Distribution Agreement

In presenting this dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I agree that the Library of the University shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to copy from, or to publish, this dissertation may be granted by the professor under whose direction it was written, or in his absence, by the Dean of the Graduate School when such copying or publication is solely for scholarly purposes and does not involve financial gain. It is understood, that any copying from, or publication of, this dissertation which involves potential financial gain will not be allowed without written permission.

_____

Signature

Structural and Functional Studies on *Salmonella Typhimurium* Ethanolamine Ammonia-Lyase

By

Adonis Bovell
Doctor of Philosophy
Physics

_____
Dr. Kurt Warncke
Advisor

_____
Committee Member
Dr. Brian Dyer

_____
Dr. Laura Finzi
Committee Member

_____
Dr. James Nagy
Committee Member

_____
Dr. Connie Roth
Committee Member

Accepted:

_____
Lisa A. Tedesco, Ph.D.
Dean of the Graduate School

_____
Date

Structural and Functional Studies on *Salmonella Typhimurium* Ethanolamine Ammonia-Lyase

By

Adonis Bovell

B.Sc., Oglethorpe University, Atlanta, GA, 2006

Advisor:

Kurt Warncke, PhD.

An abstract of

A Dissertation Submitted to the Faculty of the

James T. Laney Graduate School of Emory University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

in Physics

2013

Abstract

Structural and Functional Studies on *Salmonella Typhimurium* Ethanolamine Ammonia-Lyase

By

Adonis Bovell

Ethanolamine ammonia-lyase (EAL), a coenzyme-B_{12} (AdoCbl) dependent bacterial enzyme, catalyzes the deamination of select amino-alcohols by using a radical mechanism. Extensive high-resolution spectroscopic determinations of reactant intermediate-state structures and detailed kinetic and thermodynamic studies have been conducted for the *Salmonella typhimurium* enzyme. A statistically robust homology model for the full $[(EutB\text{-}EutC)_2]_3$ oligomer of *S. typhimurium* EAL is constructed from the *Escherichia coli* crystal structure. This structure establishes a platform for detailed, microscopic interpretation of the molecular mechanism of EAL catalysis. The model is used to describe the hierarchy of EutB and EutC subunit interactions in the native oligomer and to guide a genetic and biochemical approach to the long-standing challenge of functional oligomer reconstitution from isolated subunits. The model is used to direct site-directed mutagenesis of EAL, leading to the creation of the EutB-F258W mutant, whose fluorescence is sensitive to the binding of AdoCbl. The AdoCbl-EAL dissociation constant is determined to be 1.2 μM, which places limits on the timescale of cofactor exchange kinetics. A series of cysteine-replaced mutants of EAL was created, and progress was made towards the goal of a mutant EAL for site-directed spin labeling studies. The primary cysteine attachment site in wild-type EAL for the 4-maleimido-TEMPO spin label was identified as EutC-C37. The localization of spin labels on EAL enables the interpretation of electron paramagnetic resonance (EPR) studies that probe distal effects on protein structure caused by cofactor binding. Previously determined rate constants for decay of the cryotrapped substrate radical, and k_{cat} values at ambient temperature, for ^{1}H- and ^{2}H-labelled substrate, are united in a single model that describes the sequential radical rearrangement and hydrogen atom transfer steps, from 190 to 295 K. The model indicates that hydrogen transfer proceeds via quantum mechanical tunneling, and accounts for the anomalous hydrogen isotope effects on turnover, previously observed in EAL, and other B_{12}-dependent enzymes.

Structural and Functional Studies on *Salmonella Typhimurium* Ethanolamine Ammonia-Lyase

By

Adonis Bovell

B.Sc., Oglethorpe University, Atlanta, GA, 2006

Advisor:

Kurt Warncke, PhD.

A Dissertation Submitted to the Faculty of the

James T. Laney Graduate School of Emory University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

in Physics

2013

Acknowledgements

I am indebted to my advisor, Prof. Kurt Warncke, for the guidance he provided throughout my development at Emory. His support enabled my maturation as a scientist, and made the research presented here possible. I also owe thanks to the members of my committee, Dr. Brian Dyer, Dr. Laura Finzi, Dr. James Nagy and Dr. Connie Roth. Their encouragement and thoughtful critique were instrumental in completing my dissertation. I must also thank all the members of the Warncke lab, including Dr. Li Sun, who cheerfully acquainted me with the techniques that I would need to use over the coming years.

I am also deeply appreciative of my mother, Adonia, my father, Allan, and the rest of my family. Their unconditional love and support allowed me succeed in even the most trying times. I also owe a debt of gratitude to Ngozi, and my closest friends. You all have made the experience over my years here at Emory an enjoyable one.

It would be remiss of me, if I did not acknowledge that there are a great number of people to whom I owe a great deal of thanks, whom, because of space considerations, I cannot explicitly mention. I dedicate this monograph to all of you. Without you, this would not be possible. Thank you.

Table of Contents

List of Tables and Figures

Chapter 1

Tables

Figures

Chapter 2

Tables

Figures

Figure 2. Comparison of secondary structure representations of EutB and EutC that show the non-conserved residues in the proteins from *S. typhimurium* and *E. coli*.

Figure 3. Ramachandran plot of homology modeled *S. typhimurium* EutB and EutC structures.

Figure 4. Surface and secondary structure representations of the homology modeled structure of *S. typhimurium* EAL.

Figure 5. PyMol ribbon representation of *E. coli* 3ABO template structure overlaid on *S. typhimurium* EAL modeled structure.

Figure 6. PyMOL ribbon representation of 3ABO crystal structure highlighting non-homologous sites between *E. coli* and *S. typh.*

Figure 7. Root mean square deviation of C-α atoms between *E. coli* and *S. typh.* EAL.

Figure 8. Contact surfaces of the EutB and EutC subunits in the EutB-EutC heterodimer, in the modeled structure of *S. typhimurium* EAL.

Figure 9. Properties of residues that form the binding contact face between EutB and EutC in the EutB-EutC heterodimer.

Figure 10. Charge distribution over the surface of the EutB and EutC proteins.

Figure 11. Positions of cysteine side chain sulfhydryl groups in the EutB and EutC subunits in the modeled structure of *S. typhimurium* EAL.

Figure 12. Sodium dodecyl sulfate polyacrylamide gel electrophoresis of purified proteins.

Figure 13. Sodium dodecyl sulfate polyacrylamide gel electrophoresis of EALH6 nickel affinity column purification fractions, for different imidazole elution concentrations.

Figure 14. Native polyacrylamide gel electrophoresis of purified proteins.

Figure 15. Gel filtration retention profiles for EAL and subunits.

Figure 16. Dependence of the activity of reconstituted EAL on the presence of different disulfide reducing agents.

Figure 17. X-band CW-EPR spectra of the cryotrapped cob(II)alamin-substrate radical pair in EAL in different protein samples.

Chapter 3

Tables

Figures

Chapter 4

Tables

Figures

Chapter 5

Tables

Figures

Chapter One

Introduction

1

1.1 Background

Ethanolamine ammonia-lyase (EAL) is a coenzyme B_{12}-dependent enzyme that uses the reactive power of radicals to activate an otherwise inert carbon-hydrogen bond and induce radical rearrangement. The biology and chemistry involved in this reaction has been studied for over 40 years, primarily in *Salmonella typhimurium*,[1] with a significant amount of early work being done on the clostridial enzyme.[2, 3] From a chemical perspective, the continued interest in this B_{12} enzyme arises because of the 10^{11} fold rate enhancement to cobalt-carbon bond cleavage that the enzyme achieves over the solution reaction,[4] and the concomitant creation of reactive radicals for catalysis. The potential to understand and co-opt such extraordinary rate enhancement into rationally designed enzymes is the holy grail of enzymology. There is a great opportunity to use evolutionarily selected solutions for difficult reactions as a starting point for the design of novel synthetic catalysts.[5] One can envision the next step, in which bio-inspired artificial enzymes are developed from a set of basic components found in nature. An example of this approach is the work on serine protease subtisilin, which has been extensively studied by Wells and others, and used as a scaffold for protein engineering.[6] However, more work is necessary to broaden the library of available molecular scaffolds and accessible reaction mechanisms to be used in biomimetic catalysts. To this end, a deep understanding of the nature of catalysis of EAL, and by extension, high-energy radical-mediated catalysis, is necessary.

The proteins involved in ethanolamine processing are also of biological significance. The ability conferred by EAL to process ethanolamine, and use it as a food and nitrogen

source is crucial to the proliferation of ethanolamine utilization bacteria such as *S. typhimurium* in vivo. This key competitive advantage is mostly present in food-poisoning bacteria, as it allows these bacteria to occupy specific ethanolamine rich niches within their hosts.[7] As such, the ethanolamine utilization pathway is being reexamined as a possible target for food formulas that suppress bacterial growth.[7]

Historically, the central tenet of protein science is the sequence to structure to function paradigm.[8] In short, a protein's 3-dimensional structure is wholly determined by its primary sequence, and this structure endows it with the ability to function. More recently, examples of functional, intrinsically unstructured proteins have arisen,[9] and the possible integral role of dynamics for function,[10-13] suggest that the relationship may be more complex than originally thought. Nonetheless, knowledge of the 3-dimensional structure of a protein is essential to begin to understand the molecular mechanism of an enzyme. As of now, the structure of *S. typhimurium* EAL has yet to be solved by high-resolution crystallography.

Further complicating studies on EAL is its low solubility.[14] It remains difficult to purify functional, soluble EAL for biophysical analysis.[15] Attendant to this challenging purification, is the difficulty in producing soluble mutant enzymes to further make directed probes of the molecular mechanism of EAL function.[16] Owing to these concerns, there has not been significant mutagenesis work on EAL. This is a significant limitation, as it restricts the use of an effective technique to study EAL catalysis.

The work described here addresses the lack of an EAL crystal structure by the development of a structural model. Informed by this model, a high-throughput

purification system could be designed and implemented, easing a major limitation to EAL mutagenesis. Using this new platform, new experiments on EAL become possible, shedding further light on EAL catalysis.

1.1.1 The Ethanolamine Ammonia-Lyase Reaction

EAL, discovered by Bradbeer in 1965[17] and historically known as ethanolamine deaminase,[18] catalyzes the deamination of the vicinal end-chain amino-alcohol ethanolamine into ammonia and acetaldehyde.[17] *(S)*- and *(R)*-2-aminopropanol, two other vicinal amino-alcohols, are also compatible, albeit slower, substrates for this enzyme.[17]

Figure 1 shows a schematic depiction of the native reaction catalyzed by EAL. The cleavage of ethanolamine into the aldehyde and ammonia is at its core, an isomerization. EAL catalyzes the exchange of the amine group on the second carbon with a hydrogen atom on the adjacent carbon. The substrate then rearranges to release ammonia, and form the corresponding aldehyde.

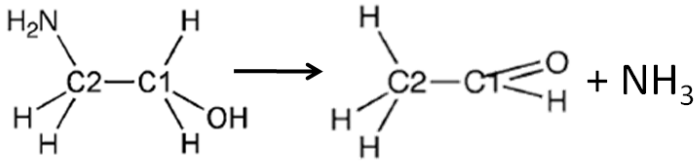


Figure 1. Native Reaction catalyzed by Ethanolamine Ammonia-Lyase.

1.1.2 Ethanolamine Utilization

The deamination of ethanolamine serves as an essential step in bacterial usage of ethanolamine, which can serve as the cell's sole source of carbon and nitrogen.[19, 20] Of

the two products, ammonia can be used directly as fixed nitrogen, and the other product, acetaldehyde, undergoes further processing by acetaldehyde dehydrogenase to convert it to acetyl-CoA.[21] Acetyl-CoA is a compound that is required for a host of other processes, including the Krebs cycle (aerobic energy generation), glyoxylate bypass (carbohydrate synthesis) and lipid biosynthesis. Alternatively, acetaldehyde may be converted to ethanol by alcohol dehydrogenase. This allows the use of ethanolamine as a carbon source, and may protect the cell from high levels of the toxic acetaldehyde intermediate.[22]

1.1.3 Genes Responsible for Ethanolamine Utilization

Genes for EAL, along with other genes required for ethanolamine utilization reside on a single *eut* operon. One comparative genomics study has revealed that *eut* genes are present in 84 fully sequenced bacterial genomes.[22] The identity of the genes present, and their organization vary, but a common thread is the presence of *eutB* and *eutC*, which are the genes responsible for the central enzyme of this pathway: the EAL protein.

In *Salmonella enterica* serovar Typhimurium (*S.typhimurium*), the strain at the focus of the present study, 15 other genes are also present on the *eut* operon.[23] These 17 genes and their functions are listed in Table 1. These genes include *eutH*, a cell membrane protein which may assist with the diffusion of ethanolamine into the cell. However, the absence of this gene does not significantly affect the functioning of the *eut* system.[23]

Table 1. Ethanolamine Utilization genes present in *Salmonella Typhimurium*.*

Gene†	Function
eutS	Carboxysome Structural Protein
eutP	Ethanolamine Utilization Protein
eutQ	Ethanolamine Utilization Protein
eutT	Corrinoid cobalamin adenosyltransferase
eutD	Phosphotransacetylase
eutM	Carboxysome Structural Protein
eutN	Carboxysome Structural Protein
eutE	Aldehyde Oxidoreductase
eutJ	Putative Chaperonin
eutG	Alcohol Dehydrogenase
eutH	Transport Protein
eutA	Reactivating factor
eutB	Ethanolamine Ammonia-Lyase Large Subunit
eutC	Ethanolamine Ammonia-Lyase Small Subunit
eutL	Carboxysome Structural Protein
eutK	Carboxysome Structural Protein
eutR	Transcriptional Regulator

* Data from Garsin[1] and Kofoid et al.[23]
† Genes listed in order of appearance on *eut* operon.

Also present on the *S. typhimurium eut* operon are *eutK*, *eutL*, *eutM*, *eutN* and *eutS* which encode proteins homologous to carboxysome shell proteins.[23] The enzymes involved in ethanolamine utilization are believed to localize within a bacterial microcompartment with a protein-based shell.[24] This protein shell is in contrast to organelles within eukaryotes which are constructed of lipid bilayers. By analogy, the extensively studied "carboxysome," an eponymous carbon-fixing microcompartment, serves to sequester and enrich CO_2 and facilitate fixation by Rubisco,[24] an enzyme notorious for its low specificity.[25] Similarly, the *eut* microcompartment is believed to assist in concentrating

the volatile intermediate, acetaldehyde, as well as protecting the cell from its toxic side effects.[26]

EutE encodes the second enzyme in the *eut* pathway, an oxidoreducatse (aldehyde dehydrogenase) that converts acetaldehyde into the useful acetyl-CoA. An alternate pathway for processing acetaldehyde is accomplished through the *eutG* gene which encodes an alcohol dehydrogenase. Further downstream processing of acetyl-CoA is accomplished by *eutD*, which encodes an enzyme responsible for converting acetyl-CoA to acetylphosphate, which leads to the production of ATP.

EAL requires the 5'-deoxyadenosylcobalamin (AdoCbl, coenzyme-B_{12}) cofactor in order to function. The *eut* operon includes *eutT*, which encodes an adenosyltransferase that can produce AdoCbl from the cobalamin. *eutA* encodes a reactivating factor that protects EAL from inhibition by the structurally similar cyanocobalamin (vitamin-B_{12}).

Other proteins encoded by the genes involved in ethanolamine utilization include the transcriptional regulator, EutR, the putative chaperone, EutJ, and two other proteins, EutP and EutQ of unknown function.

1.1.4 Biological Significance of Ethanolamine Ammonia-Lyase

Both mammalian and bacterial cell walls are rich in phosphatidylethanolamine, which is readily broken down by phosphodiesterases to give ethanolamine.[1, 27] Thus, cell wall composition and diet make the intestine an ethanolamine rich environment. The ability to use ethanolamine as an energy, carbon and nitrogen source confers an advantage for bacterial colonization of the gut. Several pathogenic bacteria, including *Salmonella*, use the ability to metabolize ethanolamine to survive in ethanolamine rich environments such

as in foods, and within their hosts.[7] Unsurprisingly, a strong correlation has been found between food-poisoning bacteria and the presence of *eut* genes.[28] Further, ethanolamine also serves as a signaling molecule and can act as the trigger for virulence in *Escherichia coli*. Thus virulent gene expression is controlled both through ethanolamine metabolism and ethanolamine signaling.[29]

Salmonella enterica accounts for 26% of food-borne infections within the U.S.[30] This arises mainly from contaminated egg products, with *S. typhimurium* being the major serotype in raw milk.[31] Thus, there has been a renewed interest in studying the *eut* operon, to find any role it may play as a possible target for the control and prevention of food-poisoning bacteria.

1.1.5 Structure of B_{12}

EAL is a member of a class of coenzyme-B_{12} dependent enzymes that use radicals generated from the homolytic splitting of the cobalt-carbon bond of coenzyme-B_{12} to achieve catalysis. B_{12} is unique in nature due to the presence of this metal-carbon, Co-C bond. Studies on B_{12} have a long history, spanning back to 1925 where evidence of its existence first originated, to 1948 when it was isolated as a red crystalline compound. The structure was discovered in 1956 by Hodgkin, and extensive study of this cofactor continues today.[32] B_{12} is, structurally, the most complex cofactor, and requires more than 20 genes for its synthesis,[33] and cannot be synthesized in humans.

Figure 2 shows the structure of B_{12} which can roughly be divided into three planes. In the central plane there is the cobalt containing corrin ring. The corrin ring resembles the porphyrin ring of heme compounds (they both originate from the same precursor), with

the exception that the corrin ring is more saturated and is missing a carbon connecting pyrrole rings in its structure. This modification leads to a smaller size, and has been speculated to be because of cobalt's smaller atomic radius when compared to iron; the metal bound by heme.[33] The cobalt is held in place by four corrin nitrogens in the equatorial plane. Usually, the cobalt is nominally in oxidation state III, and is hexa-coordinate, possessing axial ligands above and below the ring plane.

