (*49, 55*), involves a three conserved histidine coordination motif and a water molecule in a tetrahedral geometry. This coordination environment would lead to a very transiently bound iron, although this coordination environment has been observed in 2,2',3-trihydroxybiphenyl dioxygenase, another mononuclear non-heme iron dioxygenase (*92*). The iron determination data supports the model of a transiently bound iron, however it suggests that the iron is more transiently bound than the previously published K_d alludes.

Enzymatic Assay. Adjustments to the enzymatic assay were necessary for optimal performance. The mobile phase was adjusted to a 58:42 acetonitrile:water (v/v) mix containing 0.3% HFBA after an initial test run indicated that the water:methanol mixture previously published (*46*) was interfering with the detection of product. The next adjustment was the detection wavelength. An experiment was conducted by adding cysteine sulfinic acid to the mobile phase mixture and measuring the absorbance in a spectrum from 230 nm to 200 nm. The peak absorbance was observed at 208 nm, not the 215 nm reported in the literature (*46*). To ensure that no interference from L-cysteine was occurring, a duplicate experiment was performed using L-cysteine in the mobile phase. At 208 nm, no interference from L-cysteine could be detected (figure 3.3). The third adjustment was an analysis of all of the buffers (Tris, HEPES and ammonium acetate) reported in previous studies (*46*). It was determined that sodium phosphate buffer provided the least background interference at the detection wavelength.

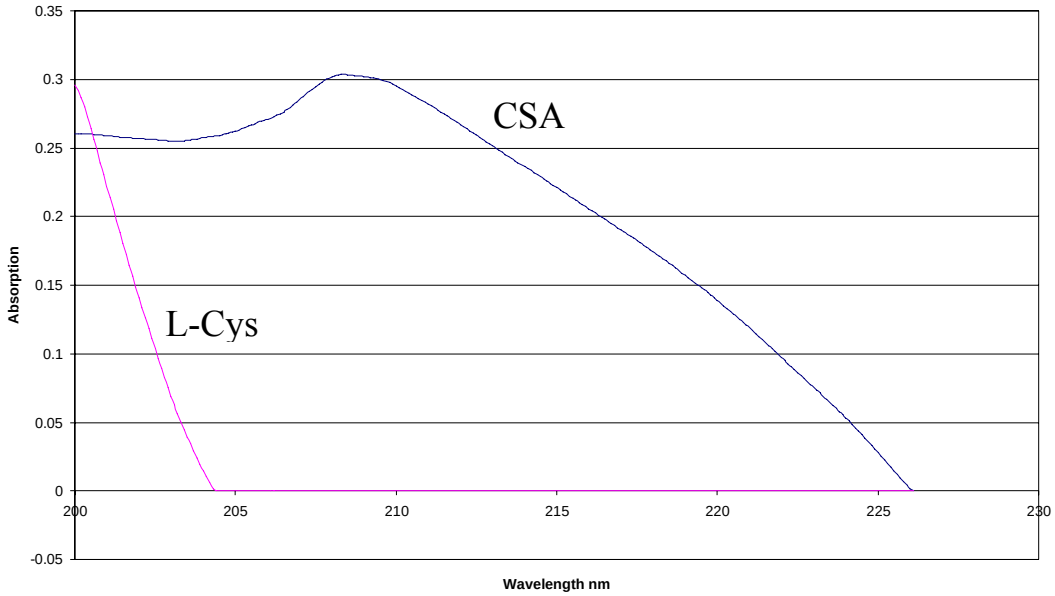

Figure 3.3: Substrate and product absorbance versus wavelength.

Once the buffer was chosen, a pH profile was in order to determine the optimal pH to run the enzymatic assay. The pH profile indicated optimums at both pH 6.5 and 7.5 (figure 3.4). It has been previously reported that when an assay contains a mixture of two isoforms, two pH optimums are observed (*93*). This does address an issue in Stipanuk's publication (*36*) that reports that the 23 kDa form of the enzyme is inactive. The pH profile suggests that the 23 kDa form is most likely active however; it is active at a different pH.

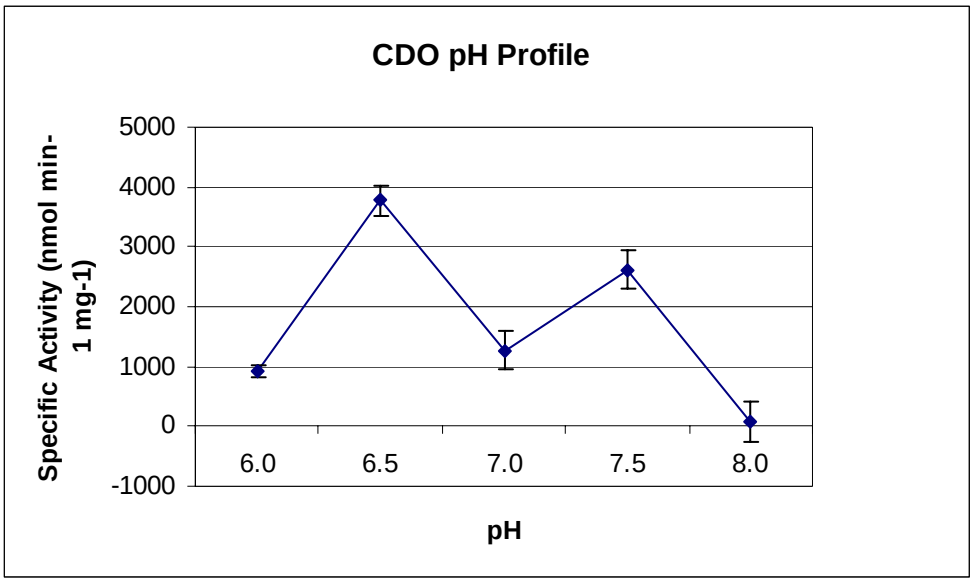

Figure 3.4: Recombinant CDO pH profile.

An explanation of the phenomenon observed in Figure 3.4 that is consistent with our ATP dependant carboxy terminal degradation hypothesis, is that the binding pocket for L-cysteine is exposed in the 23 kDa isoform. This can be further expanded by considering the available structural information. According to structures solved by McCoy et al. (*49, 55*) the L-cysteine binding pocket involves a portion of the carboxy terminal portion of the protein. The pH becomes much more important in binding substrate in the exposed binding pocket than it would in a closed pocket due to the lack of a self-created microenvironment in the full length protein.

Conclusions

Recombinant human CDO was expressed and purified in potentially two active forms. These are a 23 kDa form and 25 kDa form. While we propose that both forms are active, the pH optimum for each isoform differs. In order to properly address the enzymatic activities of each,

the isoforms must be purified and independently assayed. Due to the nature of our hypothesis it is likely that calculation of K_m for L-cysteine will result in an average unless the two forms are separated.

In order to isolate the two distinct forms, two methods are being employed. The first approach involves the addition of an ATP dependent protease inhibitor cocktail to the purification process in order to eliminate the possibility of degradation by ATP dependent proteases and potentially eliminate the 23 kDa form. Another option is to alternatively tag the protein with a 6X his-tag on the amino terminus and a maltose binding protein tag on the carboxy terminus. The expressed enzyme would be applied to a metal affinity column to purify both variants. The elution fraction would be applied to a maltose resin, the flow through would contain only the 25 kDa form and the elution fraction would contain only the 23 kDa form. Thus the variants could be independently studied.

Further characterization of both isoforms of human CDO is in order. Enzymatic rates and binding specificities will need to be calculated for each of the reported variants. Structural work on this enzyme is important in elucidating the enzymatic mechanism since spectroscopic work cannot lend this information to the field. Since an iron (II) metal coordination sphere is silent in spectroscopy, structural work is the only means by which the metal binding site can be probed. While the structures of the mouse and rat forms of CDO have provided some much needed information, the structure of human CDO could provide information to the stability and regulation of this enzyme.

CHAPTER 4

CONCLUSION

Both the α-ketoglutarate dependent dioxygenase family and the taurine biosynthetic pathway are central to many biomedical studies, therefore a firm foundation in atomic structure and biochemical kinetics is necessary to further expand mechanistic studies. The work presented in this dissertation contributes to this field through the structural and biochemical studies presented herein. This work provides new mechanistic insight into two important dioxygenase enzymes as discussed below.

TauD Structural Studies. The structure studies of TauD not only elucidated the active site of the enzyme, but also provided a basis for mutagenic studies that explore the iron coordinating residues in terms of functionality and the method of oxygen activation. Members of the α-ketoglutarate dependent dioxygenase family retain the 2-his 1-carboxylate facial triad coordination sphere for iron, however the choice of the carboxylate is not trivial. When the aspartic acid coordinating the iron is mutated to glutamic acid the enzymatic activity is silenced. Since the change in carboxylate inactivated the enzyme it is likely that the choice of carboxylate preserves the active site geometry and fold.

Structural studies of TauD have also implicated the importance of quaternary structure in the functionality of the α-ketoglutarate dependent dioxygenase family. Two possible dimeric organizations have been presented in the literature (*28*). The presence of the two-fold axis of symmetry and a hydrophobic core that extends from the dimerization site to the active site along

with evidence that water is completely excluded from the dimerization site presented in this work all suggest that the dimer that we have observed is the correct dimer.

Self-Inactivation of TauD and CDO. A number of chromophores have been noted for TauD that are reported to coincide with metal to ligand exchange and enzyme inactivation. This inactivation was initially thought to be irreversible, however treatment with ascrobic acid was found to recover enzymatic activity (*16*). Therefore, it is likely that the inactivation resulted from the oxidation of the ferrous iron and upon reduction of the iron the enzyme was able to recover activity.

Because dioxygenases activate molecular oxygen either in the presence or absence of substrate, it is not unlikely that the enzyme would react in the absence of substrate. Evidence of reactivity in the absence of substrate is reported in all of the CDO structures (*49, 55, 94*). The C-Y adduct is the result of the enzyme turning over in the absence of substrate in the crystallization conditions. Alternatively, it has also been suggested that this C-Y adduct is involved in the reaction mechanism and is utilized in the stabilization of reaction intermediates (*49, 55, 94*). However, until the substrate binding site is located and the residues involved in substrate binding are evaluated, the potential role of this adduct in catalysis can be speculated at best.

The presence of the C-Y adducts in the structures of CDO and the reported self-inactivation of TauD raise important questions regarding regulation of dioxygenases. While the inactivation of TauD is reversible, a mechanism for the degredation of CDO was thought to be a form of irreversible inactivation (*36*). During the purification of recombinant human CDO, two variants were noted, a 23 kDa form and a 25 kDa form. Bacteria possess ATP dependent proteases and since the overexpression of CDO would lead to an abundance of CSA, it is likely that CSA is shunted down an alternate pathway to form pyruvate. The excess pyruvate entering

the TCA cycle will result in an abundance of ATP that would activate ATP dependent proteases that in turn cleave proteins on the carboxy terminal end readying them for destruction. In CDO the hypothesized cleavage site is H183. This region of the protein is associated with the hypothesized binding pocket of L-cysteine (*55*). Therefore cleavage of this region may destroy the binding pocket thereby inactivating the enzyme.