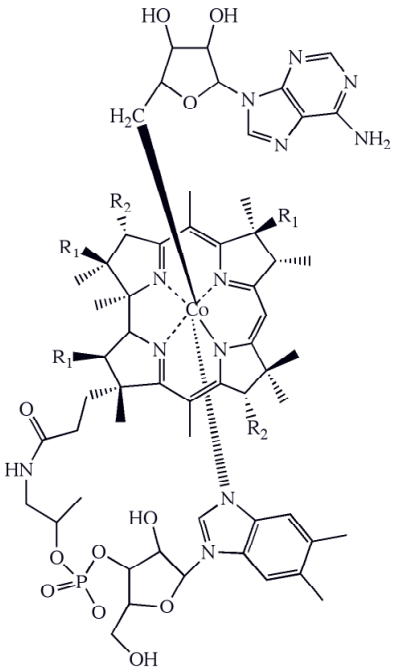


Figure 2. Structure of 5'-deoxyadenosylcobalamin.

In the plane below the corrin ring, there is the 5,6-dimethylbenzimidazole group (DMB). This group is a substituent of the corrin ring off of a propionamide side chains, and is referred to as the tail. The DMB tail also forms an axial alkyl bond with the cobalt. In the plane above the corrin ring, the 5'-deoxyadenosyl group from ATP binds axially to the cobalt, forming the weak metal-carbon bond (30 kcal/mol)[4] from which the radicals are

generated. The upper axial ligand may be substituted from the 5'-deoxyadenosyl group (in AdoCbl, coenzyme-B_{12}), creating different forms of B_{12}. Examples of ligands include -CN (cyanocobalamin, vitamin-B_{12}), the form of B_{12} that is usually ingested, $-CH_3$ (methylcobalamin) which is used by methyltransferases, and -OH (hydroxocobalamin).[34]

1.1.6 Classes of B_{12}-dependent enzymes

Enzymes using B_{12} fall into three broad categories, namely the isomerases, the methyltransferases, and the reductive dehalogenases. The isomerases, the class to which EAL belongs, catalyze 1,2 rearrangement reactions. These are reactions in which a hydrogen on one carbon is exchanged with a substituent on the adjacent carbon. In the case of EAL, this substituent is an amine ($-NH_2$). After this rearrangement, the product releases an ammonia which leads to the subcategorisation of EAL as a deaminase. Other enzymes of this class include dehydratases, such as glycerol and diol dehydratase, which exchange a hydroxyl group and lead to the release of water. Most isomerases are found in bacteria, with the exception of methylmalonyl-CoA mutase (MCM) which is present in humans. MCM is a carbon skeletal mutase, and the exchanged substituent is itself a carbon.[33] The only other B_{12}-dependent enzyme in humans is methionine synthase, a methyltransferase that requires methylcobalamin.[35]

Within the isomerases, there are two general subgroups, base-on and base-off, which refer to the binding mode of AdoCbl within the enzyme. In enzymes such as MCM, there is a binding motif (DX**H**XXG) in which a histidine replaces the DMB tail as the lower axial ligand when bound to the enzyme.[32] As such, these are categorized as base-off.

However, in EAL, this binding motif is not present and the enzyme has been shown to bind the ligand in the base-on conformation.

1.1.7 Mechanism of Action of Ethanolamine Ammonia-Lyase

Isomerases use radicals generated from the AdoCbl cofactor within their enzyme mechanism. This is in contrast to methyltransferases which use methylcobalamin, and do not proceed by a radical mechanism. The minimal mechanism for isomerases is general, with few exceptions. This minimal mechanism is shown in Figure 3 in the context of EAL catalyzing the reaction with ethanolamine.

The mechanism proceeds as follows. In short, after substrate binding to the holoenzyme the Co-C bond of AdoCbl splits homolytically, leaving an unpaired electron on the 5'-deoxyadenosyl group and on cobalt which forms low-spin Co(II). This deoxyadenosyl radical[36] migrates to the substrate where it abstracts a hydrogen from the C1 carbon of the substrate (HT1) creating a Co(II)-substrate radical pair.[37, 38] In a step termed radical rearrangement (RR),[39] the radical migrates to the adjacent carbon (C2) with the concomitant migration of the amine to C1, forming the product radical.[40] This product radical has never been directly observed, and its existence is only putative. The hydrogen on the deoxyadenosyl group is then abstracted by the product radical in the second hydrogen transfer (HT2). The deoxyadenosyl radical then migrates to the Co(II) and reforms the intact cofactor, and the acetaldehyde and ammonia products are released. The notable exception to this minimal mechanism is Ribonucleotide Reducatase (RNR), which proceeds using a thiyl radical.[34]

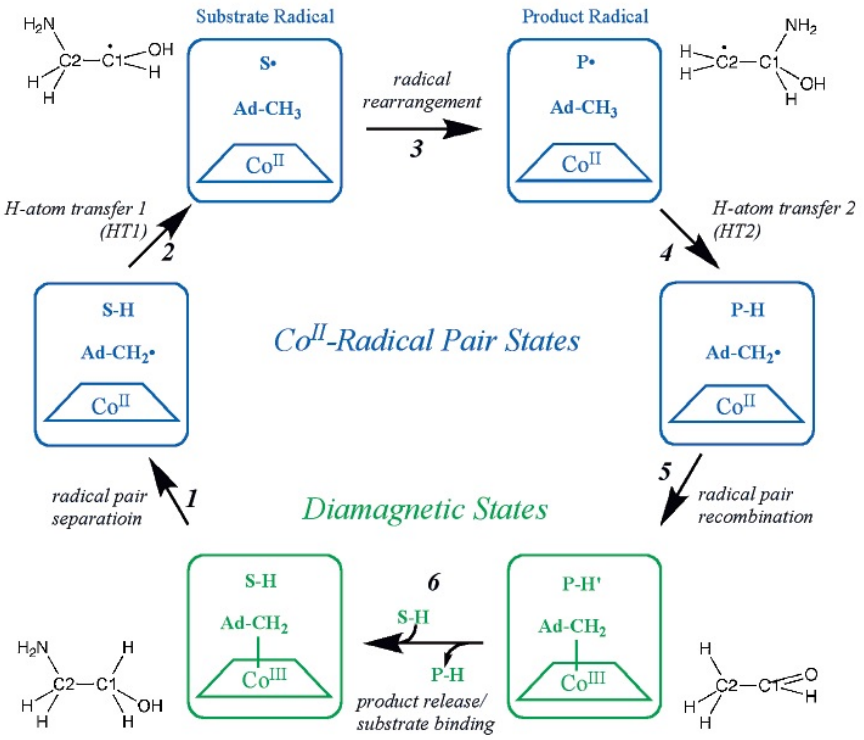


Figure 3. **Minimal Mechanism of EAL catalysis**. From bottom left: substrate binds to the holoenzyme and causes (1) Homolytic cleavage of Co-C bond and radical pair separation. (2) The 5'-deoxyadenosyl radical migrates to substrate and abstracts hydrogen atom (H-atom) to create EPR-detectable substrate radical state. (3) Amine group migrates from C2 to C1, concomitant with migration of radical from C1 to C2, to create product radical in radical rearrangement step. (4) Second H-atom transfer step; hydrogen is abstracted from 5'-deoxyadenosine, to form diamagnetic product and recreate the 5'-deoxyadenosyl radical. (5) 5'-deoxyadenosyl radical migrates back to the Co(II) and the radical-pair recombines. (6) Products are released and enzyme is ready for another reaction cycle. Note: the product radical shown in the upper right has not been observed to date, and its existence is putative.

1.1.8 Previous Structural Studies of Ethanolamine Ammonia-Lyase

EAL is an enzyme of large mass, composed of two subunits, EutB and EutC which are coexpressed from the *eutB* and *eutC* genes respectively.[41] These subunits are also referred

to in the literature as the α and β subunits respectively. The subunits form a hexameric structure (~500 kDa) of stoichiometry $EutB_6EutC_6$ in the active enzyme.[14] The larger subunit, EutB (453 residues, 49.4 kDa in *S. typhimurium*), contains the active site.[42] The other subunit, EutC, is smaller, containing 286 residues, with a mass of 32.1 kDa in *S. typhimurium.*

The 3-dimensional structure of the EutB subunit of *S. typhimurium* EAL was modeled[42] from the X-ray structure of the related isomerases, diol dehydratase and glycerol dehydratase. Later, the X-ray crystallographic structure for the EutB subunit of EAL in *Listeria monocytogenes* was solved to 2.15 Å resolution (Protein Data Bank (PDB) ID: 2QEZ). More recently, several structures of both EutB and EutC subunits of EAL from *E. coli* were solved with various ligands.[43] These structures contained cyanocobalamin, (PDB ID: 3ABR), cyanocobalamin and ethanolamine, (PDB ID: 3ABO), cyanocobalamin and 2-aminopropanol, (PDB ID: 3ABQ), and adeninylpentylcobalamin (a non-hydrolyzable B_{12} analog) and ethanolamine (PDB ID: 3ABS). The structure of *S. typhimurium* EAL remained unsolved.

Figure 4 shows the X-ray structure of EAL from PDB entry 3ABO.[43] The α-subunit contains a TIM-barrel fold which is the most common structure in the PDB.[44] It is characterized by eight β-strands surrounded by eight α-helices connected by loops. This barrel is structurally sound, and the loop regions at the end of the barrel can be extensively modified without significantly affecting the structural stability. For this reason it has been co-opted by nature to catalyze many diverse reactions.[44] The C-terminal end of the barrel contains the ethanolamine binding site, which is, in turn, covered by the AdoCbl cofactor. There is a 5'-deoxyadenosine binding pocket in the

EutB subunit providing orientation for cofactor binding.[42] The EutC subunit covers the AdoCbl, and the C-terminal end of the barrel of the EutB subunit in the complete enzyme.

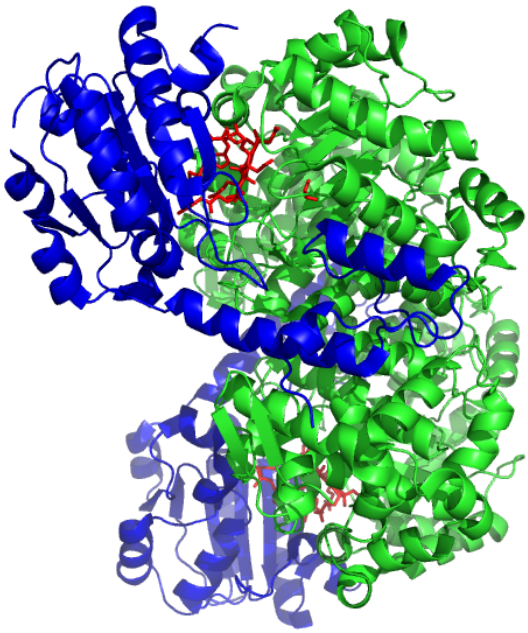


Figure 4. X-Ray Crystal Structure of *E. coli* EAL. Pymol representation of crystal structure of *E. coli* EAL $\alpha_2\beta_2$ dimer[43] shown with EutB subunits as green cartoon, EutC subunits as blue cartoon, cyanocobalamin and ethanolamine shown as red sticks.

1.2 Major Experimental Techniques Used in this Work

1.2.1 Site Directed Mutagenesis

In order to probe the mechanism of radical catalysis in EAL, many different experimental techniques have been used. These include many electron paramagnetic resonance techniques, (continuous wave,[45-47] ESEEM,[16, 37, 48] ENDOR[49]) which are sensitive to the radicals generated during catalysis, UV-visible absorbance and photolysis

measurements[50, 51] which can modify and probe the disposition of the AdoCbl cofactor, circular dichroism[52] (sensitive to structure), gel electrophoresis[14] and fluorescence. This list is by no means exhaustive. Similarly, a considerable repertoire of experimental handles exists that modify the system in a controlled manner. Most commonly, temperature-dependent studies are done. These can yield information about free energy surface on which the enzyme operates.[46, 51] In addition, techniques that take advantage of both temperature and time, such as cryotrapping,[45] and rapid-mixing, freeze-quench experiments,[53] can give kinetic information about EAL catalysis. Another probe of EAL behavior is magnetic field effect experiments[54, 55] which take advantage of the unpaired spins created during catalysis. Further, experiments have been performed where specific isotope labeling of substrates or the protein provides detailed information about molecular steps during the catalytic cycle.[47, 56, 57] This list is not exhaustive, as there are even more experimental probes, such as pH or hydrostatic pressure,[58] that allow for studies on EAL.

Along with the wide variety of experimental handles that can be used to probe EAL catalysis, the complementary approach of site-directed mutagenesis (SDM) can be used to augment the data available. SDM is a well-known molecular biology technique that enables controlled changes in the enzyme's primary structure to allow for the introduction of targeted probes of enzyme behavior. Prior to SDM, mutations were all naturally occurring, possibly after treatment with mutagens. The polymerase chain reaction (PCR) revolutionized this approach, allowing for targeted alterations of gene sequences, and thus increasing the efficiency of directed mutations from < 1% from phenotype screenings, to values between 50 and 100% achievable today.[59] The SDM

technique is amenable to all of the previously mentioned experiments, and further opens another dimension of data by allowing for targeted variations in the enzyme itself. Several examples of SDM being applied to EAL exist in the literature,[15, 16] but this technique is not prevalent, perhaps owing to the difficulty of working with a large, relatively insoluble, multi-subunit protein.

SDM by overlap extension using the PCR,[60] the main mutagenesis technique employed in this work, allows for the genetic engineering of recombinant proteins. Figure 5 shows an overview of the process. Briefly, this technique operates at the level of DNA, where two short oligonucleotide primers complementary to the gene of interest are synthesized with a mutation containing the desired modifications. Using PCR, these primers are extended and amplified to produce DNA that is identical to the parent DNA with the exception of the introduced mutations. The methylated parent DNA can be digested by *DpnI*, leaving the nascent engineered plasmid. These plasmids can be transformed into an overexpression system to produce mutated proteins.

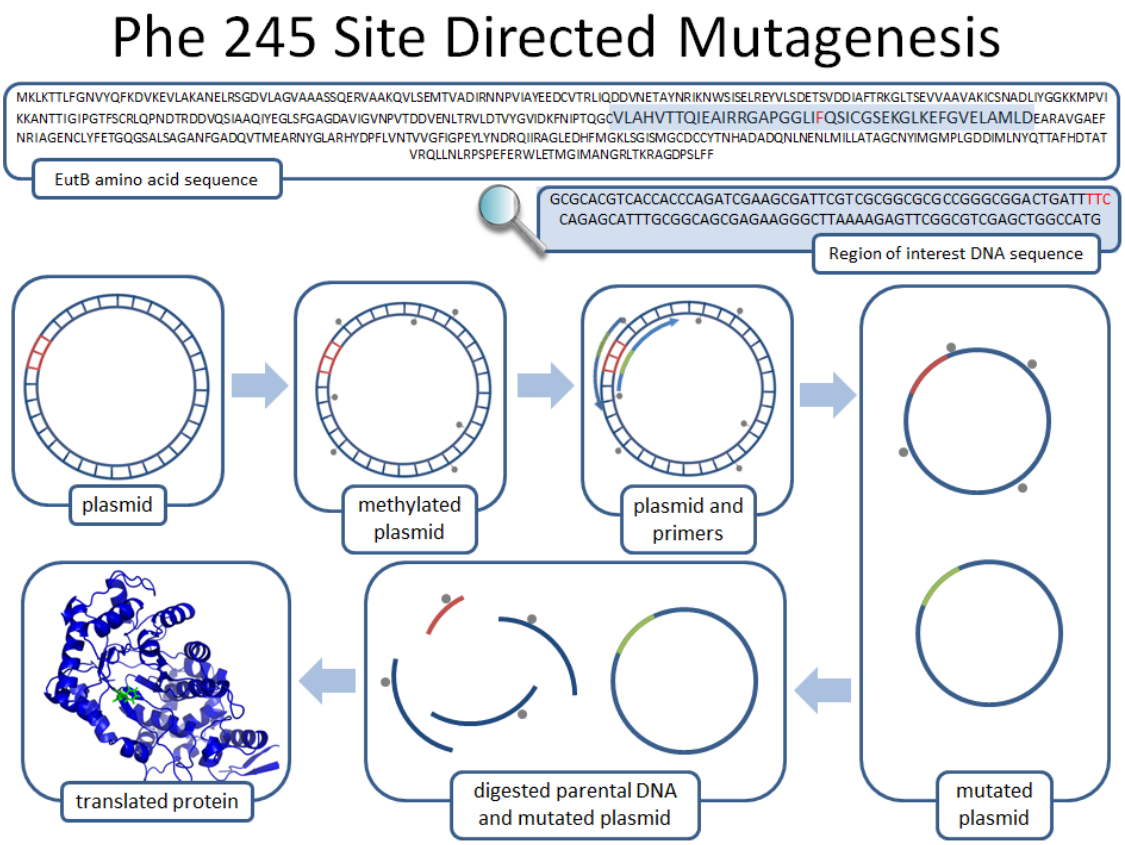


Figure 5. Site-Directed Mutagenesis Flow Chart. Starting from the amino-acid sequence of the protein (shown here is *S. typhimurium* EutB sequence, with αF245 highlighted in red), the DNA sequence around the targeted residue is analyzed, and primers complementary to this region are generated with the desired mutation (shown in green). Cells containing the parental DNA are grown in a *dam*$^{+}$ *E. coli* strain which methylates its DNA. The parental plasmid is then extracted and PCR is used to generate and amplify the mutated plasmid with staggered nicks. The parental DNA is then digested by *DpnI* which selectively targets methylated DNA, and the mutant plasmid can be transformed into expression cells to produce mutant protein.

SDM can also be accomplished by the direct synthesis of the entire gene without the use of a template. Advances in technology have made this approach plausible and accessible

at reasonable cost. Gene synthesis is preferable when a large number of point mutations are needed. For single site mutations, overlap extension is still used.

1.2.2 Electron Paramagnetic Resonance Spin Labeling

1.2.2.1 Electron Paramagnetic Resonance

Electron Paramagnetic Resonance (EPR) is the term for the resonant absorbtion of electromagnetic radiation by a paramagnet. First discovered by Zavoisky in 1945, EPR is a technique that can be used to study a host of biological, chemical and physical questions including Ångstrom level structure, and kinetics.[37, 45, 46, 61] From a classical perspective, resonance arises from Larmor precession of the electron spin magnetic moment in an external magnetic field. Quantum mechanically, it arises from transitions between non-degenerate Zeeman levels.

The central equation of EPR is

$$h\nu = g\beta B \qquad (1)$$

where h is Planck's constant, ν is the frequency of electromagnetic radiation, β is the Bohr magneton, and B is the strength of the external magnetic field. g is the g-value, and is equal to 2.0023 for a free electron.[61] Equation 1 describes when magnetic transitions are possible; the situation where the energy of the applied radiation is equal to the separation between the Zeeman manifolds, which is in turn controlled by the strength of the applied magnetic field. Table 2 shows resonance frequencies of a $g\approx2$ radical for typical fields.

Table 2. Resonance Fields for a g = 2 radical at various microwave frequencies.*

Microwave Band	Microwave Frequency (GHz)	Magnetic Field (G)
L	1	390
S	2	1,070
X	9	3,380
K	24	8,560
Q	35	12,480
W	94	33,600

* Data from Eaton and Eaton.[62]

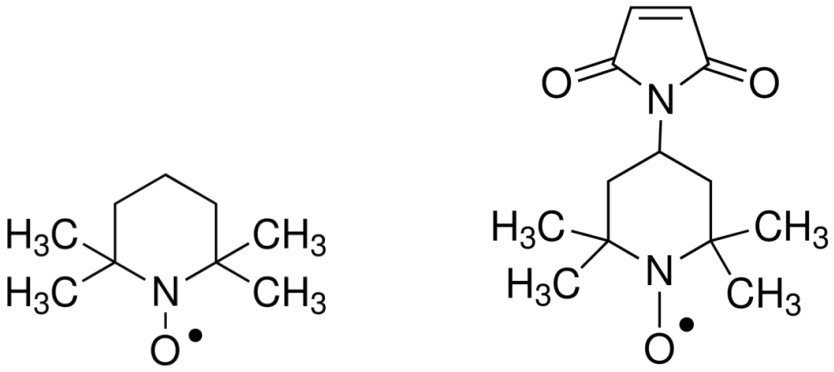


Figure 6. Structures of TEMPO and 4-maleimido-TEMPO. Spin probe TEMPO shown on left. Spin label 4-maleimido-TEMPO shown on right. Structures from Sigma-Aldrich.

For a more complex system, such as the TEMPO spin probe, (Figure 6), additional local interactions shift the position of resonance. TEMPO possesses an unpaired electron (S=1/2) on the oxygen (I=0) that is shared with the adjacent nitrogen (I=1). The electron possesses some spin density in the *s*-orbital, which has a non-zero probability density at the nucleus and thus interacts with the nuclear spin of the nitrogen. The three possible configurations of the nitrogen nuclear spin (m_I = -1, 0, 1) lead to three different energy

separations in the Zeeman levels. For traditional field-swept continuous-wave (cw) EPR, the selection rules are such that $\Delta m_S = \pm 1$, and $\Delta m_I = 0$, so only three distinct absorbances are possible. More generally, each resonance peak in the sample will be split $2nI + 1$ times, where n is the number of symmetry equivalent nuclei, and I is the nuclear spin. The degree of the energy level shift caused by interaction with the nuclear spin is controlled by the magnitude of the electron-nuclear interaction, known as the isotropic hyperfine coupling, A_{iso}. The electron also possesses spin density in the *p*-orbital. Here, the strength of the electron-nuclear interaction depends on the angle of orientation of the *p*-orbital compared to the external field.[61] This leads to local fields, all of which may vary slightly between otherwise equivalent spin probes. The end result is a multitude of slightly shifted individual lines that combine to give broadened EPR spectra. Figure 7 shows the resulting manifold from these 4 dominant interactions and their functional forms.

In solution samples, the motion of the spin probe has distinct effects on the resulting EPR spectrum. The spectral anisotropy of a system, $\Delta\omega$, is the maximum shift in the resonance positions derived from orientation changes. If the nitroxide can reorient on a rapid timescale compared with $\Delta\omega$, this anisotropy is averaged out leading to an isotropic spectrum. Table 3 shows the different dynamic regimes of spin probe motion. Note that $\Delta\omega$ depends on the frequency/field of the measurement. For X-band, $\Delta\omega/2\pi \approx 130$ MHz, which makes correlation times, τ_C, of 1 μs to 10 ps in the fast motion regime.[63]

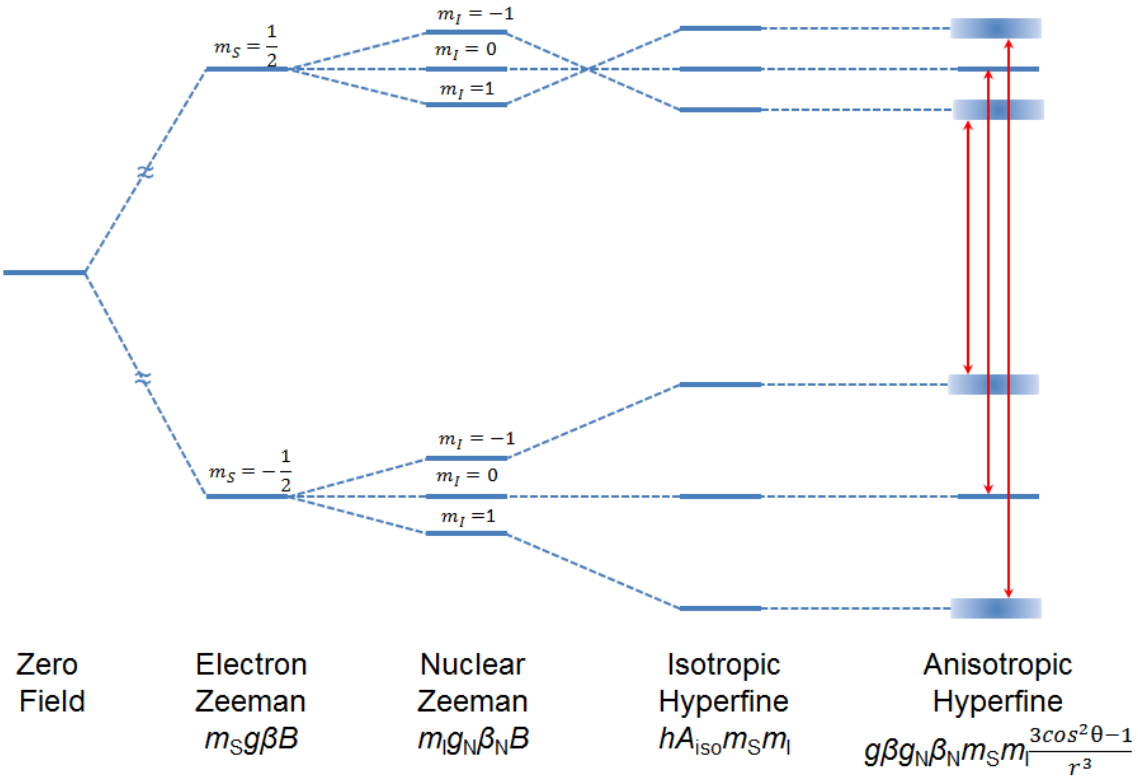


Figure 7. Energy Level Diagram of EPR transitions in a nitroxide. From left to right energy levels for a nitroxide electron in an applied field shown. Diagram is illustrative and not to scale. Functional form of interaction shown below. EPR-allowed transitions shown in red.

Table 3. Dynamic Regimes of Nitroxide Motion and Theoretical Approaches to Simulation.*

Dynamic Regime	Scale† $\tau_C^{-1}/\Delta\omega$	Theory
Isotropic limit	> 500	Breit-Rabi formula
Fast motion regime	1 to 500	Redfield theory
Slow motion regime	0.005 to 1	Stochastic Liouville equation
Rigid limit	< 0.005	Energy level modelling

* Data obtained from Stoll et al.[63]

† τ_C is the rotational correlation time of the spin probe, $\Delta\omega$ is the spectral anisotropy.

This fast motion dynamic regime is described by Bloch-Wangness-Redfield theory, BWRT.[64] Briefly, this theory uses the system Hamiltonian, which includes the commutable spin (S) and lattice (F) Hamiltonians, and the non-commuting coupling between them. Lattice fluctuations serve as a bath and these fluctuations couple to the spin system to dissipate the absorbed microwave energy. This theory predicts spin-lattice and spin-spin relaxation times, T_1 and T_2, respectively.

T_2 determines the width of Lorenztian lineshape of the absorbance spectrum. BWRT predicts this relaxation time as shown in Equation 2.[63]

$$\frac{1}{T_2} = A_0 + A + Bm_I + Cm_I^2 \qquad (2)$$

A, B and C describe the broadening from rotational motion, whereas A_0 collects broadening from other sources. A, B and C depend on both the anisotropic hyperfine coupling and the g-anisotropy. A and B are also functions of the external field, which leads to the high-field line of the nitroxide EPR spectrum being the most sensitive to correlation time, displaying greater broadening than the lower field peaks.

1.2.2.2 EPR Experimental Concerns

One determinant of EPR signal intensity is the difference in populations between m_S = +½ and -½ manifolds. This population difference is, in turn, determined by the Boltzmann distribution (Equation 3).

$$\frac{n_{+1/2}}{n_{-1/2}} = e^{-\Delta E/RT} \qquad (3)$$

Here, $n_{+1/2}$ and $n_{-1/2}$ are the number of spins in the upper and lower manifolds, ΔE is the energy level separation of each manifold, R is the gas constant and T is the absolute temperature. For typical magnetic fields used in EPR (3,000 G), the energy level separation of a g=2 radical is around 9 GHz (X-band, see Table 2). This leads to relative population levels near to unity for room temperature experiments. However, conducting experiments at lower temperature increases the Boltzmann population of the lower manifold, and increased sensitivity is achieved.

The experimental difficulty of sweeping the electromagnetic frequency over a significant range leads to fixed frequency experiments, where the magnetic field is varied to bring the sample into resonance. Such a setup has been named continuous-wave (cw) EPR. Similarly, for signal to noise concerns, phase sensitive detection is employed.[62] With this technique real signals can be distinguished from noise by the known modulation frequency. This yields the derivative spectra so commonly associated with EPR rather than the absorbance spectra expected.

1.2.2.3 Site-Directed Spin-Labeling

1.2.2.3.1 Attachment of Spin Labels to Proteins

Site-Directed Spin-Labeling (SDSL) is a technique that combines SDM with EPR and allows site-specific labeling of a protein with EPR-active nitroxide probes. The majority of nitroxide probes are stable, chemically synthesized compounds that are protected from reaction by α-carbon substituents,[61] such as the methyl groups in the TEMPO spin probe shown in Figure 6. Spin labels combine these stable organic radicals (e.g. TEMPO) with a group that permits chemical attachment to protein under study.

Most naturally occurring amino acids are chemically inert and cannot be used for protein cross-linking. Table 4 shows a list of the reactive amino acids, and the possible chemical reactions. Of those that are reactive, hydrophobic residues such as tryptophan and tyrosine are usually buried on the interior of proteins, and are not usually chosen for conjugation. Cysteine, with its sulfhydryl moiety, is the most reactive group in proteins.[65] Its low abundance makes it the usual target for conjugating spin labels with proteins. The coupling reaction is either done by alkylation, or disulfide interchange.[66]

One common cysteine conjugation group is the maleimide. In the reaction of sulfhydryls with maleimides, an example of an alkylation reaction, the maleimide double bond reacts to give a stable thioether bond as the cross-link. This reaction is specific for thiols at pH 6.5 to 7.5.[66] Figure 6 shows the structure of the 4-maleimido-TEMPO spin label (4MT), which is the EPR-active nitroxide TEMPO conjugated with the maleimide functional group.

Methane-thiosulfonate spin-label (MTSL) is another commonly used EPR spin label. MTSL contains an S-S bond and conjugates with proteins by a disulfide interchange mechanism. In this reaction a free thiol (-SH) group on the protein attacks the disulfide group (on the label), causing disulfide interchange leaving the protein and spin label linked by a disulfide bond. This reaction is reversible and the spin label can be released using reducing agents such as dithiothreitol (DTT).

1.2.2.3.2 Information Gleaned from Spin Labels Attached to Proteins

Once the spin label is attached to the protein it can report on the local environment of the spin probe. Three main classes of information are accessible; (1) the dynamics of the spin

label side-chain, (2) the solvent accessibility of the side chain and (3) the distance of the side chain from a second nitroxide or bound metal ion.[67]

The relationship between the dynamics of the spin label side-chain and the underlying protein motion is complex because the nitroxide is free to rotate about the bonds linking it to the protein. Nonetheless, detailed studies have been done where EPR lineshape has been directly assigned to structural motifs such as Hubbell's work on T4 Lysozyme with MTSL.[68] In that study spectral features could be used to identify the location of the spin label such as buried within the protein, involved in tertiary contact interactions, on surface helix residues or in loop regions.

Solvent accessibility studies using spin labels have relied on detecting the collisional quenching of the spin label by paramagnetic reagents in the solution. In a cw-EPR experiment, increased microwave power can saturate the resonance signal leading to reduced signal amplitude. Collision with a paramagnetic quencher provides the nitroxide with another relaxation pathway, and this can be seen as an increase in the power required for saturation. Access to the collisional quencher is determined by the solvent accessibility of the spin label and thus this method can be used to experimentally determine solvent accessibility. Two commonly used paramagnetic quenchers are molecular oxygen and Ni(II)ethylene-diaminediacetate (NiEDDA). An example of structural information gleaned from this model includes the periodic solvent accessibility curve from a surface exposed α-helix in cellular retinol binding protein.[67]

It is also possible to make distance measurements between pairs of spin-labels. Nitroxides can yield distance information when they are within 8 and 25 Å of each other.

At this range, the dipolar interaction dominates the *J*-coupling and the splitting of the absorption lines (*2B*) have a known distance dependence (Equation 4).[69]

$$2B = \left(\frac{3}{2}\right) g_e \beta \frac{(3cos^2\theta - 1)}{r^3} \qquad (4)$$

g_e is the electron g-value, θ is the angle between the interspin vector and the external magnetic field, and r is the separation between the spins. Thus dipolar broadening can be deconvoluted to achieve distance information. This range can be extended to 18 to 80 Å if using the pulsed EPR technique, Double Electron-Electron Resonance (DEER).[70]

Another feature of spin labels is the ability to observe the time-dependent changes in any of the parameters discussed above, with up to millisecond resolution.[71, 72] Thus, SDSL allows for a complementary set of time-dependent data, augmenting the other experimental techniques that have been used to study EAL.

With the broad array of information available, SDM and SDSL are very practical tools for the study of enzyme structure and dynamics. These tools have been utilized with success on the study of EAL, as will be described in subsequent chapters.

1.3 Overview

Herein is described a computational homology model of the 3-dimensional structure of *S. typhimurium* EAL. This model is built off of the recently solved, X-ray crystal structure of the closely related *E. coli* EAL,[43] and supersedes an earlier homology model.[42] The modeled structure can be used to interpret high-resolution data from multiple sources, and give insights as to which specific residue contacts and hydrogen bonds are important for catalysis on a molecular level. Using this structure as a guide, a protocol for

purification and reassembly of the individual subunits of EAL into the functional oligomeric enzyme is described. This experimental study serves as both a verification of the computational modeling, and as a platform to allow high-throughput mutagenesis on EAL.

Protein from this newly developed expression and purification system is characterized by numerous experimental techniques, and found to retain wild-type like parameters. One study described within, focuses on the binding of the B_{12} cofactor to EAL, which was assessed using tryptophan fluorescence quenching. Previous attempts to use this technique on EAL were unsuccessful, being stymied by the inner filter effect, and the low efficiency of quenching by the cofactor. Using a mutant EAL with a targeted phenylalanine to tryptophan mutation, these difficulties were overcome, allowing for a direct measurement of cobalamin binding.

To probe the link between dynamics and catalysis, work was done to engineer targeted cysteine labeling sites within EAL. However, *S. typhimurium* EAL contains 16 native cysteine residues which had to be removed to achieve directed labeling. Informed by structural constraints, cysteine residues were selected for removal and several new mutant constructs were created. While directed labeling could not be achieved, several important features about EAL structure were elucidated, including the order of reactivity of the most reactive native cysteine residues.

Further work on EAL function studied the native reaction of EAL with isotope labeled substrates from cryogenic to room temperatures. A single model was able to describe the data over a > 100 K temperature range, for both the heavy and light substrates. This data

revealed that both radical rearrangement and hydrogen transfer 2 contribute to rate-determination. Moreover, hydrogen transfer 2 proceeds by hydrogen tunneling. With this framework it is possible to further understand the kinetics of EAL catalysis and the role of entropy in B_{12}-dependent enzyme catalysis.

Chapter Two

The Structural Model of *Salmonella typhimurium* Ethanolamine Ammonia-Lyase Directs a Rational Approach to the Assembly of the Functional $[(EutB\text{-}EutC)_2]_3$ Oligomer from Isolated Subunits[73]

The work presented in this chapter is available at

http://pubs.acs.org/doi/abs/10.1021/bi301651n

2

2.1 Introduction

Ethanolamine ammonia-lyase (EAL)[74, 75] is a 5'-deoxyadenosylcobalamin (AdoCbl; coenzyme B_{12}; Scheme 1) –dependent bacterial enzyme [EC 4.3.1.7; cobalamin (vitamin B_{12})-dependent enzyme superfamily[76]],[18, 34, 77] that catalyzes the deamination of the short-chain vicinal amino alcohols, aminoethanol and [*S*]- and [*R*]-2-aminopropanol.[17] The coding sequence for EAL is located on the 17-gene *eut* operon[23] which contains coding regions for the proteins required for bacterial utilization of ethanolamine as a carbon and nitrogen source, in the presence of exogenous AdoCbl.[19, 20, 26] Studies of the eut metabolosome sub-organellar structure have focused on molecular understanding of the hierarchical assembly,[78] and the complex chemistry that is involved in ethanolamine processing.[1] Ethanolamine utilization has been identified as a key competitive advantage for colonization of the mammalian gut by bacterial pathogens, and ethanolamine can serve as a trigger for virulence.[7, 29] A strong correlation has been found between food-poisoning bacteria and the presence of eut genes.[28]

EAL, the central enzyme in ethanolamine utilization, is a "radical enzyme," which harnesses the high reactivity of electron-deficient species to perform catalysis.[79] Previous detailed studies of EAL structure and mechanism have elucidated the geometry of the reactant centers in the active site,[37, 48, 49, 80] and determined the kinetics of culled steps in the reaction cycle of EAL,[45-47, 53, 81] in the native enzyme from *Salmonella typhimurium*. X-ray crystallographic structures of the complete EAL from a different bacterium, *Escherichia coli*, including bound cofactor and substrate have been obtained,[43] supplanting an earlier incomplete structure of EutB from *Listeria monocytogenes* (PDB

entry 2QEZ) and a model for *S. typhimurium* EutB based on secondary structure-prediction.[42] Further advances toward elucidating the molecular mechanism of EAL require a high-resolution model for the *S. typhimurium* protein, for structure-function correlations, and development of methods for efficient site-directed mutagenesis of EAL. Progress has been thwarted by the large mass (12 subunits, ~500,000 g/mol) and low solubility (<2 mg/ml) of the EAL oligomer,[14] and individual subunit instability. Here, we present a robust, homology-modeled structure of EAL from *S. typhimurium*, which is based on the *E. coli* structure, and use it to develop and rationalize approaches for expression and reconstitution of functional EAL from purified, individual protein subunits.