The Importance of Studying Human CDO. The cloning and expression of recombinant human CDO is new research. Previously the only work done on this pathway was in rat and mouse enzymes. Successful expression of this enzyme highlighted the factors necessary in stabilizing the enzyme, predominantly the requirement for sodium chloride in the purification buffer. Interestingly, the human enzyme tends to be much more stable than the rat or the mouse enzyme. The proteins previously purified were concentrated to levels of only 10 mg ml^{-1}. The human enzyme can be concentrated to well over 200 mg ml^{-1} with sodium chloride present in the buffer at a concentration of 100mM.

The iron coordination site that is reported in the literature highly suggests an end-on method of oxygen activation (*55*). It is also noted in the literature that copper can be substituted for iron however the substitution markedly lowers the level of activity (*49*). This may make it possible to conduct structural studies to trap oxygen in the active site to elucidate the method of oxygen activation.

The Future of Dioxygenase Research and Cardiovascular Disease. Two areas of research work hand in hand to understand the underlying causes of cardiovascular disease, iron metabolism and mononuclear non-heme iron dioxygenases. The study of mononuclear non-heme iron dioxygenases is inherently dependent on iron metabolism and a falter in one will ripple an effect to the other.

It is known that family medical history is one of the most powerful tools in medicine to predict cardiovascular disease. This information given to a biochemist sends us searching for genetic markers. The cardiac pathology study by Huxtable et al. (*8, 9*) implicating altered taurine concentrations in cardiomyopathy lead us to look for this genetic predisposition in one of two areas, taurine biosynthesis or taurine recovery. Loss of taurine through urine has not been associated with any patients with notable family medical history of cardiovascular disease (*89*). With this information in hand the search is limited to taurine biosynthesis in the heart, brain, and liver. Normally mouse and rat models are sufficient in the study of proteins associated with human disease. However since there are significant differences between the stability of these models and the human model, it is important to study the human protein to look for the subtle nuances that may affect protein activity and over time lead to cardiovascular disease.

Many on-going studies are investigating the expression of CDO and CSAD. The studies are finding that diet is the main regulator of CDO expression (*34, 43, 95-99*). Therefore CDO activity, or lack thereof, in patients genetically predisposed to cardiovascular disease must be explored. Isolation of CDO from cardiac tissue in patients that have suffered from cardiovascular disease will provide information on possible mutations in CDO that may serve as genetic markers.

Errors in iron insertion into CDO may render the successfully expressed CDO inactive. Purifying CDO from cardiac tissue and measuring the concentration of iron bound would be difficult in lieu of the transient environment of the metal coordination sphere. It would be logical to assume that if iron insertion was the primary problem that other iron binding enzymes with tighter coordination would suffer as well. Evalutaion of cardiac tissue will be the only

means by which we will be able to reject the hypothesis of CDO mutations as genetic markers in cardiovascular disease.

REFERENCES

1. Administration, U. S. F. a. D. (2002), US Department of Health and Human Services.
2. Thom, T., Haase, N., Rosamond, W., Howard, V. J., Rumsfeld, J., Manolio, T., Zheng, Z.-J., Flegal, K., O'Donnell, C., Kittner, S., Lloyd-Jones, D., Goff, D. C., Jr, Hong, Y., Members of the Statistics Committee and Stroke Statistics Subcommittee, Adams, R., Friday, G., Furie, K., Gorelick, P., Kissela, B., Marler, J., Meigs, J., Roger, V., Sidney, S., Sorlie, P., Steinberger, J., Wasserthiel-Smoller, S., Wilson, M., and Wolf, P. (2006) *Circulation 113*, e85-151.
3. Sabatine, M. S., Seidman, J. G., and Seidman, C. E. (2006) *Circulation 113*, e450-455.
4. Association, A. H. (2006), American Heart Association.
5. Superko, H. R. (1996).
6. Kuller, L. H. (2006) *Circulation 113*, 598-600.
7. Association, A. H. (2006).
8. Huxtable, R., and Bressler, R. (1974) *Science 184*, 1187-8.
9. Huxtable, R., and Bressler, R. (1974) *Life Sci 14*, 1353-9.
10. Huxtable, R. J. (1992) *Physiol Rev 72*, 101-63.
11. Schuller-Levis, G. B., and Park, E. (2003) *FEMS Microbiol Lett 226*, 195-202.
12. Kertesz, M. A. (2000) *FEMS Microbiol Rev 24*, 135-75.
13. Eichhorn, E., van der Ploeg, J. R., Kertesz, M. A., and Leisinger, T. (1997) *J Biol Chem 272*, 23031-6.

14. van der Ploeg, J. R., Weiss, M. A., Saller, E., Nashimoto, H., Saito, N., Kertesz, M. A., and Leisinger, T. (1996) *J Bacteriol 178*, 5438-46.

15. 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) *Biochemistry 41*, 5185-92.

16. Ryle, M. J., Padmakumar, R., and Hausinger, R. P. (1999) *Biochemistry 38*, 15278-86.

17. Ryle, M. J., Liu, A., Muthukumaran, R. B., Ho, R. Y., Koehntop, K. D., McCracken, J., Que, L., Jr., and Hausinger, R. P. (2003) *Biochemistry 42*, 1854-62.

18. Hogan, D. A., Smith, S. R., Saari, E. A., McCracken, J., and Hausinger, R. P. (2000) *J Biol Chem 275*, 12400-9.

19. Pavel, E. G. Z., J.; Busby, R W; Gunsior, M; Townsend, C A; Solomon E I. (1998) *J Am Chem Soc 120*, 743-753.

20. Liu, A., Ho, R. Y., Que, L., Jr., Ryle, M. J., Phinney, B. S., and Hausinger, R. P. (2001) *J Am Chem Soc 123*, 5126-7.

21. Ho, R. Y., Mehn, M. P., Hegg, E. L., Liu, A., Ryle, M. J., Hausinger, R. P., and Que, L., Jr. (2001) *J Am Chem Soc 123*, 5022-9.

22. Kovacs, J. A. (2003) *Science 299*, 1024-5.

23. Karlsson, A., Parales, J. V., Parales, R. E., Gibson, D. T., Eklund, H., and Ramaswamy, S. (2003) *Science 299*, 1039-42.

24. Price, J. C., Barr, E. W., Tirupati, B., Bollinger, J. M., Jr., and Krebs, C. (2003) *Biochemistry 42*, 7497-508.

25. Sturgeon, B. E. B., D.; Chen, S.; Huynh, B.-H.; Edmondson, D. E.; Stubbe, J.; Hoffman, B. M. (1996) *J Am Chem Soc 118*, 7551-7551.

26. Siu, D. C., Orville, A. M., Lipscomb, J. D., Ohlendorf, D. H., and Que, L., Jr. (1992) *Biochemistry 31*, 10443-8.

27. Saari, R. E., and Hausinger, R. P. (1998) *Biochemistry 37*, 3035-42.

28. O'Brien, J. R., Schuller, D. J., Yang, V. S., Dillard, B. D., and Lanzilotta, W. N. (2003) *Biochemistry 42*, 5547-54.

29. Clifton, I. J., Hsueh, L. C., Baldwin, J. E., Harlos, K., and Schofield, C. J. (2001) *Eur J Biochem 268*, 6625-36.

30. Zhang, Z., Ren, J., Harlos, K., McKinnon, C. H., Clifton, I. J., and Schofield, C. J. (2002) *FEBS Lett 517*, 7-12.

31. Fernandez, A., Sosnick, T. R., and Colubri, A. (2002) *J Mol Biol 321*, 659-75.

32. Fernandez, A., and Scheraga, H. A. (2003) *Proc Natl Acad Sci U S A 100*, 113-8.

33. Tsuboyama, N., Hosokawa, Y., Totani, M., Oka, J., Matsumoto, A., Koide, T., and Kodama, H. (1996) *Gene 181*, 161-5.

34. Cresenzi, C. L., Lee, J. I., and Stipanuk, M. H. (2003) *J Nutr 133*, 2697-702.

35. Bella, D. L., Hirschberger, L. L., Hosokawa, Y., and Stipanuk, M. H. (1999) *Am J Physiol 276*, E326-35.

36. Stipanuk, M. H., Londono, M., Hirschberger, L. L., Hickey, C., Thiel, D. J., and Wang, L. (2004) *Amino Acids 26*, 99-106.

37. Bella, D. L., Hahn, C., and Stipanuk, M. H. (1999) *Am J Physiol 277*, E144-53.

38. Dominy, J. E., Jr., Hirschberger, L. L., Coloso, R. M., and Stipanuk, M. H. (2006) *Biochem J 394*, 267-73.

39. Yamaguchi, K., and Hosokawa, Y. (1987) *Methods Enzymol 143*, 395-403.

40. Yamaguchi, K., Hosokawa, Y., Kohashi, N., Kori, Y., Sakakibara, S., and Ueda, I. (1978) *J Biochem (Tokyo) 83*, 479-91.

41. Sakakibara, S., Yamaguchi, K., Ueda, I., and Sakamoto, Y. (1973) *Biochem Biophys Res Commun 52*, 1093-9.

42. Sakakibara, S., Yamaguchi, K., Hosokawa, Y., Kohashi, N., and Ueda, I. (1976) *Biochim Biophys Acta 422*, 273-9.

43. Parsons, R. B., Barber, P. C., Waring, R. H., Williams, A. C., and Ramsden, D. B. (1998) *Neurosci Lett 248*, 101-4.

44. McCann, K. P., Akbari, M. T., Williams, A. C., and Ramsden, D. B. (1994) *Biochim Biophys Acta 1209*, 107-10.

45. Ramsden, D. B., Kapadi, A., Fitch, N. J., Farmer, M. J., Bennett, P., and Williams, A. C. (1997) *Mol Pathol 50*, 269-71.

46. Chai, S. C., Jerkins, A. A., Banik, J. J., Shalev, I., Pinkham, J. L., Uden, P. C., and Maroney, M. J. (2005) *J Biol Chem 280*, 9865-9.

47. Simmons, C. R., Hirschberger, L. L., Machi, M. S., and Stipanuk, M. H. (2005) *Protein Expr Purif*.

48. Bagley, P. J., Hirschberger, L. L., and Stipanuk, M. H. (1995) *Anal Biochem 227*, 40-8.

49. McCoy, J. G., Bailey, L. J., Bitto, E., Bingman, C. A., Aceti, D. J., Fox, B. G., and Phillips, G. N., Jr. (2006) *Proc Natl Acad Sci U S A 103*, 3084-9.

50. Stipanuk, M. H., Hirschberger, L. L., Londono, M. P., Cresenzi, C. L., and Yu, A. F. (2003) *Am J Physiol Endocrinol Metab*.

51. Ito, N., Phillips, S. E., Stevens, C., Ogel, Z. B., McPherson, M. J., Keen, J. N., Yadav, K. D., and Knowles, P. F. (1991) *Nature 350*, 87-90.

52. Que, L., Jr., and Ho, R. Y. (1996) *Chem Rev 96*, 2607-2624.

53. Hong, J., and Schoneich, C. (2001) *Free Radic Biol Med 31*, 1432-41.

54. Montellano, P. R. O. (1992) *Annu. Rev. Pharmacol. Toxicol. 32*, 89 - 107.

55. Simmons, C. R., Liu, Q., Huang, Q., Hao, Q., Begley, T. P., Karplus, P. A., and Stipanuk, M. H. (2006) *J. Biol. Chem.*, M601555200.