EAL is composed of two subunits, coded by the *eutB* and *eutC* genes. The larger subunit, EutB (453 residues; 49.4 kDa; *S. typhimurium*), is a $(\beta\alpha)_8$, TIM-barrel protein,[82] that contains the substrate binding cavity at the C-terminal end of the β-barrel.[42] The substrate binding cavity is capped by the AdoCbl binding site. The cofactor site and C-terminal end of the β-barrel are covered by the smaller subunit, EutC (298 residues; 32.0 kDa; *S. typhimurium*).[43] The EutB-EutC heterodimer is the fundamental structural unit of EAL. In the $EutB_6EutC_6$, EAL oligomer (488 kDa; *S. typhimurium*), two EutB-EutC heterodimers are organized as a $(EutB\text{-}EutC)_2$ homodimer, and three homodimers are organized into a $[(EutB\text{-}EutC)_2]_3$ trimer. The native *eutb* and *eutc* are under control of the same promoter, and are co-expressed *in vivo*.

The large molecular mass of the EAL oligomer and associated low solubility[14] is linked to the quasi-solid state of the enzyme in the eut metabolosome sub-organelle, *in vivo*. *In vitro*, the high aggregation propensity interferes with expression and purification yields of

correctly folded, functional EAL (owing to inclusion body formation) during expression of plasmids that include the cloned *eutb-eutc* coding region, in *E. coli* vectors. Techniques employed for growth of bacterial in which EAL is overexpressed have used systems designed to lower the rate of gene expression, as follows: (a) Growth and expression under non-inducing conditions,[41] (b) use of the BD21 pLysS system,[83] and (c) low temperature growth.[43] Previous attempts at the alternative approach of reassembling functional EAL from its separately expressed and purified subunits, EutB and EutC, have been unable to recover activity.[14] This may, in part, be related to the observed instability of isolated EutC. The Instability Index[84] of EutC is 51.4 (a score >40 connotes high instability), compared with 27.0 for EutB. Consistent with these predictions, EutC is subject to rapid trypsinization.[15] In addition, the N-terminal region of EutC may drive the formation of inactive EAL aggregates.[15]

We generate a multi-scale homology model structure of *S. typhimurium* EAL, which describes the molecular structure of the subunits, including the active site interactions of side chains, substrate and cofactor, and the global subunit interactions in the oligomer, that is based on the X-ray crystallographic structure of *E. coli* EAL.[43] The homology model predicts that a EutB hexamer can form in the absence of EutC, and that cystine disulfide bond formation from cysteine sulfhydryl groups exposed in the isolated subunits will promote protein aggregation. Model predictions in concert with biochemical manipulations lead to development of a low-temperature growth and expression system in NEB T7 Express I^q *E. coli*, that allows the individual expression of soluble EutB and EutC. Using tandem affinity and size exclusion purification, the EutB subunit can be

purified to homogeneity, and EutC can be enriched to >95%. Judicious application of sulfydryl reagents allows the reconstitution of active EAL.

Overall, the EAL structure model predictions inform the development of a novel sequential assembly protocol for reconstitution of the functional *S. typhimurium* EAL oligomer from individual subunits, which is verified by biochemical, spectroscopic, and functional studies.

2.2 Material and Methods

2.2.1 Homology Modeling of *S. typhimurium* EAL

EAL sequence alignment was performed with the ClustalW2 software[85, 86] to map the *S. typhimurium* EAL sequence onto the *E. coli* EAL sequence. The complete alignment is given in Figure 1. Residues 4 to 43 of the N-terminal end of the EutC subunit of the *E. coli* EAL were truncated for crystallization.[43] Therefore, these residues are not present in the X-ray crystallographic structure.[43] As a result, the modeled structure also does not contain the first 46 residues of EutC. Five homology models of the $[(\text{EutB-EutC})_2]_3$ functional oligomeric structure were generated by using this alignment and the 3ABO *E. coli* crystal structure with the MODELLER program.[43, 87] MODELLER generates homology modeled protein structures by using comparative modeling based on spatial restraints. MODELLER minimizes an effective energy function, which is based not only on alignment to the template structure, but also on statistical experimental restraints from similar structural features, such as main chain bond lengths and dihedral angles from equivalent groups. In addition to these homology-derived restraints, MODELLER also takes into account stereochemical restraints using the CHARMM22 force-field.

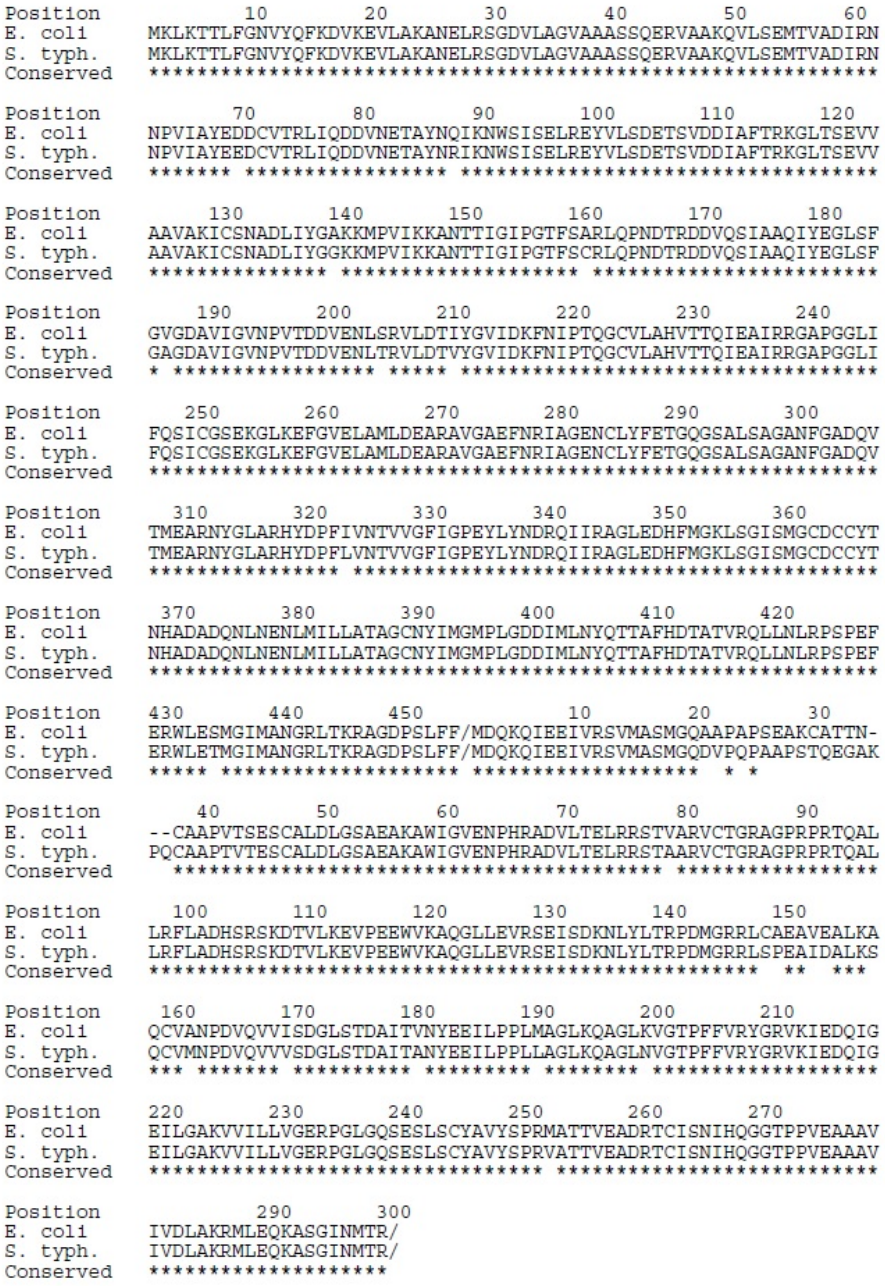


Figure 1. Amino acid sequence alignment for *E. coli* and *S. typhimurium* EAL. *EAL* Sequence alignment was performed by using the ClustalW2 software[85, 86] Junction at end of EutB (intial residues 1-453) and beginnig of EutC is denoted by "//." Conserved residues are denoted by "*." Gaps in the shorter *E. coli* EutC sequence are denoted by "–."

The models generated were first assessed by using the Discrete Optimized Protein Energy (DOPE) profile.[88] The DOPE profile allows quick evaluation of a putative structure by distinguishing high energy regions within a structure, which may indicate possible errors. However, it is not an absolute measure and is only useful to rank models generated from the same alignment. The structure with the lowest DOPE score was selected for further analysis. The modeled structure was validated by using the PROCHECK test suite.[89] PROCHECK tests the stereochemical quality of protein structures, by assessing main-chain parameters such as bond lengths, bond planarity and Ω, Φ and Ψ bond angles, in addition to side-chain parameters such as χ_1, χ_2, χ_3 and χ_4 bond angles. The structure was visually examined by using SYBYL 8.0 (Tripos Inc., St. Louis, MO) and PyMol (Delano Scientific LLC, San Fransisco, CA).

The solvent accessibility of residues in the modeled structure was judged using the ASAView web server.[90] Briefly, ASAView uses the DSSP program[91] to calculate the surface area of each amino acid residue that is accessible to solvent, and scales this solvent accessible surface area by the total area of that residue. Relative solvent accessibilities for EutB and EutC monomers, the EutB-EutC heterodimer, $(EutB\text{-}EutC)_2$ homodimer, and $[(EutB\text{-}EutC)_2]_3$ oligomer were computed, and compared to assess the changes upon hexamer formation.

The electrostatics of the modeled protein surface were calculated by using the Adaptive Poisson-Boltzmann Solver, APBS.[92] The PDB2PQR web server,[93, 94] a tool that automates the conversion of protein structural coordinates (PDB file) into files containing the calculated charge information (PQR), was used. PDB2PQR builds and assesses the hydrogen bonding network of the protein structure, and estimates protonation states,

charge and atomic radii for the biomolecule. APBS uses this input information to solve the Poisson-Boltzmann equation in a parallel manner, allowing the evaluation of the electrostatic potentials for the modeled structure. The results were visualized in PyMol.

2.2.2 Protein Library

A search of the Pfam database[95] yielded 693 sequences for EutB and 689 for EutC. A multiple sequence alignment of these sequences was downloaded, and examined with the MATLAB Bioinformatics Toolbox (Mathworks Inc., Natick, MA). In this alignment, EutB residues 11 to 452, and EutC residues 63 to 297 of *S. typhimurium* were aligned. At each position in the alignment, the most frequent residue was used to build a consensus sequence, which was used to assess the degree of conservation.

2.2.3 Construction of EALH6 Plasmid

The 8.5 kbp recombinant wild-type DNA coding for the EutB and EutC subunits of EAL from *S. typhimurium* in the *pBR322* vector was extracted from the *E. coli* overexpression strain.[41] Primers coding an N-terminal $(histidine)_6$ tag complementary to the *eutB* starting sequence were obtained from Invitrogen (Life Technologies Corp., Carlsbad, CA), as follows: **EutBH6 FP** – GGG GAA CGA CTT ATG CAT CAC CAT CAC CAT CAC AAA CTA AAG ACC ACA TTG, and **EutBH6 BP** – CAA TGT GGT CTT TAG TTT GTG ATG GTG ATG GTG ATG CAT AAG TCG TTC CCC.

Mutagenesis was carried out by using Stratagene (Stratagene, La Jolla, CA) QuikChange II XL Site-Directed Mutagenesis kit according to manufacturer's instructions. Briefly, the wild-type EAL plasmid was used as a template for PCR amplification of the $(histidine)_6$ primers with high fidelity *PfuUltra* DNA polymerase. The endonuclease,

DpnI, was used to digest the methylated template DNA, leaving the new EALH6 DNA intact. The new plasmid was then transformed into XL10 Gold ultracompetent cells (Stratagene, La Jolla, CA), and the sequence was verified by DNA sequencing (performed by Beckman Coulter Genomics, Danvers, MA).

2.2.4 Construction of EutBH6 and EutCH6 Plasmids

2.2.4.1 Preparation of pET28a Plasmid

An ampicillin resistant *pET28a* plasmid with a $(histidine)_6$ tag was used. The plasmid was digested in the multiple cloning site polylinker with the restriction enzymes *XhoI* and *NdeI* from New England Biolabs (Ipswich, MA), and the 5 kbp cleaved sticky-ended DNA fragment was purified from a 1% agarose Tris-Acetate-EDTA (TAE) gel by using the QIAquick gel extraction kit (QIAGEN, Venlo, Netherlands).

2.2.4.2 Preparation of EAL gene

The wild-type EAL *pBR322* sequences did not contain any *NdeI* (CA'T ATG) and *XhoI* (C'TC GAG) restriction sites. Primers were purchased from Integrated DNA Technologies (Coralville, IA) to insert these restriction sites to the start and end of the EutB and EutC genes separately, as follows: **EutBNdeI FP** – AAA AAA AAA A CAT ATG AAA CTA AAG ACC ACA TTG TTC GGC AAT G; **EutBXhoI BP** - AAA AAA AAA A CTC GAG TCA GAA GAA CAA TGA CGG ATC GCC CGC CCG TTT GG; **EutCNdeI FP** - AAA AAA AAA A CAT ATG GAT CAA AAA CAG ATT GAA GAA ATT GTA CG; **EutCXhoI BP** - AAA AAA AAA A CTC GAG TTA ACG GGT CAT GTT GAT GCC GGA CGC TTT. PCR was performed with *Pfu* DNA polymerase, and the *eutB* (1.5 kbp) and *eutC* (0.9 kbp) genes were purified from the result by

electrophoresis on a 1% agarose TAE gel. DNA was extracted from the gel, and then digested with the restriction enzymes, *XhoI* and *NdeI*, to obtain sticky ends complementary to the *pET28a* plasmid. These sticky ended genes were purified again using agarose gel electrophoresis.

2.2.4.3 Construction of Final Plasmids

Sticky-ended *pET28a* and *eutB* or *eutC* were ligated by using T4 DNA ligase. These constructs were transformed into XL10 Gold ultracompetent cells. The plasmid was then transferred into T7 Express I^q competent cells (New England Biolabs, Ipswich, MA), because these cells were found to be optimal for gene expression. The sequences of the EutB and EutC transformants were verified by Beckman Coulter Genomics, and were found to have 100% sequence identity between the restriction sites.

2.2.5 Bacterial Growth and Protein Purification

2.2.5.1 Growth

Cells were grown from glycerol stocks in 5 mL 2X YT starter cultures with 100 ug/mL ampicillin at 37 °C (30 °C for EutB, and EutC) overnight with shaking. This was used to inoculate 500 mL medium which was allowed to grow until OD_{600}~0.8. Following induction by 0.4 mM IPTG, cells were grown for 5 h, and then harvested by centrifugation at 3,000×*g* at 4 °C. Cells were washed with 10 mM potassium phosphate (KP_i) buffer (pH 7.5), and the pellet was frozen in liquid nitrogen and stored at -80 °C.

2.2.5.2 Purification

Cells were thawed and resuspended in 5 mL of lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, pH 8.0), which was supplemented with 1 mg/mL lysozyme, 50 uL 100x HALT protease inhibitor cocktail (ThermoFisher Scientific, Waltham, MA), and 2 μL Benzonase Nuclease (Sigma-Aldrich, St. Louis, MO) per gram of wet weight. The cells were then lysed on ice for 1 h, then sonicated at 10 W (10 s on/off cycle) for 3 minutes. The cell debris was separated by centrifugation at 10,000×*g* at 4 °C for 20 min. The supernatant was filtered with a 0.45 micron filter, then loaded onto a 5 mL HisTrap Fast Flow crude column that was mounted in an AKTA Purifier FPLC (GE Healthcare Life Sciences, Schenectady, NY). The column was washed with lysis buffer, followed by elution buffer (lysis buffer with 250 mM imidazole added) at 1:9 proportion with lysis buffer (34 mM imidazole). EAL was then eluted using 40% elution buffer (106 mM imidazole). At higher elution buffer percentages, the EAL in the eluate was deficient in EutC. For EutBH6 and EutCH6, elution was performed at 100% elution buffer.

The protein-containing fraction (10 mL) was dialyzed in 10 K membrane dialysis tubing against 1 L of 10 mM KP_i buffer (pH 7.5) for 6 h, followed by buffer replacement, and dialysis overnight at 4 °C. The protein was then concentrated using aquacide for 1 h, and spin-concentrated using a Pierce 7 mL concentrator with 10 K membrane (ThermoFisher Scientific, Waltham, MA). The resulting protein concentration was checked using the Bradford protein assay,[96] by using a 2 mg/mL bovine serum albumin standard (ThermoFisher Scientific, Waltham, MA). For purified EutB and EutC, the concentration was directly assessed using ultraviolet absorption spectroscopy, by using the sequence-predicted extinction coefficients[97] at 280 nm of 33,350 $M^{-1}cm^{-1}$ and 18,450 $M^{-1}cm^{-1}$,

respectively. Absorption spectroscopy measurements were performed on a Shimadzu UV-1601 absorption spectrophotometer.

2.2.5.3 Gel Filtration

Gel filtration was performed on an analytical Superdex 200 10/300 GL column from (GE Healthcare Life Sciences, Schenectady, NY) by using the AKTA Purifier FPLC system. Protein Enzyme in a volume of 100 μL was injected into the column, which was pre-equilibriated with 10 mM KP_i buffer (pH 7.5), and run at a flow rate of 0.5 mL/min. Absorbance was monitored at 280 nm.

2.2.6 SDS and Native PAGE

Polyacrylamide Gel Electrophoresis (PAGE) was performed by using 10% Mini Protean TGX gels from BioRad Laboratories (Hercules, CA) according to known protocols.[98, 99] Coomassie Brilliant Blue dye was used for staining.

2.2.7 Enzyme Activity Assay

Enzyme activity was determined by using the coupled alcohol dehydrogenase assay, essentially as described.[2] The assay mixture contained 10 mM KP_i buffer (pH 7.5), 2:1 5'-deoxyadenosylcobalamin/active sites, substrate, 120 μM NADH and 4.3 U of alcohol dehydrogenase at 25 °C. For EAL reconstitution, 2:1 5'-deoxyadenosylcobalamin/active sites was added to EutB in a 10 mM KP_i buffer (pH 7.5) with 5 mM reductant, while EutC was incubated separately. Incubations prior to assay were performed for 1h on ice. The two subunits were mixed and enzyme activity was assayed, as described above.

2.2.8 Electron Paramagnetic Resonance Spectroscopy

Samples for EPR spectroscopy were prepared under dim red-safe lighting, and protected from light during the experiment to prevent photo-degradation of 5'-deoxyadenosylcobalamin. All preparations were performed on ice. A proportion of 2:1 5'-deoxyadenosylcobalamin/active sites was added to 50 μM enzyme (active site concentration; corresponding EAL concentration, 8.3 μM) in 10 mM KP_i buffer (pH 7.5) with 5 mM TCEP. [*S*]-2-aminopropanol (30 mM) was added to this mixture to start the reaction. The sample was mixed, transferred to a 2 mm OD quartz epr tube and flash frozen in liquid nitrogen. Less than 15 seconds elapsed between mixing and freezing. EPR spectra were obtained by using a Bruker E500 ElexSys EPR spectrometer that was equipped with a Bruker ER4123 SHQE cavity. Temperature was controlled by using a Bruker ER4131VT liquid nitrogen/gas flow cryostat system. All spectra were collected at a microwave power of 2 mW and microwave frequency of 9.45 GHz, at 120 K.

2.3 Results and Discussion

2.3.1 Homology Modeling of *Salmonella typhimurium* EAL

EAL from *E. coli* and *S. typhimurium* display 98% sequence identity in the EutB subunit and 91% sequence identity in the EutC subunit (Figure 1).[41, 43] In EutB, there are 9 non-conserved positions (Figure 2). Two of these positions, EutB322 [I (*E. coli*)→L (*S. typhimurium*)] and EutB159 [A (*E. coli*)→C (*S. typhimurium*)] are on the interior β-strands of the $(\beta\alpha)_8$ barrel structure. These residues lie near to the N-terminal end of the barrel, and are not predicted to make substrate contacts. Two other non-conserved positions, EutB203 [S (*E. coli*)→T (*S. typhimurium*)] and EutB209 [I (*E. coli*)→V (*S.*

typhimurium)] are on the α-helices on the exterior of the barrel, and another position, EutB185 [V (*E. coli*)→A (*S. typhimurium*)] is in the loop region, that connects barrel α-helix 1 to β-strand 2 at the N-terminal end of the barrel. EutB87 [Q (*E. coli*)→R (*S. typhimurium*)] is involved in the binding interface between EutB subunits. There are a larger number, 26, of non-conserved residues in EutC. Of these, only EutC78 [V (*E. coli*)→A (*S. typhimurium*)] is involved in the EutC-EutB binding interface.

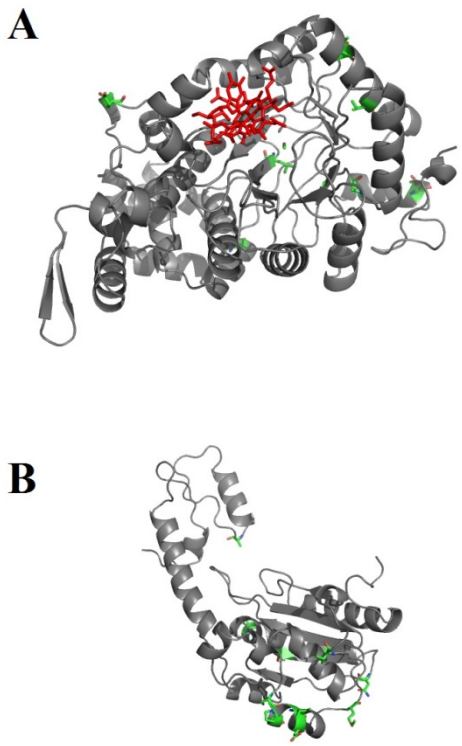


Figure 2. Comparison of secondary structure representations of EutB and EutC that show the non-conserved residues in the proteins from *S. typhimurium* and *E. coli*. The non-conserved residues are shown in green. (A) EutB. Adenosylcobalamin in the cobalamin binding site is rendered as red sticks. (B) EutC. PyMOL was used to manipulate and display the structures.

The high sequence identity allowed *S. typhimurium* EAL to be modeled directly from the template structure of the *E. coli* enzyme,[43] following sequence alignment.[85, 86] The selected template was the X-ray crystallographic structure of EAL, Protein Data Bank code, 3ABO. The 3ABO structure includes the complete oligomer structure [(EutB-EutC)$_2$]$_3$, and the EutB-EutC dimer units contain the bound substrate, ethanolamine, and the cobalamin cofactor analog, cyanocobalamin. Five homology models were generated using MODELLER.[100] The models were analyzed by using the Discrete Optimized Protein Energy (DOPE), which is an effective energy function that allows comparison between modeled structures.[88] The model with the lowest DOPE score was selected. The selected model was also validated by using the PROCHECK model stereochemistry test suite (Figure 3).[89] The model showed 99.9% of residues in allowed regions of the ϕ, ψ-, or Ramachandran, plot, of which 93.5% were in favored regions. The 2 residues that are in disallowed regions (EutC(L197), EutC(S156)) are present in loop segments outside of the α- and β-secondary structure, and therefore do not have a significant impact on the structure of the model. The model structure was also assessed by using the G-factor, a log-odds score based on observed distributions[101] of stereochemical parameters. Values of the G-factor were all significantly larger than the limits of -0.5 (probability of being incorrect) (Table 1).

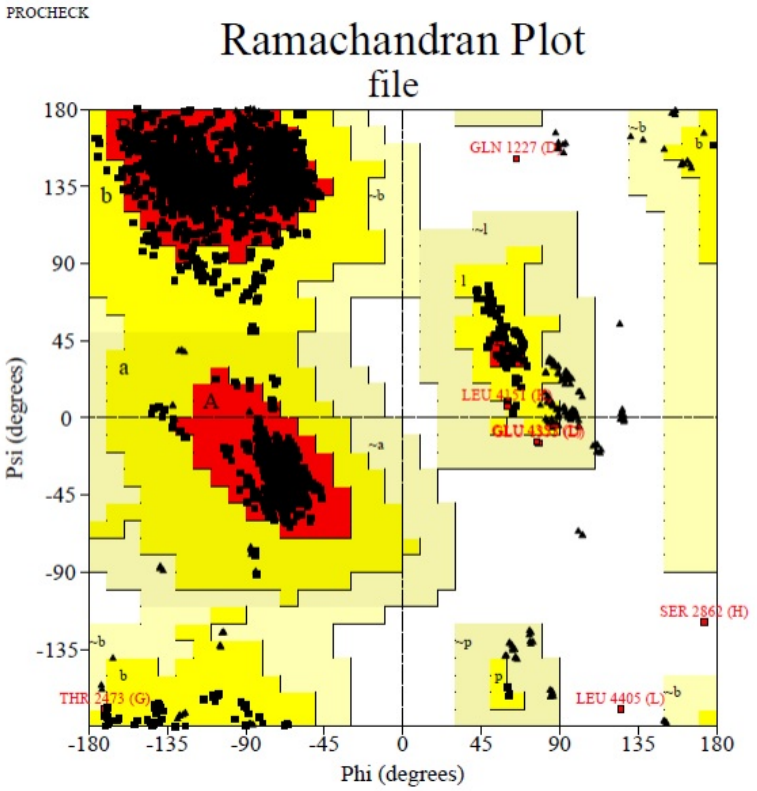


Figure 3. Ramachandran plot of homology modeled *S. typhimurium* EutB and EutC structures. Obtained by using PROCHECK stereochemistry test suite.[89] Most favored regions are shown in red. Additional allowed regions are shown in yellow. Generously allowed regions are shown in light brown. Disallowed regions are shown in white. Residues that reside in the disallowed regions are labeled.

The homology model of the *S. typhimurium* EAL oligomer structure is presented in Figure 4. An overlay of the homology model and E. coli secondary structures is shown in Figure 5. The positions of nonhomologous residues are highlighted in Figure 6, and the root-mean-square deviation of backbone Cα atoms is presented in Figure 7.

Table 1. The G-factors for the homology modeled *S. typhimurium* protein structure.

Parameter	Score
Dihedral Angles	
Φ-Ψ distribution	0.35
χ_1- χ_2 distribution	-0.35
χ_1 only	0.02
χ_3 & χ_4	0.48
Ω	-0.16
Main-Chain covalent forces	
Main-chain bond lengths	-0.05
Main-chain bond angles	-0.27

The G-factor is a log-odds score for stereochemical parameters based on observed distributions. A low G-factor is indicative of a low probability of finding that particular conformation. Values less than –0.5 are unusual, and values below –1.0 are highly unusual.

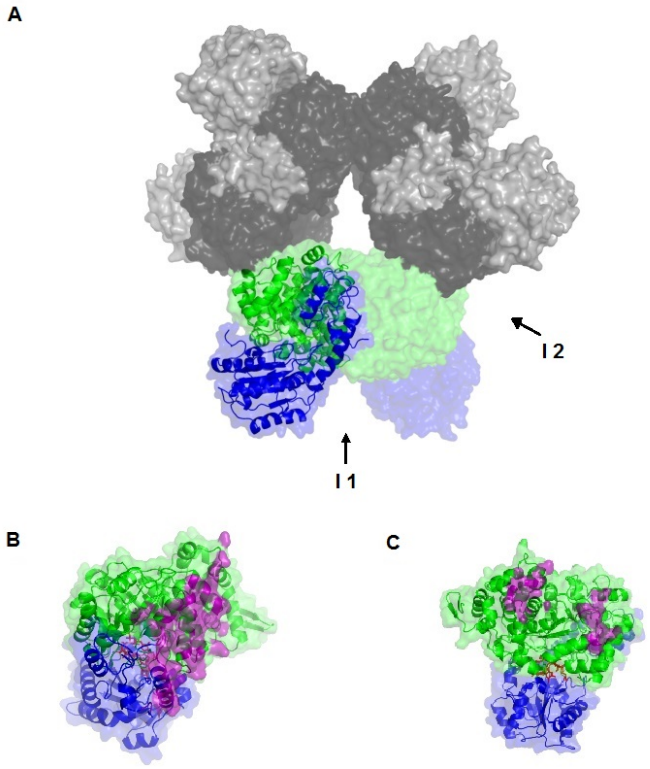


Figure 4. Surface and secondary structure representations of the homology modeled structure of *S. typhimurium* EAL. (A) The $[(EutB\text{-}EutC)_2]_3$ EAL oligomer. One $(EutB\text{-}EutC)_2$ heterodimer is shown as green (EutB subunits) and blue (EutC subunits), and other two heterodimers in grey tone (EutB, dark grey; EutC, light grey). (B and C) Two EutB-EutC heterodimers, that compose one $(EutB\text{-}EutC)_2$ homodimer. EutB is colored green and EutC is blue. (B) The intra-heterodimer interface (denoted I1, in panel A), with residues involved in the contact surface represented in purple. (C) The inter-heterodimer interface (denoted I2, in panel A), with residues involved in the contact surface represented in purple.

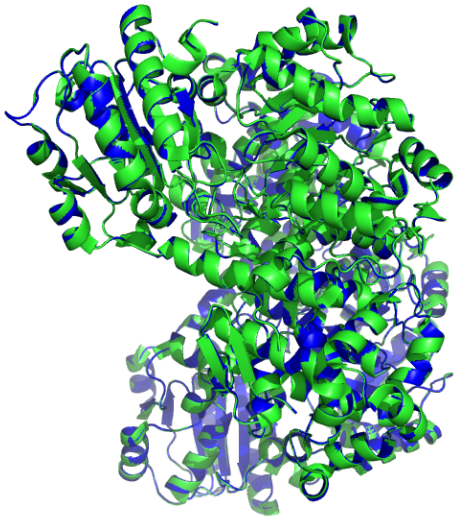


Figure 5. PyMol ribbon representation of *E. coli* 3ABO template structure[43] (green) overlaid on *S. typh.* EAL modeled structure (blue). Structures were aligned in PyMol using a structure based alignment algorithm. 70% of the atoms could be aligned, with a root mean square deviation of 0.197.

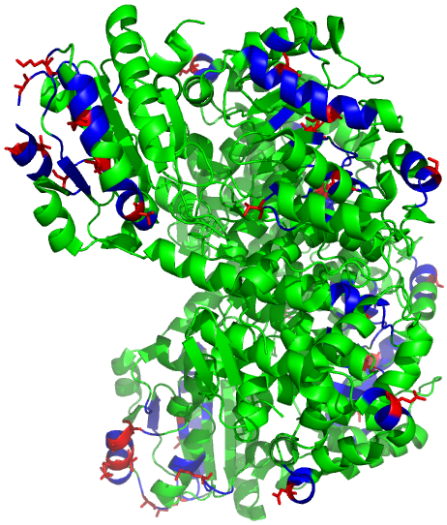


Figure 6. PyMol ribbon representation of 3ABO crystal structure[43] shown in green. Non-homologous sites between *E. coli* and *S. typh*. shown in red as sticks. Nearby residues, within a 5 Å sphere of influence from non-homologous sites, are shown in blue.

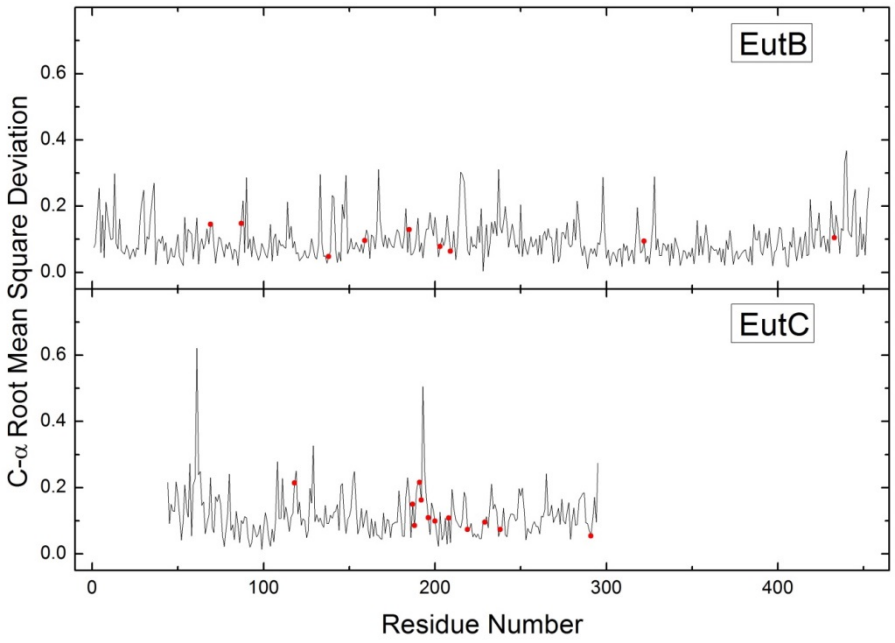


Figure 7. Root mean square deviation of C-α atoms between *E. coli*[43] and *S. typh.* EAL. Non-homologous residues indicated by red dots. Structural alignment of the 3ABO *E. coli* EAL structure and the modeled *S. typh.* EAL structure computed in PyMol using a structure based alignment algorithm.

2.3.2 Subunit Interaction Hierarchy from the Homology Model of *Salmonella typhimurium* EAL

Figure 4 illustrates the different hierarchies of organization of the EutB and EutC subunits in the oligomer: (a) EutB and EutC monomers form the EutB-EutC heterodimer. (b) Two heterodimers form a $(EutB\text{-}EutC)_2$ homodimer, and (c) three $(EutB\text{-}EutC)_2$ homodimers form the biologically active oligomer, which is described as a trimer of dimers, $[(EutB\text{-}EutC)_2]_3$. Thus, there are two different types of contact faces between EutB-EutC units. The $(EutB\text{-}EutC)_2$ intra-homodimer contact face has an area of 1293 $Å^2$, and is shown in Figure 4B. The edge-to-edge contact from inter-homodimer

formation of the trimer, shown in Figure 4C, has a smaller surface area of 804 Å^2. The primary contacts for both the intra- and inter-homodimer interfaces are on the EutB subunits. The model shows that the EutC subunit is not involved in the edge-to-edge, intra-homodimer interaction. Therefore, the model predicts that the hexameric $(EutB_2)_3$ trimer will form in solution, in the absence of EutC.