56. Heinamaki, A. A., and Piha, R. S. (1980) *Acta Chem Scand B 34*, 363-7.

57. Guion-Rain, M. C., Portemer, C., and Chatagner, F. (1975) *Biochim Biophys Acta 384*, 265-76.

58. Heinamaki, A. A., Peramaa, A. K., and Piha, R. S. (1982) *Acta Chem Scand B 36*, 287-90.

59. Tang, X. W., Hsu, C. C., Sun, Y., Wu, E., Yang, C. Y., and Wu, J. Y. (1996) *J Biomed Sci 3*, 442-453.

60. Tappaz, M., Almarghini, K., and Do, K. (1994) *Adv Exp Med Biol 359*, 257-68.

61. Do, K. Q., and Tappaz, M. L. (1996) *Neurochem Int 28*, 363-71.

62. Prescott, A. G., and Philip, J. (1996) *Annu. Rev. Plant Physiol Plant Mol. Biol. 47*, 245-271.

63. Buzenet, A. M., Fages, C., Bloch-Tardy, M., and Gonnard, P. (1978) *Biochim Biophys Acta 522*, 400-11.

64. Myllyharju, J., and Kivirikko, K. I. (1997) *Embo J 16*, 1173-80.

65. Jaakkola, P., Mole, D. R., Tian, Y. M., Wilson, M. I., Gielbert, J., Gaskell, S. J., Kriegsheim, A., Hebestreit, H. F., Mukherji, M., Schofield, C. J., Maxwell, P. H., Pugh, C. W., and Ratcliffe, P. J. (2001) *Science 292*, 468-72.

66. Epstein, A. C., Gleadle, J. M., McNeill, L. A., Hewitson, K. S., O'Rourke, J., Mole, D. R., Mukherji, M., Metzen, E., Wilson, M. I., Dhanda, A., Tian, Y. M., Masson, N., Hamilton, D. L., Jaakkola, P., Barstead, R., Hodgkin, J., Maxwell, P. H., Pugh, C. W., Schofield, C. J., and Ratcliffe, P. J. (2001) *Cell 107*, 43-54.

67. Borovok, I., Landman, O., Kreisberg-Zakarin, R., Aharonowitz, Y., and Cohen, G. (1996) *Biochemistry 35*, 1981-1987.

68. Valegard, K., van Scheltinga, A. C., Lloyd, M. D., Hara, T., Ramaswamy, S., Perrakis, A., Thompson, A., Lee, H. J., Baldwin, J. E., Schofield, C. J., Hajdu, J., and Andersson, I. (1998) *Nature 394*, 805-9.

69. Roach, P. L., Clifton, I. J., Hensgens, C. M., Shibata, N., Schofield, C. J., Hajdu, J., and Baldwin, J. E. (1997) *Nature 387*, 827-30.

70. Hegg, E. L., and Que, L., Jr. (1997) *Eur J Biochem 250*, 625-9.

71. Pavel, E. G., Zhou, J., Busby, R. W., Gunsior, M., Townsend, C. A., and Solomon, E. I. (1998) *J. Am. Chem. Soc. 120*, 743-753.

72. Hegg, E. L., Whiting, A. K., Saari, R. E., McCracken, J., Hausinger, R. P., and Que, L., Jr. (1999) *Biochemistry 38*, 16714-16726.

73. Barton, G. J. (1993) *Protein Eng 6*, 37-40.

74. Zhou, J., Gunisor, M., Bachmann, B. O., Townsend, C. A., and Solomon, E. I. (1998) *J. Am. Chem. Soc. 120*, 13539-13540.

75. Tabor, S. (1990) in *Current Protocols in Molecular Biology* (Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K., Eds.) pp 16.2.1-16.2.11, Green Publishing Associates and Wiley-Interscience, New York.

76. Chan, J. M., Wu, W., Dean, D. R., and Seefeldt, L. C. (2000) *Biochemistry 39*, 7221-8.

77. Powell, H. R. (1999) *Acta Crystallogr D Biol Crystallogr 55 (Pt 10)*, 1690-5.

78. Winn, M. D. (2003) *J Synchrotron Radiat 10*, 23-5.

79. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr D Biol Crystallogr 54 (Pt 5)*, 905-21.

80. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard. (1991) *Acta Crystallogr A 47 (Pt 2)*, 110-9.

81. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 945-949.

82. McRee, D. E. (1999) *J Struct Biol 125*, 156-65.

83. Merritt, E. A., and Bacon, D. J. (1997) *Methods Enzymol. 277*, 505-524.

84. Zhang, Z., Ren, J., Stammers, D. K., Baldwin, J. E., Harlos, K., and Schofield, C. J. (2000) *Nat Struct Biol 7*, 127-33.

85. Hogan, D. A., Smith, S. R., Saari, E. A., McCracken, J., and Hausinger, R. P. (2000) *J. Biol. Chem. 275*, 1200-12409.

86. Arnesano, F., Banci, L., Bertini, I., Ciofi-Baffoni, S., Molteni, E., Huffman, D. L., and O'Halloran, T. V. (2002) *Genome Res 12*, 255-71.

87. Burzlaff, N. I., Rutledge, P. J., Clifton, I. J., Hensgens, C. M., Pickford, M., Adlington, R. M., Roach, P. L., and Baldwin, J. E. (1999) *Nature 401*, 721-4.

88. 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) *Chem Rev 100*, 235-350.

89. Huxtable, R. J. M., D. (1994) *Advances ini experimental medicine and biology 359*.

90. Laemmli, U. K. (1970) *Nature 227*, 680-5.

91. Bradford, M. M. (1976) *Anal Biochem 72*, 248-54.

92. Bertini, I., Capozzi, F., Dikiy, A., Happe, B., Luchinat, C., and Timmis, K. N. (1995) *Biochem Biophys Res Commun 215*, 855-60.

93. Laine, R., and Montellano, P. R. O. d. (1998) *Mol Pharmacol 54*, 305-312.

94. Simmons, C. R., Hao, Q., and Stipanuk, M. H. (2005) *Acta Crystallograph Sect F Struct Biol Cryst Commun 61*, 1013-6.

95. Shimada, M., Koide, T., Kuroda, E., Tsuboyama, N., Hosokawa, Y., and Watanabe, M. (1998) *Amino Acids 15*, 143-50.

96. Bagley, P. J., and Stipanuk, M. H. (1995) *J Nutr 125*, 933-40.

97. Kohashi, N., Yamaguchi, K., Hosokawa, Y., Kori, Y., Fujii, O., and Ueda, I. (1978) *J Biochem (Tokyo) 84*, 159-68.

98. Loriette, C., Pasantes-Morales, H., Portemer, C., and Chatagner, F. (1979) *Nutr Metab 23*, 467-75.

99. Yamaguchi, K., Hosokawa, Y., Niizeki, S., Tojo, H., and Sato, I. (1985) *Prog Clin Biol Res 179*, 23-32.

APPENDIX A

INSIGHT INTO THE MECHANISM OF THE B_{12}-INDEPENDENT GLYCEROL DEHYDRATASE FROM *CLOSTRIDIUM BUTYRICUM*; PRELIMINARY BIOCHEMICAL AND STRUCTURAL CHARACTERIZATION.[2]

[2] O'Brien, J, R, Raynaud, C, Soucaille, P, and Lanzilotta, WN. *Biochemistry.* 2004 Apr 27;43(16):4635-45
Reprinted here with permission of publisher.

Abstract

The molecular characterization of a B_{12}-independent glycerol dehydratase from *Clostridium butyricum* has recently been reported (Raynaud et al. *P.N.A.S.* (2003) 100, 5010-5015). In this work, we have further characterized this system by biochemical and crystallographic methods. Both the glycerol dehydratase (GD) and the GD-activating enzyme (GD-AE) could be purified to homogeneity under aerobic conditions. In this form both the GD and the GD-AE were inactive. A reconstitution procedure, similar to what has been reported for pyruvate formate lyase activating enzyme (PFL-AE), was employed to reconstitute the activity of the GD-AE. Subsequently, the reconstituted GD-AE could be used to reactivate the GD under strictly anaerobic conditions. We also report here the crystal structure of the inactive GD in the native (2.5 Å resolution R_{cryst} = 17 %, R_{free} = 20 %), glycerol-bound (1.8 Å resolution R_{cryst} = 21 %, R_{free} = 24 %), and 1,2-propanediol-bound (2.4 Å resolution R_{cryst} = 20 %, R_{free} = 24 %) form. The overall fold of the GD monomer was similar to what has been observed for pyruvate formate lyase (PFL) and anaerobic ribonucleotide reductase (ARNR), consisting of a 10-stranded β/α barrel motif. Clear density was observed for both substrates and a mechanism for the dehydration reaction is presented. This mechanism clearly supports a concerted pathway for -OH group migration through a cyclic transition state that is stabilized by partial protonation of the migrating -OH group. Finally, despite poor alignment (RMSD ~ 6.8 Å) of the ten core strands that comprise the barrel structure of the GD and PFL, the C-terminal domains of both proteins align well (RMSD ~ 0.7 Å) and have structural properties consistent with this being the docking site for the activating enzyme. A single point mutation within this domain, at a strictly conserved arginine residue (R782K) in the GD, resulted in formation of a tight protein-protein complex between the GD and the GD-AE *in vivo*, thereby supporting this hypothesis.

Radical-catalyzed reactions in biological systems are of great significance. In particular, the arrival of the genomic and proteomic era, coupled with advances in scientific methods and technology, has had a dramatic impact on our understanding of protein radicals in catalysis. Among one of the most recent developments is the recognition of a new superfamily of enzymes that utilizes *S*-adenosylmethionine (SAM) in place of adenosylcobalamin (B_{12}) to catalyze radical reactions at enzymatic sites (For reviews see refs. (*1-4*). Despite intensive biochemical and spectroscopic investigation, the exact chemical details of radical generation in this new superfamily, termed the "radical SAM" superfamily (*5*), remains unclear. Similar to the B_{12}-dependent reactions, this new superfamily of enzymes has been identified in all kingdoms of life and has been shown to catalyze a diverse array of chemical reactions of significant medical and biotechnological importance (*5*). It has been further suggested that the diversity of chemical reactions catalyzed by this new superfamily exceeds those catalyzed by B_{12} (*1*). Although different enzymes in the radical SAM superfamily appear to require different cofactors specific to their "primary" metabolic substrate, it appears that a common mechanism is at work in the generation of the initial radical species. The cofactors required for this common activation mechanism are a $[4Fe\text{-}4S]^{1+}$ cluster (formally three Fe^{2+} and one Fe^{3+}) and *S*-adenosylmethionine (SAM). This catalytic cluster is similar to the cluster found in aconitase in that one iron is uncoordinated. The role of this open coordination site has been demonstrated by both spectroscopic and crystallographic methods and is to coordinate the SAM prior to reductive cleavage (*6,7*). Reductive cleavage of SAM results in the transient formation of a 5'-deoxyadenosyl radical (*8,9*). While this is reminiscent of the B_{12} reaction mechanism, the exact

chemical details are not well understood as this radical is highly reactive and rapidly transferred to the active site where conversion of the primary metabolic substrate can proceed.

Two areas of intense research have been initiated to investigate 1) the mechanism of radical generation by the Fe-S cluster and SAM as well as 2) the general mechanism of radical transfer and catalysis during the conversion of the primary substrate to product. For some radical SAM systems, all the components necessary for activation and conversion of the primary metabolic substrate are found within a single poly-peptide. In contrast, some systems require an additional enzyme for the initial radical generation. Of particular interest to this work are anaerobic ribonucleotide reductase (ARNR) and pyruvate formate lyase (PFL). In these radical SAM systems, the activating enzyme contains the catalytic [4Fe-4S] cluster and the SAM binding site, while the enzyme proper contains the active site for the prime metabolic substrate. Details regarding the complex formed between these two enzymes are essential to completing an accurate description of the activation mechanism and the mechanism of radical transfer.

One reaction that has, until most recently, been shown to be exclusively catalyzed by B_{12} is the dehydration of glycerol to 3-hydroxyl-propionaldehyde (3-HPA). In most systems the subsequent reduction of 3-HPA to 1,3-propanediol (1,3-PD) is catalyzed by a NADH-dependent reductase. The bacterial conversion of glycerol to 1,3-PD has become of significant interest due to the recent development of a new polyester called poly(propylene terephtalate). The polymer is essentially made by substituting 1,3-PD for ethylene glycol in producing polyester and results in a new fiber with superior properties (*10,11*). However, chemical processes for producing 1,3-PD involve toxic intermediates and chemical conversion under dangerous and expensive conditions. Not too surprisingly several patents have recently been filed that describe environmentally friendly approaches for conversion of renewable resources such as glucose to

1,3-PD (*12*,*13*). All of these approaches utilize recombinant microorganisms expressing the appropriate genes to convert a renewable resource, such as glucose, to 1,3-PD. The penultimate step of this pathway is the dehydration of glycerol to 3-HPA. The mechanism of both the B_{12}-dependent glycerol and diol dehydratases has been extensively studied and is fairly well understood (*14*). A severe limitation of the B_{12}-dependent glycerol and diol dehydratases is their rapid inactivation during enzyme turnover with glycerol (*15-18*). This limitation is overcome in the native organism by the presence of a reactivase enzyme that couples the energy of MgATP hydrolysis to the exchange of inactivated cofactor with new coenzyme B_{12} (*19*). Presently, engineered strains carrying the B_{12}-dependent system do not contain the reactivase, the inactivation problem is therefore overcome by the addition of excessive amounts of vitamin B_{12} to the culture medium. An exciting alternative has been presented by Raynaud et al. through their identification and molecular characterization of a B_{12}-independent glycerol dehydratase system from the strict anaerobe *Clostridium butyricum.* Similar to the B_{12}-dependent glycerol dehydratases, this system was also capable of dehydrating 1,2-propanediol to propionaldehyde in addition to glycerol dehydration (*20*). Further sequence analysis of this operon suggested that this system was also a member of the radical SAM superfamily. Specifically, that this system was similar to PFL in that it contained genes for the glycerol dehydratase (GD) as well as a glycerol dehydratase activating enzyme (GD-AE).

In order to further advance our understanding of the radical SAM superfamily as well as accurately describe a mechanism for the B_{12}-independent dehydration of glycerol, we have further isolated the B_{12}-independent GD and GD-AE. We report here the preliminary biochemical properties of this system as well as the crystallographic models for the radical-free form of the GD in the absence of substrate and with the substrates glycerol and 1,2-propanediol

bound. A detailed mechanism for the B_{12}-independent dehydration of glycerol is presented. In addition, the ability to follow the activity of the system *in vivo* and rapidly isolate either component has led to the identification of a single point mutation in the GD that forms a tight complex with the GD-AE. Taken together, the work presented here provides the first direct evidence for the site of complex formation in the two-component radical SAM systems and the first look at the atomic factors that catalyze B_{12}-independent glycerol dehydration in biology.

EXPERIMENTAL PROCEDURES

Plasmid Preparation and Enzyme Expression. The polycistronic plasmids used for co-expression of the glycerol dehydratase (GD) and the glycerol dehydratase activating enzyme (GD-AE) from *Clostridium butyricum* were prepared as previously described (*20*), except that an additional eleven residues (MWSHPQFEKRS) were added to the N-terminus of either the GD (plasmid pST11) or the GD-AE (plasmid pST12). This epitope is more commonly known as the Strep-Tag™ (IBA, Göttingen Germany) and is used for affinity purification. Similar to their parent plasmids (pSPD5 and pIMP1), the expression plasmids also carry the *dhaT* gene that codes for the 1,3-propanediol dehydrogenase. Consistent with what has also been previously reported (*20*), conversion of glycerol to 1,3-propanediol was observed during the anaerobic growth of *Escherichia coli* cells harboring either of these plasmids.

While the pST11 plasmid was sufficient for the isolation of the GD (~10 mg of pure enzyme from 20 g of wet cell paste), relatively small amounts of the GD-AE were obtained using the pST12 plasmid (<1 mg of pure enzyme from 50 g wet cell paste). In order to obtain better expression of the GD-AE, the *dhaB2* gene was amplified using PCR methods and the primers were designed to incorporate a NheI and HindIII restriction site at the 5' and 3' end of the gene respectively. The PCR product was then subcloned into the commercially available

expression vector pTRCHisA (Invitrogen Corporation, Carlsbad, CA). Instead of the Strep-Tag, this vector incorporates a 6x poly His tail at the N-terminal end of the GD-AE.

Purification of the GD. The following procedures could be performed under either aerobic or anaerobic conditions. If the procedure was performed under anaerobic conditions then *E. coli* cells harboring the pST11 plasmid were grown on LB media supplemented with nitrilotriacetic acid (200 $mg \cdot L^{-1}$), $FeSO_4$ (50 $mg \cdot L^{-1}$), K_2HPO_4 (0.5 $g \cdot L^{-1}$), Na selenate (30 $\mu g \cdot L^{-1}$) and glycerol (40 $g \cdot L^{-1}$). A 50 mL overnight culture was used to inoculate a 2000 mL culture that was contained in a 4000 mL glass carboy. Following inoculation, the carboy was sealed with a screw cap and the culture was incubated at 34°C and 110 rpm for 18 hours. The cells were harvested anaerobically by centrifugation and stored at -80°C. In addition, for anaerobic purifications, all buffers were made anaerobic by several rounds of degassing and flushing with oxygen-scrubbed argon and contained 1 mM Ti(III)citrate. All subsequent purification procedures were then performed in a glove box under an atomosphere of 95 % nitrogen and 5 % hydrogen.

Approximately 50 g of wet cells were resuspended in 350 mL of buffer containing 150 mM NaCL, 2 mM TCEP, and 50 mM TRIS pH 8.0 (Buffer A). Cells were then broken at 16,000 psi using a French press. The pellet was then discarded and 5 mg of avidin (Sigma, St. Louis MO) was added to the supernatant. The supernatant was then gently stirred for an additional 30 minutes at 4 ° C before being centrifuged at 10,000 x g for 10 minutes. This supernantant was passed through a Strep-Tactin® sepharose (IBA, Göttingen Germany) column that had been equilibrated with buffer A. Once the supernatant was loaded onto the column, the column was then washed with 20 volumes of buffer A. Buffer A containing 2.5 mM *d*-desthiobiotin (Sigma,

St. Louis MO) was then used to elute the bound protein. Purified protein was then exchanged into fresh buffer A and concentrated to 10 mg/mL.

Isolation and reconstitution of the GD-AE. Inactive GD-AE was isolated using the Talon™ metal affinity resin (Clonetech, Palo Alto CA). 50g of wet cells were resuspended in buffer containing 20% 1,2-propanediol, 100 mM HEPES pH 7.5, and 100 mM KCl (Buffer B). The resuspended cells were broken using a French press at 16,000 psi and centrifuged at 50,000 x g for one hour. The supernatant was then applied to affinity resin equilibrated with buffer B. The column was then washed with 20 column volumes of buffer B and the bound GD-AE was eluted with buffer B containing 150 mM imidazole. Isolated enzyme was immediately degassed on an argon manifold and concentrated to 1-2 mL. In this form the GD-AE was unable to activate the GD.

In order to reconstitute the GD-AE activity we employed the same procedure used for PFL-AE (*21*) except that the Mops buffer was replaced with HEPES and the reconstitution buffer also contained 20 % 1,2-propanediol. An incubation time of 12 hours was used for reconstitution and the reconstituted GD-AE was separated from excess iron and sulfide by passage through a sephadex G25 column equilibrated with 100 mM HEPES pH 7.5, 100 mM KCl, 2 mM mercaptoethanol and 2 mM sodium dithionite. The reconstituted GD-AE was concentrated to 1 mL. Typical yields for the entire isolation and reconstitution procedure were 20 mg of GD-AE from 50 g of wet cell paste. All enzymes used in this study were homogeneous as judged by SDS-PAGE and their concentrations were determined by a modified biuret method (*22*).

Activation of the GD by the GD-AE. It has been previously reported (*20*) that this system, like the B_{12}-dependent glycerol dehydratases, can also catalyze the dehydration of 1,2-

propanediol to propionaldehyde. In order to test the ability of the GD-AE to reactivate the GD, we modified a procedure that had been previously reported (*20*). The assay was performed under strictly anaerobic conditions in sealed quartz cuvettes at 25 ° C and the final assay volume was 0.5 mL. The assay solution contained 5 % 1,2-propandiol, 100 mM HEPES pH 7.5, 1.13 □M GD, 1 mg/mL yeast alcohol dehydrogenase (YADH) (Sigma Chemical Companay, lot# 070K7432), 220 μM NADH and 2 mM sodium dithionite. This solution was prepared with and without 5 mM SAM. Prior to initiating the assay, the spectrophotometer was blanked against buffer containing 100 mM KCl, 100 mM HEPES pH 7.5, and 2 mM sodium dithionite. The assay was initiated by addition of varying amounts of purified GD-AE and the absorbance at 340 nm was monitored. The GD-AE used in this study was purified as described above and concentrated to a final concentration of 18 $mg \cdot mL^{-1}$ prior to being used in the assay.

Crystallization of the GD. The best diffracting crystals of the GD were obtained using aerobically prepared protein and a precipitating solution containing 200 mM tri-ammonium citrate pH 7.0 and 10 % PEG 3350. Diffraction quality crystals could be obtained under anaerobic or aerobic conditions using either the capillary batch or hanging drop methods. Crystals of the GD belonged to the space group $C222_1$ and had unit cell dimensions of a=104, b=212, and c=199 Å (See Table A.1). These crystals were taken to be the substrate free GD crystals. These crystals could be soaked with up to 20% glycerol (glycerol-bound crystals) or 20% 1,2-propanediol (1,2-propanediol bound crystals) without any significant changes in the diffraction quality. For data collection on the native or substrate-free form of the enzyme, crystals were mounted in quartz capillaries and data was collected at room temperature.