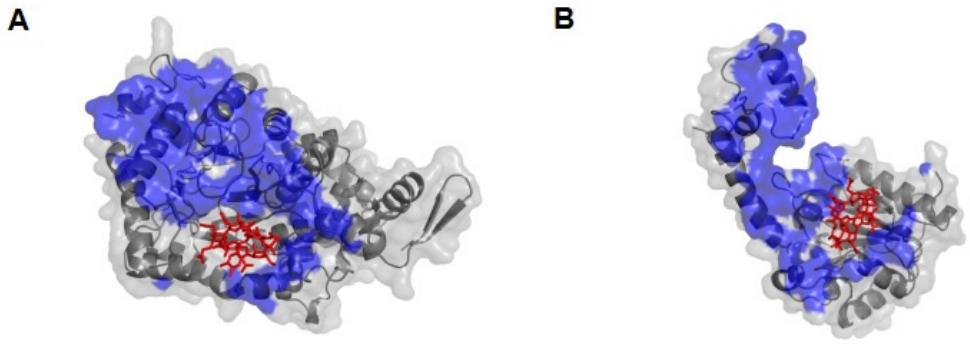


Figure 8. Contact surfaces of the EutB and EutC subunits in the EutB-EutC heterodimer, in the modeled structure of *S. typhimurium* EAL. (A) EutB subunit. (B) EutC subunit. The contact face is shown as a blue surface. The 5'-deoxyadenosylcobalamin cofactor is represented as red sticks.

Figure 8 shows the common contact face between EutB and EutC, which has a surface area of 2080 Å^2. The EutC contact face has a higher incidence of positively charged residues (23% positive, 10% negative), compared with the corresponding face on EutB, which carries the opposite charge (6% positive, 20% negative), where the percentages are relative to all contact residues. Figures 9 and 10 present the classes of contact surface residues and the location in EutB and EutC, respectively. In addition, two salt bridges are identified in this binding interface: EutB-D398 – EutC-R80 and EutB-D411 – EutC-R89. These results suggest that static charge-charge interactions contribute significantly

to EutB-EutC heterodimer assembly, and therefore, that low ionic strength conditions will favor inter-subunit interactions.

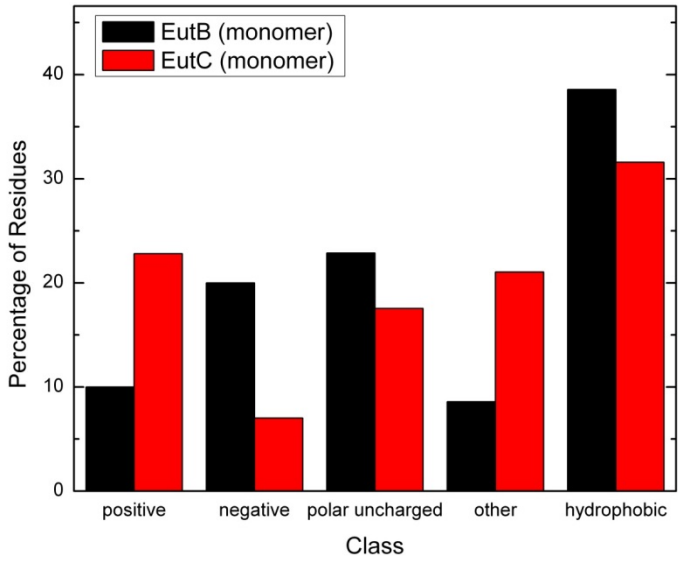


Figure 9. Properties of residues that form the binding contact face between EutB and EutC in the EutB-EutC heterodimer. The histrogram represents the percentage of residues that have a given character, that are located in the binding contact face between EutB and EutC subunits. A residue is deemed to be part of the contact face if the solvent accessibility changes with and without the adjacent subunit. The properties listed are identified by the following amino acid types (single-letter amino acid code). Positive – R, H, K. Negative – D, E. Polar uncharged – S, T, N, Q. Hydrophobic – A, I, L, M, F, W, Y, V. Other – C, G, P.

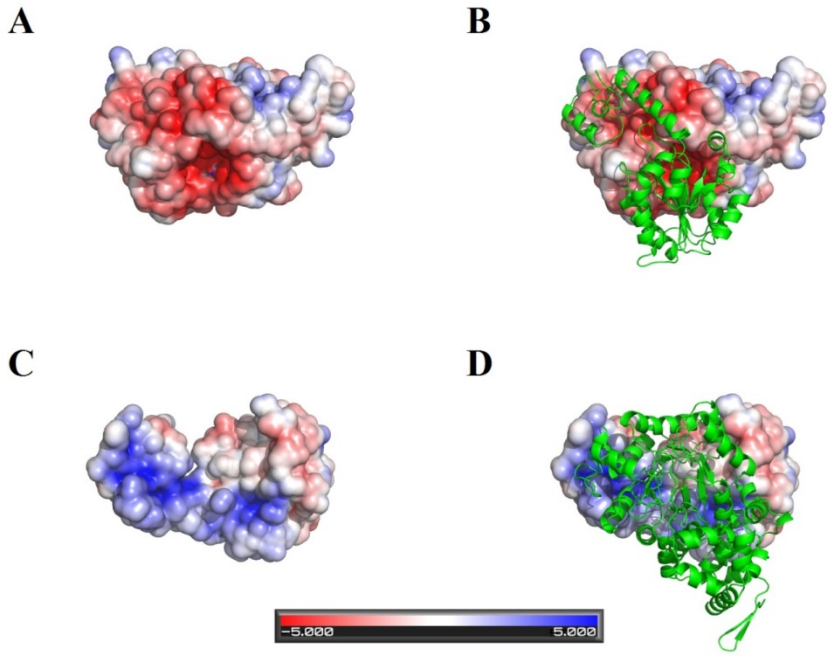


Figure 10. Charge distribution over the surface of the EutB and EutC proteins, and specifically, on the surface in the region of the binding contact face between EutB and EutC in the EutB-EutC heterodimer. (A) EutB subunit surface. (B) Same as Panel A, but with EutC subunit overlaid as green secondary structure rendering. (C) EutC subunit surface. (D) Same as Panel C, but with EutB subunit overlaid as green secondary structure rendering. The electrostatics of the modeled protein surface were calculated by using the Adaptive Poisson-Boltzmann Solver, APBS.[92] The PDB2PQR web server,[93, 94] was used to convert protein structural coordinates (PDB files) into files containing the calculated charge information (PQR). Scale bar shows electrostatic potential of surface using color, with units of k_BT/e. PyMOL was used to manipulate and display the structures.

The EAL from *S. typhimurium* contains 16 cysteine residues. Eight of these cysteine residues (EutB: C223, C361, C363, C364, C388; EutC: C158, C243, C261) are highly conserved, based on a survey of 693 EutB sequences and 689 EutC sequences from different bacterial species and strains.[95] Figure 11 illustrates the positions of the cysteine

residues in the EutB and EutC structures. The conserved residues include an unusual CDCC motif, in the active site region in EutB (residues C361, C363, C364), and C261 of EutC, which is in the cofactor binding site. Both of these active site regions become solvent exposed in the isolated subunits, and therefore, during the separate expression of EutB and EutC. Native EAL appears to be devoid of cystine disulfide bonds, with the possible exception of cysteines in the active site, and C37 and C46 of EutC, which were not modeled, because they are present on a cleaved leader sequence in *E. coli*. However, the EAL homology model indicates that, when the subunits are grown separately, opportunistic disulfide linkages may form, which indicates a requirement for a disulfide-reducing environment, during the reconstitution of isolated EutB and EutC subunits.

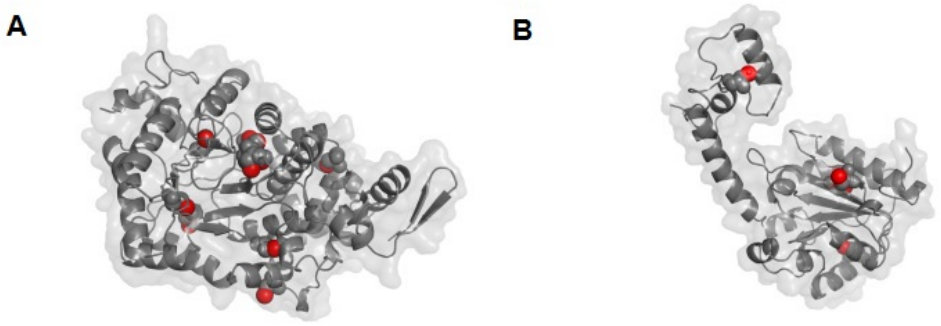


Figure 11. **Positions of cysteine side chain sulfhydryl groups in the EutB and EutC subunits in the modeled structure of *S. typhimurium* EAL.** (A) EutB subunit. (B) EutC subunit. Sulfhydryl groups shown as red spheres. The subunit orientations are the same as shown in Figure 8.

2.3.3 Expression and Purification of Wild-type EAL

The (histidine)$_6$–tagged, wild-type EAL construct, EALH6, allows expression of EAL at 37 °C without significant inclusion body formation. An SDS-PAGE gel of overexpressed EAL purified from this construct is shown in Figure 12. The EutB and EutC protein subunits have calculated masses of 48 kDa, and 30 kDa, which are comparable to the expected masses of 50 kDa and 32 kDa, respectively. The differences in masses between EutB and EutC from the expressed EAL oligomer and the individually purified subunits are explained by the presence of the pET28 (histidine)$_6$–tag sequence in the plasmids used for the purification of the individual protein subunits.

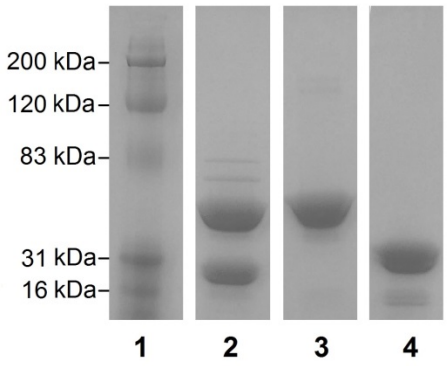


Figure 12. Sodium dodecyl sulfate polyacrylamide gel electrophoresis of purified proteins. Lanes contain the following: (1) Protein molecular mass ladder. (2) EAL purified from the EALH6 construct. (3) EutB. (4) EutC. Lanes 4 was adjusted to correct for dye front curvature.

EAL purified from the EALH6 construct had a higher proportion of the EutB protein, relative to EutC. This phenomenon has been previously noted in other EAL overexpression systems,[15] and attributed to more efficient expression of genes that are closer to the promoter region, relative to genes further downstream.[15] In the EALH6 pBR322-based construct, *eutB* precedes *eutC*, and is therefore closer to the *lac* promoter. Alternative explanations are that incompletely constructed EAL shows a higher affinity

for the nickel column than the complete oligomer, or that EutC is released from column-bound EutB during the washing procedures. In order to accomplish the purification of the homogeneous $[(EutB\text{-}EutC)_2]_3$ complex, elution from the nickel affinity column was performed at a relatively low imidazole concentration (100 mM). At higher imidazole concentrations (250 mM), the eluted EAL was deficient in EutC (Figure 13). Densitometric analysis[102] of SDS PAGE gels showed that the subunit molar ratio of the purified EAL was 1.00:1.01, following elution at 100 mM imidazole, compared to a molar ratio of 1.00:0.32, for elution with 250 mM imidazole (Table 2). In addition, native PAGE and gel filtration experiments show no evidence of excess EutB or EutC in purified EALH6 (Figures 14 and 15). In summary, the results indicate the following: (*a*) EutC-deficient EutB-EutC oligomer-like complexes bind with higher affinity to the nickel affinity column. (*b*) The purified $[(EutB\text{-}EutC)_2]_3$ complex can be obtained by elution with 100 mM imidazole.

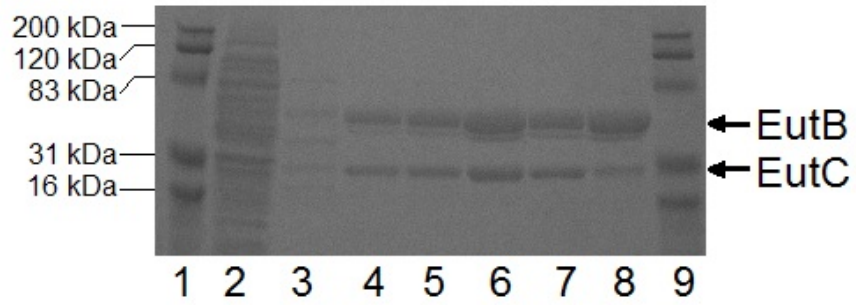


Figure 13. Sodium dodecyl sulfate polyacrylamide gel electrophoresis of EALH6 nickel affinity column purification fractions, for different imidazole elution concentrations. Lanes represent the following conditions: (1) and (9) Protein molecular mass marker ladder (Biorad). (2) Unbound protein flow-through (10 mM imidazole). (3) 25 mM imidazole. (4)–(7) 100 mM imidazole. (8) 250 mM imidazole.

Table 2. Relative intensity of EutB and EutC bands on an SDS-PAGE gel, obtained by densitometric analysis. *

Imidazole Concentration	100 mM	250 mM
Relative Intensity		
EutB	0.607	0.830
EutC	0.392	0.0.170
Molar Ratio		
EutC/EutB	1.01	0.321

* Protein bands were stained by using Coomassie Brilliant Blue dye. The columns represent the EAL eluted at different concentrations of imidazole from the nickel affinity column. The molar ratio of EutC/EutB (scaled for the difference in EutB and EutC mass) is also presented. Densitometry was performed on gel images by using the MATLAB program, GelBandFitter.[102]

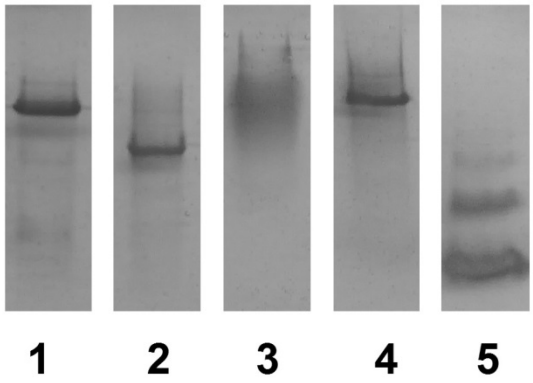


Figure 14. **Native polyacrylamide gel electrophoresis of purified proteins.** Lanes contain the following: (1) EAL purified from the EALH6 construct. (2) EutB. (3) EutC. (4) EAL reconstituted from the purified EutB and EutC subunits. (5) Bovine serum albumin standard. Lane 5 was adjusted to correct for dye front curvature.

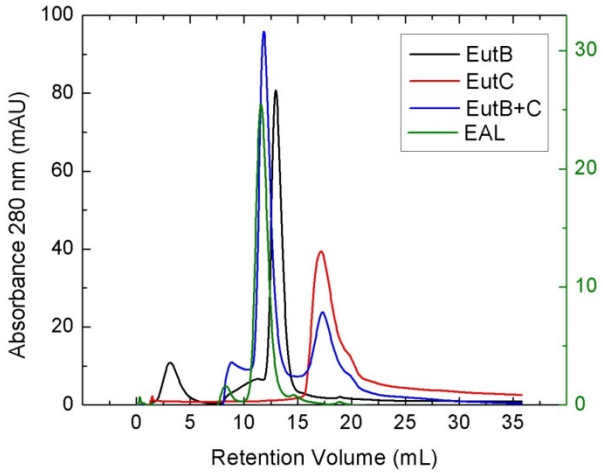


Figure 15. Gel filtration retention profiles for EAL and subunits. The profiles correspond to the following proteins: 100 µM EutB (black), 100 µM EutC (red), 100 µM reconstituted EAL (blue) and 30 µM EAL (green). The absorbance scale is on the left for EutB, EutC, and reconstituted EAL, and on the right for EAL. The peak retention volumes are as follows: EAL (11.6 mL), reconstituted EAL (11.8 mL, 17.2 mL), EutB (12.9 mL), EutC (17.2 mL).

2.3.4 Expression and Purification of Individual EutB and EutC Proteins

The formation of insoluble aggregates during the separate, individual expression of EutB and EutC proteins from the respective constructs, was prevented by expression at low temperature (30 °C) in NEB T7 Express I^q *E. coli* cells. The tight control of EAL expression by the upregulated *lac* repressor in this strain[103] prevents the formation of inclusion bodies, and leads to formation of soluble EutB and EutC. Using a single step of nickel affinity column chromatography, over 90% pure protein could be obtained as judged by densitometric analysis of SDS-PAGE gels. The EutB and EutC proteins were purified further, by using size exclusion chromatography. EutB was purified to homogeneity, and EutC was purified to >95%. An overloaded SDS-PAGE gel of the purified protein preparations is shown in Figure 12. The mass of the EutB protein was

calculated to be 50 kDa, which agrees well with the sequence-predicted mass of 52 kDa. The EutC protein had a calculated mass of 33 kDa, which is comparable to the sequence-predicted value of 34 kDa. Following purification, the EutC protein showed visible aggregation over the course of 2 h, when at high concentrations (> 5 mg/mL). Therefore, care was taken to ensure that EutC was not kept at high concentration for periods > 15 min. The EutB protein was not observed to form higher molecular mass aggregates, up to a protein concentration of 40 mg/mL, over the course of 2 h. The results show that the individual EutB and EutC proteins can be expressed and purified, separately, by the developed protocol.

2.3.5 Characterization of Oligomeric State of Individually Expressed and Purified EutB and EutC

Figure 14 shows non-denaturing PAGE results for protein purified from expression of the EALH6, EutB and EutC constructs. The EALH6 protein oligomer (494 kDa; lane 1) migrates as a single dominant band, with a relative mobility (R_f) of 0.187. The purified EutB protein (Figure 14, lane 2) displays a single dominant band with a larger R_f value (0.247). The presence of a single band for purified EutB is consistent with the presence of EutB as the 310 kDa $(EutB_2)_3$, or $EutB_6$ hexameric, oligomer. This interpretation is supported by gel filtration experiments (Figure 15), which show that purified EutB elutes closer to the authentic EALH6 oligomer (+1.0 ml), than to purified EutC (+5.5 mL). Although differences in protein size, structure and charge influence R_f values in non-denaturing PAGE, so that comparisons of different proteins are only approximate, we note that the bovine serum albumin control in Figure 14 (lane 5), shows characteristic

monomer (66 kDa), dimer (132 kDa) and trimer (198 kDa; R_f = 0.265) bands,[104] which display relatively large R_f values.