Data Collection, Phasing, Model Building, and Refinement. Data for the substrate-free GD and the 1,2-propanediol bound form of the GD were collected at the University of Georgia

on a Rigaku RU-200 rotating anode equipped with Osmic focusing mirrors and a R-axisIIc image plate detector. Data for the glycerol-bound form of the GD were collected at the advanced light source (ALS, Berkeley CA) on beamline 8.2.2. All data were processed using HKL2000 (*23*) and the CCP4 suite of programs (*24*). While both glycerol and 1,2-propanediol are excellent cryo-protective agents, we were interested in comparing the data between the substrate-free enzyme and the glycerol- or 1,2-propanediol-bound enzyme. Therefore in order to collect data for both the substrate-free and a mercury derivative, data was collected at room temperature on crystals mounted in quartz capillaries. Under these conditions about 30 frames worth of data could be collected before the diffraction deteriorated. Five different crystals were used to complete the data sets for both the native (substrate-free) and Hg derivative data sets. A heavy atom derivative was used for phasing because, despite the conserved sequences shared between PFL and the GD (~23.0 % identical), attempts to obtain phases by molecular replacement were unsuccessful. Due to the high cysteine content of the GD (14 cysteines per monomer) a single Hg derivative was sufficient to obtain phase information good to 2.5 Å resolution, and therefore no phase extension was required prior to model building and refinement (Table A.1). Subsequent rounds of model building and refinement were performed using the programs O (*25*) and CNS (*26*). All graphical representations of structural elements were made using the programs molscript (*27*), Bobscript (*28*), and Raster3d (*29*).

Table A.1: Data Collection, Phasing Statistics, and Refinement Statistics for the Native Glycerol Dehydratase and Its Substrate-Bound Forms. [a]$R_{sym} = \sum_{hkl}[\sum_{I}(|I_{hkl,I} - \langle I_{hkl}\rangle|)] / \sum_{hkl,I} \langle I_{hkl}\rangle$, where I_{hkl} is the intensity of an individual measurement of the reflection with indices hkl and $\langle I_{hkl}\rangle$ is the mean intensity of that reflection. [b]The numbers in parenthesis refer to the outer resolution bin used in data processing. [c]The numbers refer to cell edges a, b, and c. The space group is $C222_1$ so all angles are 90°.

	native glycerol-free	glycerol-bound	1,2-propanediol-bound	Hg-derivative
number of crystals used	5	1	1	5
wavelength (Å)	1.54	0.98	1.54	1.54
resolution range (Å)	50.0-2.5	50.0-1.8	50.0-2.4	50.0-2.5
unique observations	74,994	199,038	83,330	73,024
total observations	427,249	1,002,737	491,798	422,078
completeness (%)	98.6(93.5)[b]	99.5(96.5)	98.0(90.8)	96.9(90.3)
R_{sym} (%)[a]	9.9(38.5)	8.3(28.2)	9.6(31.2)	12.7(44.7)
I/σ	29.7(2.2)	26.7(3.3)	20.0(3.2)	27.8(1.7)
R_{cullis} (cen/acen/anom)				0.84/0.80/0.81
phasing power (cen/acen/anom)				1.24/1.47/1.45
isomorphous difference (%)				27.8
number of sites				8

model	substrate-free	glycerol-bound	1,2-propanediol-bound
unit cell[c]	103.4 212.9 199.4	103.0 213.4 199.5	102.6 212.3 198.7
protein atoms	12,370	12,390	12,392
solvent atoms	511	969	696
resolution limits	50.0-2.5	50.0-1.8	50.0-2.4
R_{cryst} (%)	17.1	21.7	19.8
R_{free} (%)	20.3	23.8	23.5
rmsd bonds (Å)	0.005	0.005	0.005
rmsd mainchain angles (°)	1.21	1.21	1.19
average B factor (Å^2)	28.4	31.1	27.0

RESULTS AND DISCUSSION

Isolation of the Glycerol Dehydratase (GD) from C. butyricum and Activation of the GD by the GD-AE. Expression and isolation of the GD could be performed under both aerobic and anaerobic conditions. However, as was previously reported by Raynaud et al., in order to observe any activity *in vivo* or in crude extracts, *E. coli* cells expressing both the GD gene

(*dhaB1*) and the GD-AE gene (*dhaB2*) had to be grown under strictly anaerobic conditions using a modified media containing glycerol (*20*). Using the Strep-tactin® sepharose affinity resin, the GD could be isolated from aerobically grown cells under aerobic conditions or anaerobically grown cells under anaerobic conditions. However, if an aliquot of the elution fraction from anaerobically prepared enzyme was removed from the glove box and immediately exposed to air, then a cleavage product was observed. This is shown in Figure A.1 and is reminiscent of the oxygenolytic cleavage of activated PFL (*30*). Therefore, in order to insure that we had a homogenous population of enzyme for crystallization trials, all of the GD used in the biochemical and crystallographic work was expressed and purified by aerobic methods. In this state the enzyme is inactive, presumably because it does not contain the active site radical. The aerobic isolation of the GD-AE is also shown in Figure A.1.

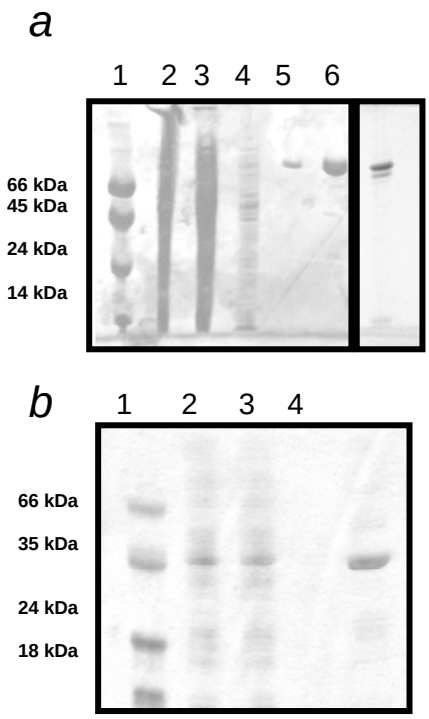

Figure A.1: Isolation of the GD and the GD-AE. ***Panel A;*** Isolation of recombinant GD under aerobic conditions and the affect of oxygen exposure on activated enzyme. *Lane 1.* Molecular weight markers. *Lane 2.* Crude extract from areobic preparation. *Lane 3.* Flow-through from Strep-Tactin® sepharose column (aerobic preparation). *Lane 4.* Wash fraction after 10 column volumes of buffer A (aerobic preparation). *Lane 5 and 6.* First and Second elution fractions using buffer A containing 2.5 mM *d*-desthiobiotin (aerobic preparation). *Lane 7.* GD isolated under strictly anaerobic conditions from anaerobically grown *E. coli* expressing both the GD and the GD-AE after a five minute exposure to air. ***Panel B;*** Isolation of recombinant GD-AE under aerobic conditions. *Lane 1.* Molecular weight markers. *Lane 2.* Crude extract. *Lane 3.* Flow-through from Talon™ metal affinity resin during the load. *Lane 4.* Wash fraction after 10 column volumes of buffer B. *Lane 7.* Elution fraction using buffer B containing 150 mM imidazole.

In order to test whether or not the GD used in this investigation could be activated by the GD-AE, we developed a simple assay based on the observation of Raynaud et al. that the system was also capable of dehydrating 1,2-propanediol to propionaldehyde. This is a coupled assay that utilizes a ten-fold excess of yeast alcohol dehydrogenase (YADH) to catalyze the NADH-dependent conversion of propionaldehyde to 1-propanol. The coupled reaction is monitored at 340 nm and the results from these preliminary studies are shown in Figure A.2. Following the addition of the GD-AE, a delay was observed prior to a decrease in the absorbance at 340 nm. No decrease was observed if either SAM or the GD-AE was excluded from the assay.

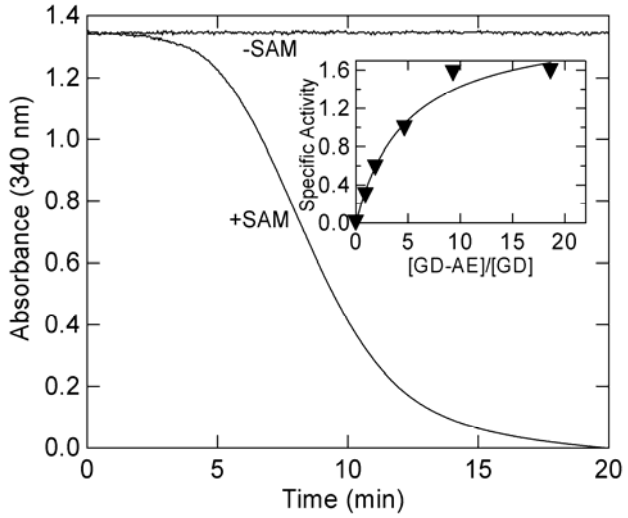

Figure A.2: *S*-adenosylmethionine (SAM) dependent activation of the GD by the GD-AE. The coupled assay was performed as described in the Experimental Procedures. The absorbance at 340 nm was monitored to follow the NADH-dependent reduction of propionaldehyde formed by GD-catalyzed dehydration of 1,2-propanediol (+SAM trace). No change in the absorbance was observed if SAM was left out of the assay solution (-SAM trace). ***Inset;*** The affect of the

GD-AE concentration on activation of the GD. The specific activity was calculated from the maximum slope of the coupled assay using 6.2 mM^{-1} cm^{-1} as the molar extinction coefficient for NADH.

Interestingly, the maximum activity was achieved once the GD-AE:GD ratio exceeded 10. Two important pieces of information must be considered in interpreting our data. First, our assay does not utilize the light-activated reductant 5'-deazariboflavin, as has been routinely used for PFL-AE (*21*). Second, it has been shown that activation of PFL by the PFL-AE requires an intact $[4Fe\text{-}4S]^{1+}$ cluster to catalyze the reductive cleavage of SAM and generate the catalytic radical on PFL (*31*). At this time, we do not know if sodium dithionite is sufficient for reduction of the catalytic Fe-S cluster in the GD-AE. Moreover, at the present time we do not know if the reconstitution procedure has restored the catalytic Fe-S cluster in 100 % of the apo GD-AE. In addition, the sequence of the GD-AE suggests that a second cluster is also present in the enzyme (*20*). The exact role of this second cluster has not been characterized in any radical SAM system. However, it should be noted that a second cluster binding site has also been identified in the sequences of TutE and BssD, these are also radical-SAM enzymes involved in the benzyl succinate synthase system from two *Thauera* species (*32*,*33*). We are currently investigating the properties of the activated GD-AE to address these issues. What can be said is that the reconstituted GD-AE can catalytically re-activate the GD in a SAM-dependent manner. A maximum specific activity for the GD of 1.56 ± 0.09 $mmoles \cdot min^{-1} \cdot mg^{-1}$ is obtained when the ratio of GD-AE to GD exceeds ten in our assay. Because glycerol is both a good substrate and a potent inhibitor of the B_{12}-dependent glycerol dehydratase (*34*), a comparison of the 1,2-propanediol dehydratase activity in the two systems is valid. Interestingly, the specific activity

we have observed for the B_{12}-independent system is approximately thirteen times that of the B_{12}-dependent glycerol dehydratase (*35*). Moreover, this specific activity also exceeds that typically reported for the B_{12}-dependent diol dehydratases (*15,16,36*). The biochemical parameters for glycerol dehydration are currently being pursued by this laboratory using purified DhaT in place of YADH in the coupled assay.

Overall Structure of the Glycerol Dehydratase Monomer. In order to provide a solid foundation for future biochemical and biophysical investigations, we have determined the crystal structure of the B_{12}-independent GD in the native, glycerol-bound and 1,2-propanediol-bound forms. In all of the structures reported here the complete translated peptide (residues 2-787) is clearly visible in the electron density. The overall structure of the GD is shown in Figure A.3. Similar to PFL and ARNR, the GD monomer forms a core 10-stranded β-barrel motif that is assembled in an antiparallel manner from two parallel five-stranded β-sheets. This β-barrel core is entirely surrounded by a α-helices as was also reported for PFL and ARNR and has been termed a β/α-barrel structure. Despite the observation of a highly conserved core, if the ten strands that compose this motif are aligned between PFL (PDB ID 1H16) and the GD, then a root mean square deviation (RMSD) of 6.8 Å is obtained for the backbone α-carbon atoms. In addition to the significant variation in the core motif, there are also dramatic differences in the number and position of the surrounding helices. If the conserved helical regions are also included in this alignment the RMSD for backbone α-carbon atoms is substantially worse (>8.0 Å). A substantial finding of this work is the observation that the only region of the GD that aligns well with PFL is the C-terminal domain (exploded view, Figure A.4, Panel A). For PFL and the GD this includes residues 731-782 (GD) and residues 702-754 (PFL) respectively. Alignment of this domain results in an RMSD of 0.7 Å for the backbone α-carbon atoms. This

domain is highlighted in yellow in Figure A.3 and contains the radical glycine loop, a small anti-parallel sheet section and four helices. Notably, two of the helices are on the surface of the enzyme. This domain can therefore be described as extending from the surface of the protein to the substrate binding site at one end of the 10-stranded β-barrel.

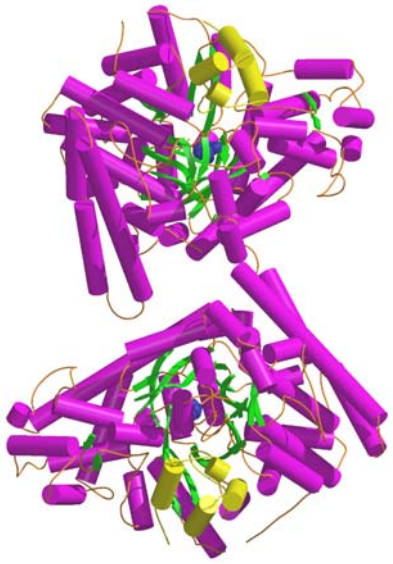

Figure A.3: Overall fold of the glycerol dehydratase (GD) monomer and the proposed biological dimer. Helices are shown as purple cylinders, sheet structure is represented by green strands, and random coil is represented by orange coils. The topology of the C-terminal region (Residues 731-787) is represented similarly, but is colored yellow in the monomer and dimer representations. *Panel A.* Overall topology of the monomer and exploded view of the C-terminal region of the enzyme. *Panel B.* The proposed biological dimer of the GD, for clarity monomer 1 and monomer 2 are labeled.

Oligomeric Structure and a Potential Site for Binding of the Activating Enzyme. Similar to the most recent model of PFL (*37*), our GD crystals also belong to the space group $C222_1$ and contain two monomers in the asymmetric unit. Like PFL, the two monomers are related by a nearly perfect two-fold rotation and represent the closest contact between any two monomers in the unit cell with a buried surface area that exceeds 1000 $Å^2$ per monomer. Consistent with what has been proposed for PFL, we propose that this arrangement also represents the biological dimer (monomer 1 and 2 in Figure A.3, Panel B). Further efforts to investigate whether or not the GD exists as a dimer in solution and whether or not this oligomeric arrangement has any impact on function are currently being pursued by this laboratory. However, if this is the correct biological dimer than the C-terminal domain of each monomer is found on the opposite side of the dimer interface. One significant property that both the GD and PFL have in common is that they must be activated by another enzyme. This involves docking of the activating enzyme and therefore a significant unanswered question is "where is the site of complex formation?" Given i) the overall structural similarities of PFL and the GD, ii) the observation that a similar reconstitution procedure can be used to reconstitute the PFL-AE as well as the GD-AE, and iii) the observation that the only region of PFL and the GD that aligns well is the C-terminal domain, we propose that this is the site of complex formation between the activating enzyme and the enzyme proper in both systems. To further examine this hypothesis, a sequence and structural alignment of this domain between PFL and the GD is shown in Figure A.5. All of the side chains that are shown in Figure 4 are strictly conserved between PFL and the GD and are further found in identical conformations in both crystallographic models. Of particular interest is the *cis* conformation of the strictly conserved glycine residue. Another noteworthy residue that is strictly conserved in sequence and structure is R782 in the GD (R753 in PFL). In the structure of

the GD and PFL this residue extends inward from the C-terminal helix to form a hydrogen bond with the backbone carbonyl of a residue that is only two amino acids away from the active site glycine (Dashed lines Figure A.4, ~2.9 Å). Again, in both crystallographic models this interaction is further stabilized through a salt bridge to an acidic residue found within the same C-terminal helix. The distance from the surface of the GD to the active site along this pathway is ~17 Å, however in both the PFL and the GD models, a substantial amount of water molecules are clearly seen on the interior side of this helix. This observation may suggest that this helix moves during activation. Further evidence for this hypothesis comes from the single point mutation R782K in the GD.

GD 731-GFHVQFNVIDKKILLAAQKNPEKYQDLIVRVAGYSAQFISLDKSIQNDIIART-783

PFL 703-GQHLNVNVMNREMLLDAMENPEKYPQLTIRVSGYAVRFNSLTKEQQQDVITRT-754

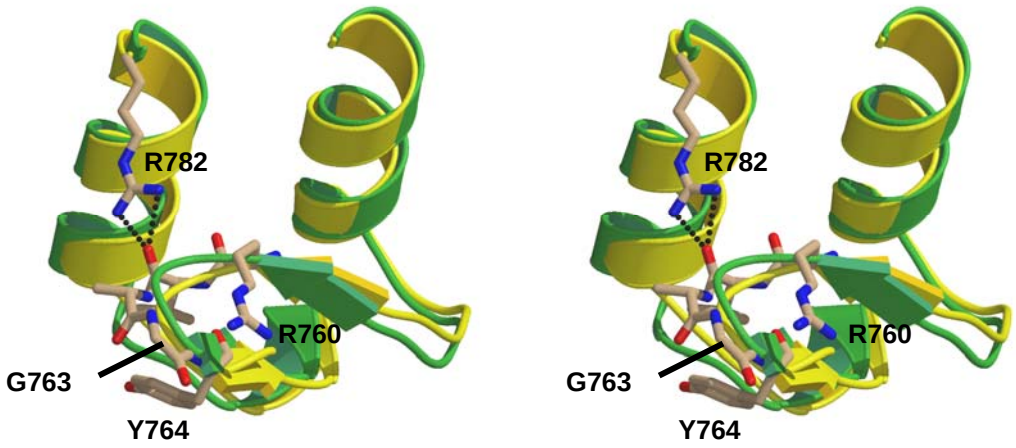

Figure A.4: Sequence and structural comparison of the C-terminal domains of PFL and the GD. Identical residues are colored red and similar residues are colored blue in the sequence alignment. The structural topology of residues in the sequence alignment is shown in a cross-eyed stereo view with PFL colored green and the GD colored yellow. The side chains of strictly conserved residues in the GD model are represented in ball and stick with the carbon atoms colored light brown, oxygen atoms colored red and nitrogen atoms colored blue. The side chain of R782 comes within 3.0 Å of the backbone carbonyl oxygen of residue 761 (dashed line).

During our initial attempts to construct the pST12 plasmid, a single point mutation was serendipitously introduced in the GD gene and sequence analysis confirmed that AAG was substituted for AGG in the codon for residue 782. Sequence analysis also confirmed that the sequence for the affinity-tagged GD-AE was correct. Since the point mutation was on the GD and the Strep-Tag™ was on the GD-AE, we thought that this plasmid (pG782KGD-ST12) might still be useful for production of wild-type GD-AE as well as to investigate the impact of this point mutation on the *in vivo* activity of the glycerol dehydratase system. Surprisingly, no activity could be detected in crude extracts or *in vivo* for this construct. As was observed for the parent plasmids (*20*) reconstruction of this plasmid (without the R782K point mutation) resulted in full activity *in vivo*. To further confirm that the loss of activity was due to the R782K point mutation in the GD and not the affinity tag on the N-terminus of the GD-AE, we have shown in Figure A.1 and Figure A.2 it is possible to isolate the affinity-tagged GD-AE and use it to activate the wild-type GD. We then took a step backwards and attempted to isolate the GD-AE anaerobically from crude extracts of *E. coli* cells harboring the pR782KGD-ST12 plasmid that were grown under the conditions previously described (*20*). The results of this purification are shown in Figure A.5. Despite the fact that the affinity tag is on the GD-AE, both the GD and the GD-AE eluted from the column after washing the column with several column volumes of buffer. These data suggest that a tight protein-protein complex has been formed *in vivo* as a result of the point mutation R782K on the GD. This observation further supports our hypothesis that this region is the site of complex formation between the GD and the GD-AE. The factors that are required to form the complex are currently being investigated by isolating the R782K GD and wild-type GD-AE independently and then initiating complex formation *in vitro.* An exciting possibility is that the complex forms after electron transfer and radical generation and

that this arginine residue has a role in dissociation. Structural and spectroscopic characterization of such a complex with the radical intact would be extremely valuable.

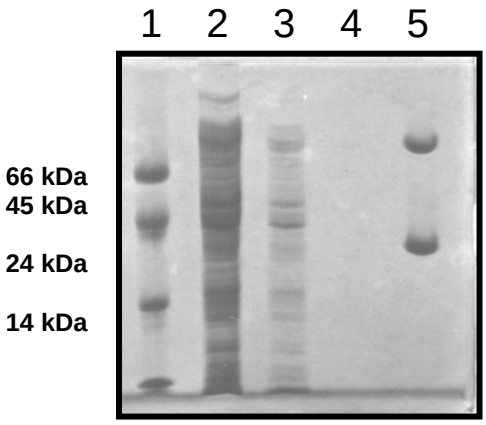

Figure A.5: Isolation of a tight protein-protein complex between the GD-AE and the R782K GD *in vivo.* *Lane 1.* Molecular weight markers. *Lane 2.* Crude extract from *E. coli* cells harboring the *pR782KGD-ST12* plasmid. *Lane 3.* First wash fraction. *Lane 4.* Last wash fraction. *Lane 5.* Elution fraction.

Substrate Binding. Because both glycerol and 1,2-propanediol have been shown to be good substrates for the GD (*20*), the interaction of both of these substrates was investigated. The binding mode of glycerol is shown in Figure A.6.