Figure 14 also shows that the purified EutC subunit migrates as a diffuse band in the non-denaturing gel. Under denaturing and reducing conditions, EutC migrates as a single band in SDS-PAGE, as shown in Figure 12, and under reducing conditions, it elutes in a single fraction in Gel Filtration (Figure 15). Therefore, the diffuse band on the non-denaturing gel is caused by non-specific aggregation of isolated EutC. The first 27 residues of EutC have been associated with the low solubility and aggregation of wild-type EAL in *E. coli*.[15] It was found that these residues are not essential for activity, and can be removed.[15] The first 19 residues of *E. coli* EutC are homologous to those in *S. typhimurium*, and have been identified as a signaling sequence that targets EAL to the eut metabolosome microcompartment, where EAL is found *in vivo*.[105] The residues in the N-terminal region of EutC may induce the formation of aggregates of varying sizes and conformations.

2.3.6 Assembly of EAL from Individually Expressed and Purified EutB and EutC

Figure 14 shows the non-denaturing PAGE results for reconstitution of 1:1 mol/mol purified EutB and EutC proteins. The reconstituted EAL (lane 4) has R_f =0.178, compared with R_f=0.187 for wild type EAL (lane 1). The closely similar R_f values indicate that the $[(\text{EutB-EutC})_2]_3$ EAL oligomer is reconstituted from its component subunits. The reconstituted EAL shows some evidence of diffuse staining at R_f values less than the principal band. This may represent improperly constructed EAL, owing to aggregation, because EutB is present in stoichiometric amounts with EutC. In summary,

the results demonstrate that the developed purification and reconstitution protocols lead to the native $[(EutB\text{-}EutC)_2]_3$, oligomeric structure of EAL, from isolated EutB and EutC subunits.

2.3.7 Steady-state Enzyme Activity of EAL Reconstituted from Isolated EutB and EutC

The homology model structures in Figure 11 show that cysteine residues are exposed on the contact surfaces of both the EutB and EutC subunits. Thus, EutB and EutC subunits, that are expressed and purified individually, will have solvent exposed cysteine residues in the active site region, and, in the absence of the reducing intra-cellular environment,[106] these cysteines may undergo oxidative reactions to form intra- and/or inter-subunit disulfide bond formation. In-line with this prediction, cystine disulfide bond reducing agents were found to be crucial for the functional reconstitution of EAL from isolated EutB and EutC subunits, as shown by the k_{cat} values for different reconstitution conditions in Figure 16. A k_{cat} value of 0.51 ± 0.09 s^{-1} is obtained, when no cystine disulfide reducing agent is added to the post-affinity column eluate, prior to loading onto the gel filtration column, and when no cystine disulfide reducing agent is present in the kinetic assay buffer. Figure 16 shows that addition of the disulfide reducing agent, *tris*-(2-carboxyethyl) phosphine (TCEP), to the sample, prior to loading on the gel filtration column, leads to 25-fold increase in the k_{cat} value to 12.6 ±0.7 s^{-1}. An additional increment to 15.6 ±0.3 s^{-1} is achieved when TCEP is added to the separate, isolated EutB and EutC, and incubated for 2-5 min, prior to mixing the subunits for the steady-state kinetic assay. Figure 16 shows that pre-incubation with TCEP prior to the steady-state kinetic assay mixture is more effective in raising k_{cat} than the disulfide reducing agents,

β-mercaptoethanol (β-ME) and D,L-dithiothreitol (DTT) at equivalent concentrations. β-ME is a monothiol reductant, and is therefore less effective than the dithiol, DTT.[107, 108] From the selectivity of TCEP for reducing cysteine residues in the protein,[109] we can conclude that the primary mechanism of inactivation of EAL is through oxidation of cysteines, rather than other pathways, such as proteolytic degradation.

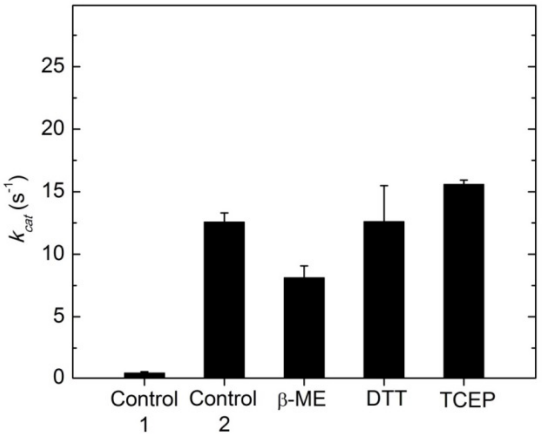


Figure 16. **Dependence of the activity of reconstituted EAL on the presence of different disulfide reducing agents.** Activity reprents the maximum turnover number (k_{cat}) attained at excess ethanolamine substrate (concentration>>K_M), under the assay conditions described in Methods. The EutB and EutC proteins were present at stoichiometric concentrations. The bars correspond to the following conditions: *Control 1*: Reductant was omitted during the purification procedure and during activity measurement. All other samples were purified in the presence of 5 mM TCEP reductant, which was removed as a consequence of the gel filtration step of the purification procedure. *Control 2*: No reductant added to the kinetic assay mixture. *β-ME*: β-mercaptoethanol (5 mM) added to kinetic assay mixture. *DTT*: Dithiothreitol (5 mM) added to assay mixture. *TCEP*: Tris-2-carboxyethyl phosphine (5 mM) added to assay mixture. Values represent the averages and standard deviation (error bars) of three measurements.

Although the enhancement of k_{cat} by DTT is comparable to TCEP, to within the standard deviation of the activity measurements, the stablity of DTT is reduced by nickel

contamination from the affinity purification step.[108] In addition, TCEP is compatible with sulfhydryl-specific spin labeling reagents at the TCEP concentration used,[110] which is a consideration for downstream EPR spectroscopy-associated experiments. Therefore, TCEP was selected as the cystine disulfide reductant for use in the purification and manipulation of EutB and EutC.

Table 3. Steady-state enzyme kinetic parameters for the purified EAL oligomer and EAL oligomer reconstituted from EutB and EutC subunits.

	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M ($M^{-1}s^{-1}$)
EAL	22.7 ± 0.3	31.6 ± 4.5	7.2 (± 1.1) x 10^5
Reconstituted EAL	16.0 ± 2.4	22.0 ± 4.2	7.3 (± 2.5) x 10^5

* Results represent the average and standard deviation for at least three measurements.

Table 3 shows the steady-state enzyme kinetic parameters for the EutB/EutC-reconstituted EAL and the histidine$_6$-tagged EALH6 oligomers. The k_{cat} value of reconstituted EAL is (70 ±16)% of the k_{cat} for EALH6, and the K_M value was reduced by by 1.4-fold. The difference in K_M between EALH6 and reconstituted EAL may be caused by a subtle effect on the protein structure, that arises from the different His$_6$ tags. In EALH6, there is one His$_6$ tag, which is located on the EutB subunit. The individual EutB and EutC subunits each have a His$_6$ tag, which is adjacent to a linker region. Thus, the reconstituted EAL includes two His$_6$ tags. The reduction in k_{cat} is consistent with the level of functional reconstitution from isolated subunits, as described in more detail, below. . The catalytic efficiency, or "specificity constant,"[111] which is given by k_{cat}/K_M,

does not differ significantly for the two EAL oligomers. Isolated, purified EutB and EutC showed no detectable activity (< 0.03 s^{-1}). In summary, the steady-state enzyme kinetics results demonstrate that the developed purification and reconstitution protocol leads to the formation of a functional EAL, staring from the isolated EutB and EutC protein subunits. This result is consistent with the formation of the $[(EutB\text{-}EutC)_2]_3$ EAL oligomer structure, which is evidenced by the gel electrophoresis results in Figure 14.

2.3.8 EPR Spectroscopy of EAL Reconstituted from Isolated EutB and EutC

The cob(II)alamin-substrate radical pair intermediate accumulates to a level of >90%, relative to active sites, in EAL during steady-state turnover on the substrate, (*S*)-2-aminopropanol.[39] The substrate radical accumulates, because the radical rearrangement is the rate-limiting step in the catalytic cycle.[39] The cob(II)alamin-substrate radical pair intermediate can be cryotrapped, and the characteristic radical pair line shape can be detected by using low temperature (120 K), continuous-wave (CW) EPR spectroscopy.[37, 39, 112] The line shape arises from the isotropic *J*-coupling and anisotropic dipolar coupling of the two unpaired electron spins, for a defined, rigid orientation of the substrate radical and cob(II)alamin over an 11 Å electron-electron separation,[37, 38] and is therefore exquisitely sensitive to the structure of the protein in the active site region.[37, 112] Figure 17 shows the characteristic native cob(II)alamin-substrate radical pair line shape, with features at $g_{\perp}$ =2.35 for cob(II)alamin and around g=2.00 for the substrate radical, as reported previously.[37] Figure 17 shows that the characteristic cob(II)alamin-substrate radical pair line shape is also present in the reconstituted EAL sample, and that the individual EutB and EutC proteins did not support the formation of paramagnetic species. The spectra of wild type EAL and reconstituted EAL overlap, upon rescaling the wild

type spectrum by a factor of 0.73. The results suggest the following: (*a*) The functional reconstitution of EAL activity from isolated EutB and EutC subunits is achieved at a level of 73%, relative to the EALH6 enzyme. This value is consistent with the reduction in k_{cat} of 70% (Table 3). The origin of the 70% level of functional reconstitution may be caused by incomplete binding of EutC to the oligomer. (*b*) The appearance of the radical pair indicates that the free energy surface for the reaction, and specifically, rate limitation by the radical rearrangement step, is the same in the reconstituted and control EAL oligomers. (*c*) The structure of EAL in the active site region of the reconstituted enzyme is comparable to the structure in native EAL.

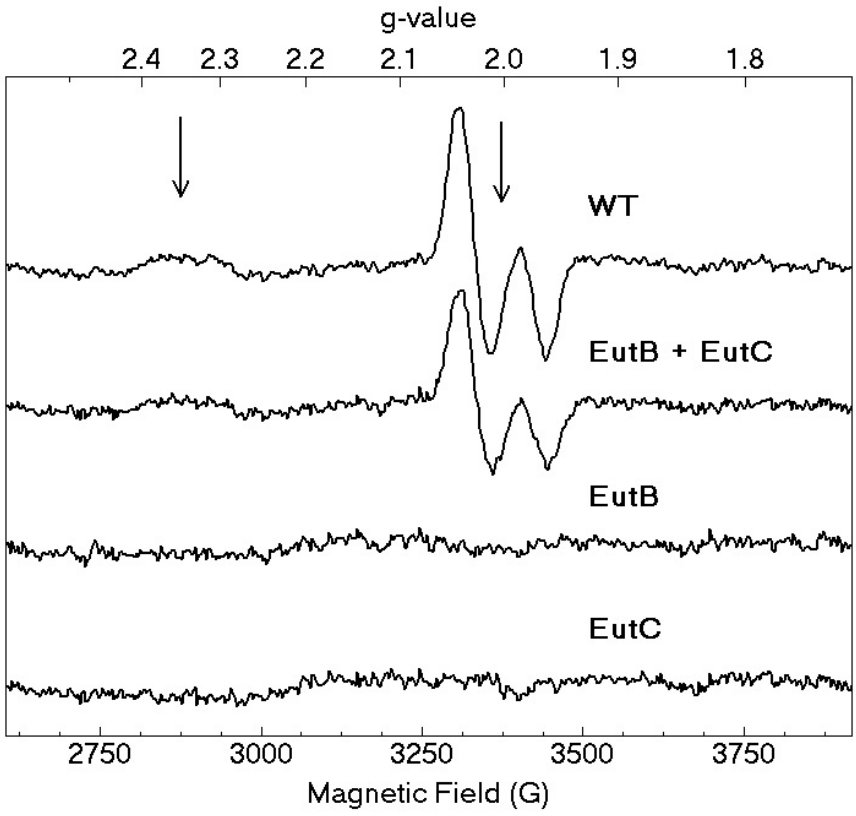


Figure 17. X-band CW-EPR spectra of the cryotrapped cob(II)alamin-substrate radical pair in EAL in different protein samples. From top, spectra shown are as follows: *WT*: EALH6 protein. *EutB+EutC*: EAL enzyme reconstituted from purified EutB and EutC

subunits. *EutB*: Purified EutB subunit, alone. *EutC*: Purified EutC subunit, alone. Arrows show $g = 2.35$ ($g_{\perp}$ for cob(II)alamin) and g=2.0023 (free electron g-value). *Experimental conditions*: microwave frequency, 9.458 GHz; temperature, 120 K; microwave power, 2.04 mW; magnetic field modulation, 10.0 G; modulation frequency, 100 kHz; scan rate, 71 G/s; time constant, 10.24 ms.

2.4 Conclusions

A homology model for the *S. typhimurium* EAL is constructed from the X-ray crystallographic structure of EAL from *E. coli*,[43] and demonstrated to be robust by statistical criteria. The model structure is used to describe the hierarchy of EutB and EutC subunit interactions that construct the native EAL oligomer, and specifically, to address the long-standing[14, 41] challenge of reconstitution of the functional oligomer from isolated, purified subunits. The model predicts that $(EutB_2)_3$ oligomer assembly will occur from isolated EutB, and that this hexameric structure will template the formation of the complete, native $[(EutB\text{-}EutC)_2]_3$ oligomer. SDS- and non-denaturing PAGE and gel filtration verify this prediction. The homology model also shows multiple cysteine residues in the subunit-subunit contact surfaces of EutB and EutC, which predicts that formation of non-native, cystine disulfide bond interferes with reconstitution. This is verified by the activation of steady-state enzyme activity, by addition of disulfide reducing agents during purification of the isolated EutB and EutC, and during reconstitution of the EAL oligomer. The significant static charge interactions at the interface between EutB (net negative) and EutC (net positive) in the heterodimer interface dictate low ionic strength conditions during reconstitution. The Ångstrom-scale congruence of the reconstituted and native EAL in the active site region, which is revealed by EPR spectroscopy, demonstrates correct assembly of the catalytic machinery.

Overall, the hierarchy of subunit interactions and microscopic features of the contact surfaces, that are revealed by the homology model, guide the establishment of a refined genetic and biochemical approach to reconstitution of the functional $[(EutB\text{-}EutC)_2]_3$ EAL oligomer, and provide a rationale for the structure. This is a prerequisite for further advances toward understanding of the molecular mechanism of EAL catalysis, by using site-directed mutagenesis, and for insight into approaches to therapeutic manipulation of eut-associated metabolosome and disease pathways.

Chapter Three

Cobalamin Binding to Ethanolamine-Ammonia Lyase

3

3.1 Introduction

3.1.1 Background

EAL requires 5'-deoxyadenosylcobalamin (AdoCbl) for catalysis. EAL uses radicals derived from the homolytic splitting of the cofactor's Co-C bond to catalyze the 1,2 rearrangement of the amine and hydrogen of ethanolamine and 2-aminopropanol, leading to the elimination of ammonia, and formation of the corresponding aldehyde. In this chapter, the dissociation constant (K_D) of the AdoCbl cofactor with EAL is directly assessed by using tryptophan fluorescence quenching. Analysis of this binding interaction allows the estimation of cofactor release kinetics, and the binding energy available for catalysis. As such, this work provides further insights into EAL catalysis.

The AdoCbl binding site in EAL is at the C-terminal end of the $(\beta\alpha)_8$ barrel of the EutB subunit.[42, 43] Within the EutB subunit, there is an adenine binding pocket, which accommodates the upper axial ligand (5'-deoxyadenosine) of AdoCbl. This site is capped by EutC, with several residues making critical cofactor contacts.[73] In certain B_{12}-dependent isomerases, such as glutamate mutase,[113] and methylmalonyl-CoA mutase (MCM),[114] there is a distinct binding mode for the AdoCbl cofactor, in which the lower dimethylbenzimidazole (DMB) axial ligand is displaced. These are referred to as base-off enzymes. Base-off enzymes contain a conserved DX**H**XXG cobalamin binding motif where the central histidine residue replaces DMB as the Co axial ligand.[115] DMB remains covalently linked to a pyrrole ring of the cofactor, and the displaced nucleotide ligand is

sequestered in a groove at the active site.[116] However, DMB may not contribute significantly to the cofactor-protein binding in base-off enzymes, as in MCM where the displaced tail only contributed ~1 kcal/mol to binding.[117] The structural changes in the cofactor are accompanied by significant structural rearrangement of the protein to accommodate the changed binding mode.[117] Other enzymes, such as diol dehydratase,[118] and ribonucleotide reductase[119] do not contain the cobalamin binding motif, and bind AdoCbl with the DMB ligand in place as the cobalt's lower axial ligand. This base-on binding motif is the cofactor binding mode employed by EAL. UV-visible, EPR and X-ray crystallographic studies show the DMB ligand bound to the cobalt, and remaining in place throughout the EAL catalytic cycle.[43, 120] Further, in EAL, even the binding of substrate which initiates the reaction, does not significantly destabilize the AdoCbl equilibrium structure.[121] Since both the upper and lower axial ligands of B_{12} are not significantly perturbed in cofactor binding to EAL, it is likely that the protein structure is altered as a result of the binding interactions and the binding energy is instead stored within the EAL protein.[121]

The energy of the binding interaction between the cofactor and EAL has not been directly measured. Binding affinity, ΔG, is determined by the dissociation constant, K_D, as shown in Equation 1, where R is the gas constant, T is the temperature in Kelvin, and c is a standard concentration of 1 M used to render the value in the logarithm dimensionless.

$$\Delta G = -RTln\frac{K_D}{c} \qquad (1)$$

The dissociation constant of the cofactor binding to EAL has been estimated from the cofactor Michaelis constant, K_M, to be on the order of 1.5 μM.[122] This measurement

followed the activity of EAL as a function of AdoCbl concentration. Equating K_M and K_D assumes that the kinetics of the AdoCbl binding interaction is faithfully reported by the activity measurement. Equation 2 shows the kinetic scheme involved in the experimental determination of K_M for AdoCbl.

$$E\text{-}S + B_{12} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} E\text{-}B_{12}\text{-}S \xrightarrow{k_2} E\text{-}B_{12} + P \qquad (2)$$

In Equation 2, the enzyme-substrate complex, *E-S*, is added to AdoCbl, denoted B_{12}, which forms the ternary enzyme-AdoCbl-substrate complex, $E\text{-}B_{12}\text{-}S$. This reacts to release the product, *P*. The microscopic rate constants for AdoCbl binding and release are given by k_1 and k_{-1} respectively, with k_1 representing the bimolecular reaction and dependent on B_{12} concentration and having units of $M^{-1}s^{-1}$. The rate of enzymatic turnover is given by k_2. In this system, K_M and K_D are given by Equation 3 and 4 respectively.

$$K_M = \frac{k_2 + k_{-1}}{k_1} \qquad (3)$$

$$K_D = \frac{k_{-1}}{k_1} \qquad (4)$$

K_D is a good approximation for K_M when k_2 is sufficiently slow compared to k_{-1}. Otherwise, the subsequent catalytic steps influence the value obtained for K_M, and the estimated K_D will be unreliable. As such, a direct assay for cofactor binding is preferred. The K_D for the cofactor in the related B_{12}-dependent enzyme, MCM, has been directly measured by fluorescence quenching to be 0.17 μM.[117] MCM binds AdoCbl in the base-

off conformation, whereas EAL does not significantly change AdoCbl structure upon binding. Despite this difference, and the likely change in interaction energy, the fluorescence quenching technique can be used to measure the affinity of B_{12} for EAL.