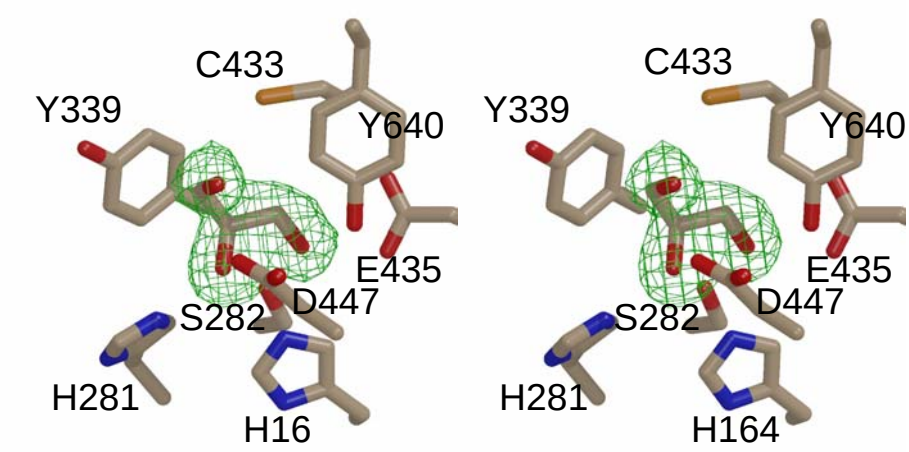

Figure A.6: Cross-eyed stereoview of the active site with glycerol bound. All atoms are represented by ball-and-stick models with the carbon atoms colored light brown, oxygen atoms colored red, nitrogen atoms colored blue, and sulfur atoms colored orange. The green cage represents a fourier difference (F_o-F_c) map contored at 4σ using the observed glycerol-bound structure factors and the calculated phases from the substrate-free model.

A number of hydrogen bonds position the glycerol molecule so that both of the terminal carbon atoms are each positioned 3.7 Å away from the γ-sulfur atom of the active site cysteine (C433). This residue is strictly conserved in PFL and ARNR. For PFL, Kozarich and co-workers have shown that a thiyl radical is generated at this site as the result of hydrogen atom abstraction by the active site glycyl radical (*38*). Comparison of the substrate-free and substrate-bound form of the GD suggests that there is some movement of the γ-sulfur atom of C433 towards the α-carbon atom of G763 as this distance is 4.4 Å in the substrate-free structure and 4.0 Å in the glycerol-bound structure. A thiyl radical on the γ-S atom of C433 would be perfectly positioned to abstract a hydrogen atom from either end of the glycerol. The question that arises is then "which

end of the substrate is the site of hydrogen atom abstraction?" This is easily answered by examining the binding mode of the alternative substrate 1,2-propanediol. The binding mode of 1,2-propanediol is shown in Figure A.7. Of particular interest is the purple density, that essentially represents the missing –OH group when the 1,2-propanediol and glycerol data are compared. These data clearly place the 1-OH group of 1,2-propanediol in a hydrogen bond with the carboxylic acid oxygen of E435 (2.6 Å). At the opposite end of either substrate, a direct connection between the bulk solvent and the active site is observed. This is accomplished through a series of hydrogen-bonded water molecules that extend from the side chain of D447 to the surface of the enzyme and into the bulk solvent.

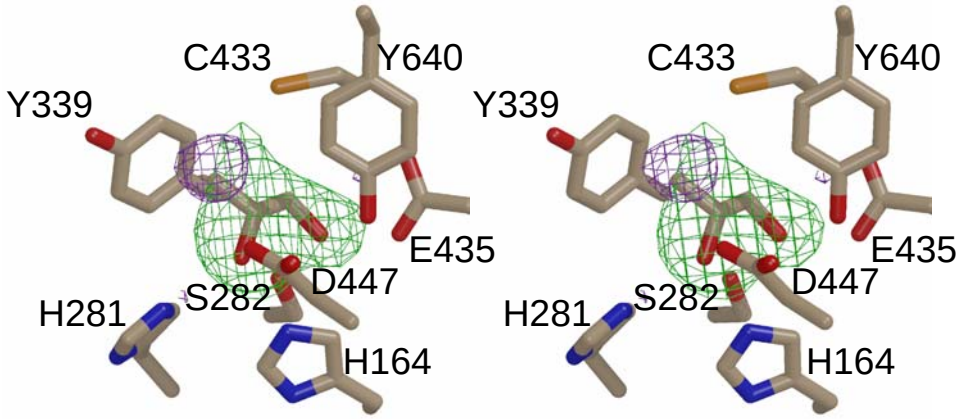

Figure A.7: Cross-eyed stereoview of the active site with 1,2-propanediol bound. All atoms are represented by ball-and-stick models with the carbon atoms colored light brown, oxygen atoms colored red, nitrogen atoms colored blue, and sulfur atoms colored orange. The green cage represents a fourier difference (F_o-F_c) map contored at 4σ using the observed 1,2-propandiol-bound structure factors and the calculated phases from the substrate-free model. The purple cage represents a fourier difference (F_o-F_c) map contored at 4σ using the observed

glycerol-bound structure factors and the calculated phases from the 1,2-propanediol-bound model.

The Mechanism of B_{12}-Independent Glycerol Dehydration. Given the structural details revealed here and the extensive information available for the B_{12}-dependent glycerol and diol dehydratases, a reasonable hypothesis for the B_{12}-independent mechanism of glycerol dehydration can be surmised. Given that the GD system is structurally similar to PFL and the observation that the activation of the GD by the GD-AE is SAM-dependent, we propose that the activation mechanism will be similar to PFL. Specifically, that activation of the GD by the GD-AE will result in a stable glycyl radical at G763 and that this radical will be in equilibrium with a thiyl radical at the γ-sulfur atom of C433. Whether or not this is the case, and the mechanism of radical propagation from the GD-AE to the active site of the GD, is currently under investigation. In particular, the observation that the key glycine residue is structurally conserved in the models of the GD, ARNR, and PFL would suggest that these enzymes will all share a common activation mechanism. While PFL appears to require two cysteine residues (C418 and C419) for activity, only one of these residues (C419) is conserved in the sequence and structures of the GD and ARNR. This would be equivalent to C433 in the GD and would represent the site of the thiyl radical. Therefore our proposed mechanism for B_{12}-independent glycerol dehydration is shown in Figure 8, and begins at the point where, by analogy with PFL, the active site glycyl radical has already abstracted a hydrogen atom from the γ-sulfur atom of C433.

The binding modes of the substrates examined here indicate that a thiyl radical at C433 would be perfectly positioned to abstract the *proS* hydrogen atom from the C1 carbon of either

substrate. A significant mechanism for lowering the activation barrier for this event may come from the resonance of the peptide bond between residue 433 and residue 434. Since the amide proton of this peptide is hydrogen bonded to the acidic side chain of E435, the strength of this hydrogen bond is associated with the overall resonance of the peptide bond. This will further have a direct impact on the hydrogen bond between the hydroxyl group of the substrate and side chain of E435 (See Figure A.8, Steps A & B). At the present time it is not clear how this equilibrium might influence the initial abstraction of the hydrogen atom.

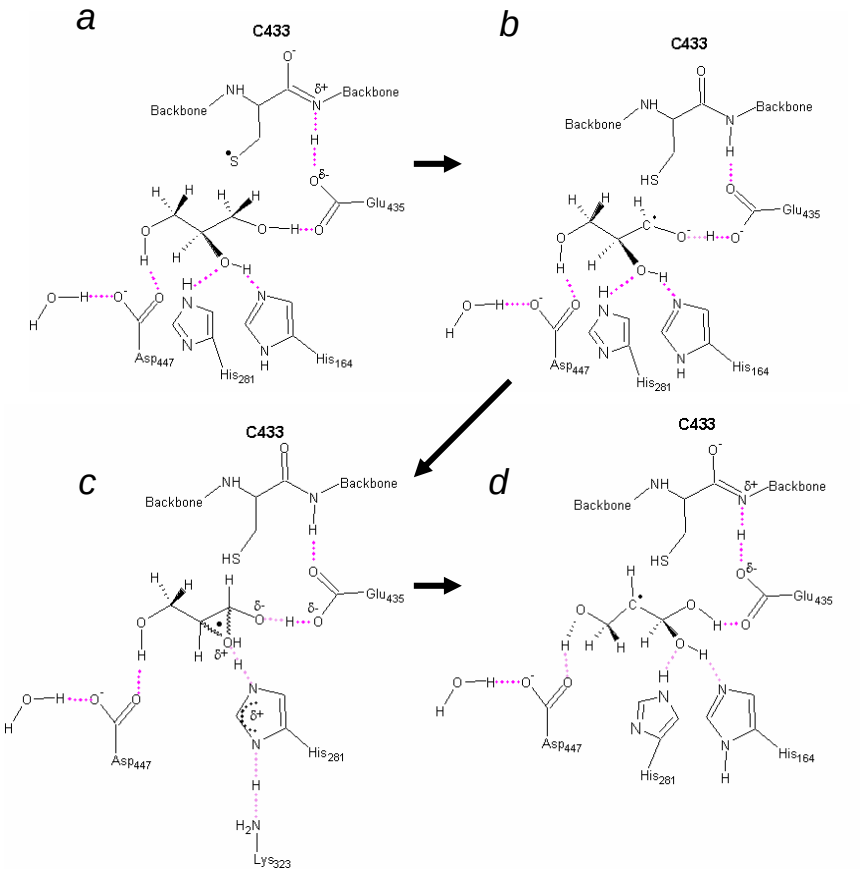

Figure A.8: Proposed mechanism of vitamin B_{12}-independent glycerol dehydration by the *Clostridium butyricum* glycerol dehydratase.

Although the B_{12}-dependent reactions are quite diverse, all of them, except for the ribonucleotide reductase reaction, involve an intramolecular group transfer reaction. In the case of the diol dehydratase reaction mechanism, this involves the overall migration of a hydrogen atom in exchange for the movement of the OH group in the opposite direction. Due to the nature of these reactions, the mechanism of group migrations is difficult to study and has been the subject of considerable debate (Recently reviewed in (*14*). In particular, it is difficult to get direct evidence for a specific mechanism from biochemical studies. Therefore, model reactions, X-ray structures, and theoretical calculations have provided significant insight. Of particular interest to this work is the proposal by Golding and Random (*39*) that the barrier in the intramolecular 1,2-shifts of an OH group, through a cyclic transition state, is reduced by protonation of the migrating OH group. A significant problem is that the nature of an acid required to protonate this species would be outside the range of acidic groups found in proteins. Although the active site of an enzyme can modulate pK_a values substantially, the pH at which the enzymes are functioning in general would indicate that the majority of the acidic groups would exist as $–COO^-$ and therefore the conjugate acids of the active site residues are more likely. A Lewis acid such as K^+ might also lower the energy of this transition state. The latter proposal is consistent with the requirement for potassium in the B_{12}-dependent dehydration of glycerol and the position of K^+ and 1,2-propanediol in the binding site of the B_{12}-dependent glycerol dehydratase model (*35*). In both the glycerol and 1,2-propanediol bound forms of the B_{12}-independent GD presented here, we clearly see the interaction of two histidine residues with what would ultimately be the migrating OH group. Because H164 is involved in a hydrogen bond with a backbone carbonyl it would better serve as the hydrogen bond acceptor for the OH group at the C2 carbon of substrate. H281 is a hydrogen bond acceptor for the side chain amino

group of K323 and therefore would be better suited to serve as a conjugate acid for the stabilization of the transition state shown in Figure A.8, Panel C. The nitrogen atom of H281 is also slightly closer to the migrating OH group in both the glycerol- and 1,2-propanediol-bound structures (2.8 Å compared with 3.0 Å). This mechanism is further supported by computational results that indicate that the presence of imidazolium is effective for transition state stabilization in the absence of K^+ (*40*). In order for these histidine residues to maintain their interaction with this OH group, the product must reorient itself in the active site. This would essentially amount to a 180 ° rotation that would place the 2° carbon radical in proximity to the γ-S atom of C433 (Figure A.8, Panel D). In this sense, the migration of the OH group is essential to regenerating the catalytic radical at the active site. Whether or not the migration of this OH group is stereospecific, as has been shown for the B_{12}-dependent glycerol and diol dehydratases, remains to be addressed.

Summary. In this work we present the first isolation and preliminary biochemical characterization of a B_{12}-independent glycerol dehydratase. The 1,2-propandiol dehydratase activity is significantly greater than what has previously been reported for the B_{12}-dependent glycerol and diol dehydratases. We also report the crystal structure for the B_{12}-independent GD and show that it is similar to PFL in both structure and the SAM-dependent activation mechanism. The observation that only the C-terminal domains of PFL and the GD crystal stuctures align, coupled with the observation that a single point mutation (R782K) in this domain results in the formation of a tight complex *in vivo*, provides support for our hypothesis that this is the binding site for the activating enzyme. Further crystallographic analysis of the binding modes of glycerol and 1,2-propanediol have provided critical insight into the mechanism of B_{12}-independent dehydration reactions. Taken together, the biochemical and structural data

presented here provide a solid foundation for further understanding the B_{12}-independent dehydration reactions and the radical SAM activation mechanism in the two-component systems such as PFL.

REFERENCES

1. Frey, P. A. (2003) S-Adenosylmethionine: A wolf in sheep's clothing or a rich man's adenosyl cobalamin. *Chem.Rev. 103*, 2129-2148.
2. Cheek, J., and Broderick, J. B. (2001) Adenosylmethionine-dependent iron-sulfur enzymes: versatile clusters in a radical new role *J. Biol. Inorg. Chem. 6*, 209-226.
3. Stubbe, J., and van der Donk, W. A. (1998) Protein Radicals in Enzyme Catalysis *Chem. Rev. 98*, 705-762.
4. Fontecave, M., Mulliez, E., and Ollagnier-de-Choudens, S. (2001) Adenosylmethionine as a source of 5'-deoxyadenosyl radicals *Curr Opin Chem Biol 5*, 506-11.
5. Sofia, H. J., Chen, G., Hetzler, B. G., Reyes-Spindola, J. F., and Miller, N. E. (2001) Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: functional characterization using new analysis and information visualization methods *Nucleic Acids Res 29*, 1097-106.
6. Walsby, C. J., Ortillo, D., Broderick, W. E., Broderick, J. B., and Hoffman, B. M. (2002) An Anchoring Role for FeS Clusters: Chelation of the Amino Acid Moiety of S-Adenosylmethionine to the Unique Iron Site of the [4Fe-4S] Cluster of Pyruvate Formate-Lyase Activating Enzyme *J. Amer. Chem. Soc. 124*, 11270-11271.
7. Layer, G., Moser, J., Heinz, D. W., Jahn, D., and Schubert, W. D. (2003) Crystal structure of coproporphyrinogen III oxidase reveals cofactor geometry of Radical SAM enzymes *Embo J 22*, 6214-6224.

8. Frey, M., Rothe, M., Wagner, A. F., and Knappe, J. (1994) Adenosylmethionine-dependent synthesis of the glycyl radical in pyruvate formate-lyase by abstraction of the glycine C-2 pro-S hydrogen atom. Studies of [2H]glycine-substituted enzyme and peptides homologous to the glycine 734 site *J Biol Chem 269*, 12432-7.

9. Wagner, A. F., Demand, J., Schilling, G., Pils, T., and Knappe, J. (1999) A dehydroalanyl residue can capture the 5'-deoxyadenosyl radical generated from S-adenosylmethionine by pyruvate formate-lyase-activating enzyme *Biochem Biophys Res Commun 254*, 306-10.

10. Miller, H. (2000) *Int. Fiber J. 15*, 14-16.

11. Rudie, R. (2000) *Int. Fiber J. 15*, 8-12.

12. Sarcabal, P., Croux, C., and Soucaille, P. (2001) in *PCT Int. Appl.*, France.

13. Emptage, M., Haynie, S., Laffend, L., Pucci, J., and Whited, G. (2001) in *Patent Corporation Treaty (PCT)*.

14. Toraya, T. (2003) Radical catalysis in coenzyme B12-dependent Isomerization (Eliminating) reaxtions *Chem.Rev. 103*, 2095-2127.

15. Toraya, T., Shirakashi, T., Kosuga, T., and Fukui, S. (1976) Substrate specificity of coenzyme B12-dependent diol dehydrase: glycerol as both a good substrate and a potent inactivator *Biochem Biophys Res Commun 69*, 475-80.

16. Bachovchin, W. W., Eagar, R. G., Jr., Moore, K. W., and Richards, J. H. (1977) Mechanism of action of adenosylcobalamin: glycerol and other substrate analogues as substrates and inactivators for propanediol dehydratase--kinetics, stereospecificity, and mechanism *Biochemistry 16*, 1082-92.

17. Schneider, Z., and Pawelkiewicz, J. (1966) The properties of glycerol dehydratase isolated from Aerobacter aerogenes, and the properties of the apoenzyme subunits *Acta Biochim Pol 13*, 311-28.

18. Poznanskaya, A. A., Yakusheva, M. I., and Yakovlev, V. A. (1977) Study of the mechanism of action of adenosylcobalamindependent glycerol dehydratase from Aerobacter aerogenes. II. The inactivation kinetics of glycerol dehydratase complexes with adenosylobalamin and its analogs *Biochim Biophys Acta 484*, 236-43.

19. Toraya, T., and Mori, K. (1999) A reactivating factor for coenzyme B12-dependent diol dehydratase *J Biol Chem 274*, 3372-7.

20. Raynaud, C., Sarcabal, P., Meynial-Salles, I., Croux, C., and Soucaille, P. (2003) Molecular characterization of the 1,3-propanediol (1,3-PD) operon of Clostridium butyricum *P.N.A.S. 100*, 5010-5015.

21. Kulzer, R., Pils, T., Kappl, R., Huttermann, J., and Knappe, J. (1998) Reconstitution and characterization of the polynuclear iron-sulfur cluster in pyruvate formate-lyase-activating enzyme. Molecular properties of the holoenzyme form *J Biol Chem 273*, 4897-903.

22. Chromy, V., Fischer, J., and Kulhanek, V. (1974) Re-evaluation of EDTA-chelated biuret reagent *Clin Chem 20*, 1362-3.

23. Otwinowski, Z., and Minor, W. (1997) Processing of x-ray diffraction data collected in oscillation mode. *Methods in Enzymology 276*, 307-326.

24. Potterton, E., Briggs, P., Turkenburg, M., and Dodson, E. (2003) A graphical user interface to the CCP4 program suite *Acta Crystallogr D Biol Crystallogr 59*, 1131-7.

25. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models *Acta Crystallogr A 47 (Pt 2)*, 110-9.

26. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination *Acta Crystallogr D Biol Crystallogr 54 (Pt 5)*, 905-21.

27. Kraulis, P. J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr. 24*, 945-949.

28. Gouet, P., Robert, X., and Courcelle, E. (2003) ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins *Nucleic Acids Res 31*, 3320-3.

29. Merritt, E. A., and Bacon, D. J. (1997) *Methods Enzymol. 277*, 505-524.

30. Wagner, A. F., Frey, M., Neugebauer, F. A., Schafer, W., and Knappe, J. (1992) The free radical in pyruvate formate-lyase is located on glycine-734 *Proc Natl Acad Sci U S A 89*, 996-1000.

31. Henshaw, T. F., Cheek, J., and Broderick, J. B. (2000) The [4Fe-4S]1+ Cluster of Pyruvate Formate-Lyase Activating Enzyme Generates the Glycyl Radical on Pyruvate Formate-Lyase: EPR-Detected Single Turnover *J. Am. Chem. Soc. 122*, 8331-8332.

32. Coschigano, P. W., Wehrman, T. S., and Young, L. Y. (1998) Identification and analysis of genes involved in anaerobic toluene metabolism by strain T1: putative role of a glycine free radical *Appl Environ Microbiol 64*, 1650-6.

33. Hermuth, K., Leuthner, B., and Heider, J. (2002) Operon structure and expression of the genes for benzylsuccinate synthase in Thauera aromatica strain K172 *Arch Microbiol 177*, 132-8.

34. Toraya, T., Ushio, K., Fukui, S., and Hogenkamp, P. C. (1977) Studies on the mechanism of the adenosylcobalamin-dependent diol dehydrase reaction by the use of analogs of the coenzyme *J Biol Chem 252*, 963-970.

35. Yamanishi, M., Yunoki, M., Tobimatsu, T., Sato, H., Matsui, J., Dokiya, A., Iuchi, Y., Oe, K., Suto, K., Shibata, N., Morimoto, Y., Yasuoka, N., and Toraya, T. (2002) The crystal structure of coenzyme B12-dependent glycerol dehydratase in complex with cobalamin and propane-1,2-diol *Eur. J. Biohem. 269*, 4484-4494.

36. Tobimatsu, T., Sakai, T., Hashida, Y., Mizoguchi, N., Miyoshi, S., and Toraya, T. (1997) Heterologous expression, purification, and properties of diol dehydratase, an adenosylcobalamin-dependent enzyme of Klebsiella oxytoca *Arch Biochem Biophys 347*, 132-40.

37. Becker, A., and Kabsch, W. (2002) X-ray Structures of Pyruvate Formate-Lyase in Complex with Pyruvate and CoA *J. Biol. Chem. 277*, 40036-40042.

38. Reddy, S. G., Wong, K. K., Parast, C. V., Peisach, J., Magliozzo, R. S., and Kozarich, J. W. (1998) Dioxygen inactivation of pyruvate formate-lyase: EPR evidence for the formation of protein-based sulfinyl and peroxyl radicals *Biochemistry 37*, 558-63.

39. Golding, B. T., and Radom, L. (1976) The mechanism of action of adenosylcobalamin. *J. Amer. Chem. Soc. 98*, 6331-6338.

40. Masataka, E., Takashi, K., Yoshizawa, K., and Tetsuo, T. (2002) Theoretical study on the mechanism of catalysis of coenzyme B12-dependent diol dehydratase *Bull. Chem. Soc. Jpn. 75*, 1469-1481.