3.1.2 Origins of Tryptophan Fluorescence

In MCM, AdoCbl cofactor binding was measured by following the quenching of the intrinsic protein fluorescence.[117] The aromatic side chains present in the amino acids tryptophan (Trp), phenylalanine (Phe) and tyrosine (Tyr) are fluorescent. The fluorescent properties of these are given in Table 1.[123] The double ring structure of Trp allows it to have a lower energy excited state, and slightly red-shifted absorbances and emissions compared to the other amino acids. The high quantum yield of Trp allows it to be the dominant source of fluorescence in unlabeled proteins. Though Tyr has a similar quantum yield, in proteins Tyr fluorescence tends to be quenched, or the excitation energy is transferred to lower energy Trp residues where the fluorescence is emitted. Phe has a much lower quantum yield, and its fluorescence is rarely observed, except in the absence of Tyr and Trp.

Table 1. Fluorescence Parameters of Aromatic Amino Acids in Water at Neutral pH

Amino Acid	λ_{ex} (nm)	λ_{em} (nm)	Quantum Yield	Lifetime (ns)
Phenylalanine	260	282	0.02	6.8
Tyrosine	275	304	0.14	3.6
Tryptophan	295	353	0.13	3.1

Data from Lacowicz.[123]

Trp fluorescence arises from the indole group, which has two low energy excited states, ^1L_a and ^1L_b, which may contribute to the absorption and emission bands.[124] These states can be treated independently.[125] The ^1L_a state is the one that is sensitive to the environment owing to its large transition dipole, and is the primary state observed in Trp fluorescence measurements.[126] The mechanism of this environmental sensitivity is as follows. On the absorption of an excitation photon, the electronic distribution of Trp instantly reacts (~1 fs); however, the heavier nuclei cannot respond on this timescale (Franck-Condon principle). This relative redistribution of charge causes a change in the permanent dipole moment of Trp. The orientations of the local fields present at each excited Trp causes inhomogeneous broadening of the Trp excitation spectrum. As the excited Trp relaxes, both intrinsic and extrinsic factors come into play; internally there are vibrations from the now unbalanced forces within Trp, and externally neighboring molecules reorient to respond to Trp's new dipole moment (dielectric relaxation). The external reaction field from the dielectric surroundings is the cause of the sensitivity of Trp fluorescence to the environment. The timescale and magnitude of rearrangements are critical to the resulting Trp fluorescence spectra. If the ground state Trp contained a dipole in the same direction as the excited dipole, the excitation is pre-stabilized, and only an increase in the magnitude of the stabilizing environmental dipoles is necessary, without significant rearrangements yielding red-shifted fluorescence spectra. In a fluid solvent with large dipoles (like water), the reaction field can respond within femtoseconds. In a protein, this field may take nanoseconds to fully form, which is longer than the excited state lifetime. This yields the red-shifted fluorescence emissions

(stabilized, lower energy) shown by Trp in water when compared to those of Trp buried within a protein.

The Trp emission's environmental sensitivity, stemming from the response time for the reaction field, yields fluorescence lifetimes from ~20 fs to nanoseconds, and conceivably every value in between.[127] Hydroxyl groups and the orientation of their large dipole relative to the Trp, in addition to solvent accessibility, can yield a very sensitive label-free method of protein conformation analysis. This sensitivity has been exploited to allow the use of endogenous Trp residues as intrinsic reporters of protein conformations, including a study in which wavelength shifts in the spectral maximum (Λ_m) for Trp residues buried in a non-polar environment (Λ_m = 330 – 332 nm) could be differentiated from solvent exposed Trp residues (Λ_m = 350 – 353 nm), and partially solvent exposed Trp residues (Λ_m = 340 – 342 nm).[128]

3.1.3 Tryptophan Fluorescence Quenching

In addition to changes in the spectral profile of Trp fluorescence caused by the properties of the local dielectric interactions, energy transfer interactions change the fluorescence lifetime of Trp. The lifetime of Trp (τ_F) is given by the exponential decay of fluorescence intensity upon pulse excitation. This lifetime is affected by the rate of non-radiative de-excitation of the fluorophore (k_{nr}) and the radiative decay rate (Γ) as shown in Equation 5.

$$\tau_F = \frac{1}{\Gamma + k_{nr}} \qquad (5)$$

Due to the difficulty in measuring the typical nanosecond fluorescence lifetimes, measuring steady-state intensity (I_{SS}) is a suitable experimental proxy.

$$I_{SS} = \int_0^\infty I_0 e^{-t/\tau_F} dt \qquad (6)$$

In Equation 6, I_0 is a constant that depends on the fluorophore concentration and instrumental parameters. Thus, changes in fluorescent lifetimes are manifested as changes in the steady-state fluorescence intensity. One drawback of following intensity is that there may be the trivial effect, wherein fluorescence intensity is reduced with the lifetime being unaffected.[123] A high concentration of molecules that absorb in the region of the fluorescence emission (or excitation) can artificially reduce I_{SS}. This "inner filter effect" can be corrected for,[129] but it places an effective limit to the concentrations that can be probed by fluorescence quenching.

The resonant energy transfer interaction is a more informative effect that changes τ_F, and, by extension, I_{SS}. Resonant excitation energy transfer to the absorbing molecule (quencher) occurs with a rate that is proportional to the square of the overlap of the transition dipoles, and inversely proportional to the sixth power of their separation.[126] The energy transfer efficiency (E_{ET}) is given by Equation 7.

$$E_{ET} = \frac{1}{1+(r/R_0)^6} \qquad (7)$$

In Equation 7, r is the separation between the fluorophore and quencher, and R_0 is a system dependent constant that accounts for the relative orientation of the dipoles, quantum yield, spectral overlap and the intervening medium. If the absorbing molecule is itself fluorescent, this is termed Förster Resonance Energy Transfer (FRET). However, if the absorbing molecule is non-fluorescent, this energy transfer still occurs, and the

acceptor (quencher) undergoes another method of de-excitation. In either case, resonance energy transfer serves as a mechanism for the reduction in the donor moleculue's τ_F.

Another mechanism of energy transfer is a collision quenching interaction, where energy transfer occurs after contact, during the excited state lifetime of Trp. Collisional quenchers include molecular oxygen, and electron scavengers such as protons, histidine and cysteine.[123] This quenching interaction may occur by a variety of mechanisms, including abstraction of the excited electron.[130] Despite mechanistic variations, collisional interactions are concentration dependent, and are described by the Stern-Volmer equation.

$$\frac{F_0}{F} = 1 + k_q \tau_0 [Q] \qquad (8)$$

In Equation 8, F_0 and F are fluorescent intensities in the absence and presence of quencher, $[Q]$ is the concentration of quencher, and τ_0 is the fluorescence lifetime of the Trp without quencher. k_q is a bimolecular quenching constant, which accounts for the efficiency of quenching and includes the dependence on steric shielding and viscosity.

Non-radiative quenching of Trp may also occur by intersystem crossing to the ^3L_a phosphorescent state,[126] solvent quenching, and excited state proton transfer.[131] This yields τ values that can vary by more than two orders of magnitude,[127] making Trp uniquely sensitive to local environmental conditions. Despite the wide variation in timescales and mechanisms, both static and dynamic quenching yield variations of fluorescence intensity that are well-described by a Stern-Volmer-like equation. For dynamic, collisional quenching, the quenching concentration prefactor $k_q\tau_0$ is termed K_{SV}. For static, (post-binding) energy transfer interactions, the prefactor is termed K_S, which

takes into account the association constant of complex formation.[123] Thus, even in the absence of detailed mechanistic information, changes in steady-state fluorescence intensity can be used to follow the structural changes, such as the binding of the AdoCbl cofactor.[117]

3.2 Methods

Binding of AdoCbl to EAL was measured by using the quenching of endogenous Trp fluorescence. A limiting concentration of EAL (100 nM active sites) was added to 100 nM to 1 mM cobalamin, and the samples' fluorescence assayed by using a Fluoromax-4 spectrofluorometer (Horiba Jobin Yvon). The samples were excited at 280 nm, and the full fluorescence spectrum from 330 nm to 450 nm was collected. Individual samples were made to minimize sample degradation, and sample preparation was performed on ice, under dim, red safe lights. Protein samples without the cobalamin were used to measure baseline fluorescence. AdoCbl showed no detectable fluorescence under these conditions. To measure and compensate for inner filtering, control experiments with the free amino acid, $_{L}$-tryptophan in place of protein, were conducted.

To increase sensitivity, an EAL mutant, αF258W was created. Starting from the EALH6 pBR322 plasmid previously described, mutational PCR was performed by using the Stratagene QuikChange II XL Site-Directed Mutagenesis kit by using the primers listed in Table 2. These primers replaced the phenylalanine codon at amino acid position 258 on EutB with a tryptophan codon. This mutant plasmid was transformed into XL10 Gold competent cells, for gene expression. Sequence analysis of the mutant plasmid confirmed 100% sequence identity in the EAL reading frame. Bacteria harboring this EAL αF258W

mutant plasmid were grown and mutant protein was purified according to the protocol described for the purification of wild-type EAL. Kinetic analysis was performed by using the ADH-NADH coupled assay.[2]

Table 2. Primers used in the Development of αF258W*

Primer Name	5' to 3' Primer Sequence
αF258Wfp	GGG CTT AAA AGA GTG GGG CGT CGA GCT GGC C
αF258Wbp	GGC CAG CTC GAC GCC CCA CTC TTT TAA GCC C

* Primers obtained from Integrated DNA Technologies

Fluorescence quenching experiments were repeated on this mutant as described above. The concentration of bound complex was calculated by fitting the normalized fluorescence quenching with a binding isotherm for independent binding sites. From this, the dissociation constant of AdoCbl was extracted.

3.3 Results and Discussion

3.3.1 Structural Considerations for Tryptophan Quenching

EAL has 4 Trp residues, αW91, αW430, βW58 and βW117. The positions of these residues are shown in Figure 1. These Trp residues are distant from the active site, with the closest one being 30.1 Å away from the cobalamin. (All distances are measured from the Trp C-α to the central cobalt atom.) Table 3 shows the cobalt-C-α distances for the Trp residues in EAL. For comparison, MCM has 11 native Trp residues in the cobalamin binding α-subunit, and 9 in the β-subunit. Analysis of the crystal structure of MCM[116] shows that the nearest Trp (αW334) is 8.4 Å removed from the cobalamin. As stated previously, the efficiency of Trp quenching energy transfer follows an r^{-6} dependence.

This strong distance dependence causes Trp residues further away from the cobalamin to be less efficiently quenched.

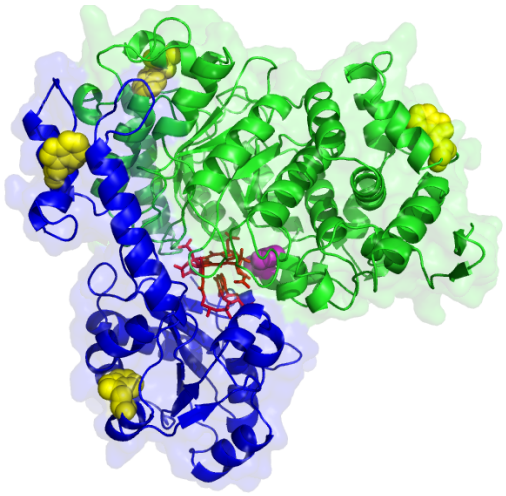


Figure 1. Location of Tryptophan Residues in EAL. PyMOL cartoon representation of EAL, with EutB subunit shown in green, EutC subunit shown in blue, and cobalamin cofactor shown as red sticks. Location of native tryptophan residues shown as yellow spheres. Location of phenylalanine residue, αF258 shown as purple spheres.

Table 3. Tryptophan to cobalamin Distances in EAL.

Residue	Distance (Å)*
αW91	40.0
αW258†	9.4
αW430	33.4
βW58	31.8
βW117	30.1

* Distances measured from C-α to cobalt
† Residue is Phe in native EAL

In the interface between EutB and EutC, there is a Phe residue (αF258) that is 9.4 Å from the cobalamin. The location of this residue makes it well-suited to report on cobalamin binding in EAL. In addition, though αF258 is near the active site, it is also positioned near a cavity in the interface of EutB and EutC. Thus, it may be possible to mutate this

residue to the larger Trp without significant deleterious effects on protein structure and activity.

3.3.2 Tryptophan Fluorescence in EAL

Figure 2 shows the excitation spectrum of EAL, EutB and αF258W while monitoring the 340 nm fluorescence. In all three spectra, there is a broad excitation peak centered at 278 nm. With 280 nm excitation, EAL and EutB show a broad emission peak centered at 335 nm. This is characteristic of Trp fluorescence,[126] as expected. The αF258W fluorescence maximum is red-shifted to 337 nm. This suggests that the environment of αW258 is more polar than the Trp residues in EutB and EAL.[132] This is consistent with its location near a cavity which may contain polar solvent molecules. Alternatively, the shift in fluorescence in this mutant may arise from a Trp other than αW258, which is caused by a structural change in the environment of this Trp.

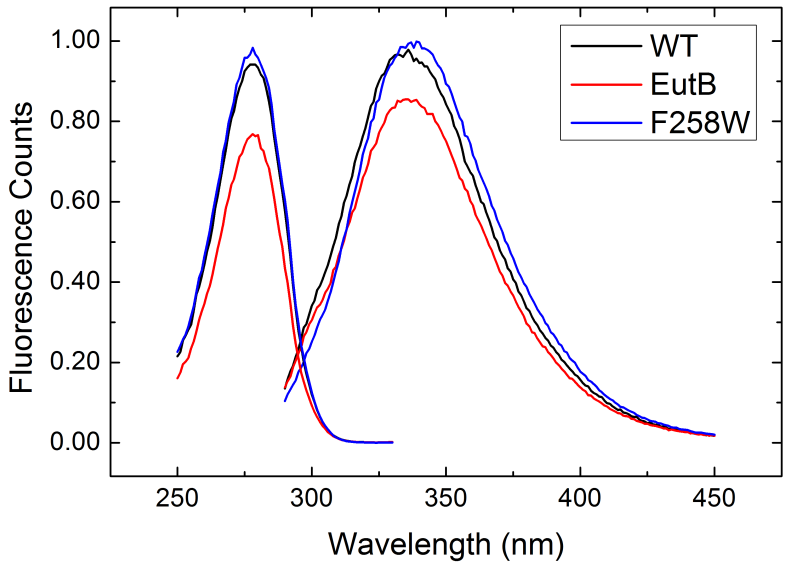


Figure 2. Excitation and Emission Spectra of Protein Tryptophan Fluorescence. Excitation spectrum (left) for 340 nm fluorescence.

Emission spectrum (right) for 280 nm excitation. Protein concentration is 1 μM in 10 mM KP_i at pH 7.5. Wild-type EAL shown in black, EutB shown in red, and αF258W shown in blue. Fluorescence counts measured have been scaled to the maximum excitation and emission of αF258W.

Another feature that is apparent from the data is the reduction in intensity for EutB, compared with wild-type, and a subtle increase in intensity for αF258W. Self-quenching of the fluorescence occurred at higher protein concentrations (10 μM), and the experiment was conducted in the linear region of the detector, so these effects cannot account for the change in intensity. The variation can be reconciled by considering the number of Trp residues in each protein. EutB contains two Trp residues, while wild-type EAL also includes two Trp residues from the EutC subunit for a total of four. αF258W also includes the eponymous fifth Trp. The fluorescence intensity depends on the number of Trp residues in the protein. However, the change in fluorescence intensity is non-linear. Since the Trp residues are in different environments, they possess differing fluorescence lifetimes and quantum yields, leading to the non-linear fluorescence intensities seen in the different proteins. This does not affect fluorescence quenching experiments, which are internally calibrated with the maximum fluorescence of each protein in the absence of quencher.

In fluorescence experiments, improved signal-to-noise ratios are possible by increasing the excitation intensity, resulting in greater fluorescence emission. However, there is a competing effect in which charge transfer interactions damage the fluorescing Trp residues (photobleaching). In addition, AdoCbl can be easily photolyzed. To mitigate these effects, the experiments were conducted under dim, red safe lighting, and experimental conditions were explored to select a condition that minimized damage while

maximizing signal. The final experimental conditions were selected to be 3 nm slit widths on the excitation and emission monochromators, with 0.1 s integration time at each wavelength. Under the selected experimental conditions, the photolysis of AdoCbl was monitored by UV-visible spectroscopy, and less than 0.1% (sensitivity limit) of the coenzyme was affected. In addition, Trp fluorescence was unaffected for the duration of the experiment. Thus, sample degradation is minimal, and the system is appropriate for fluorescence quenching studies.

3.3.3 Inner Filter Effect

The inner filter effect is caused by the UV absorbance of both the excitation and emission photons by the quencher. Absorbance of the excitation beam reduces the number of photons available to excite Trp, and absorbance of the fluorescence emission reduces the photons available for collection by the detector. The absorbance of the quencher is well-described by the Beer-Lambert law (Equation 9).

$$A = -\log_{10}\left(\frac{I}{I_0}\right) = \varepsilon c l \qquad (9)$$

A is the absorbance of the sample, I is the intensity of the light transmitted through the sample from the initial I_0 value, ε is the extinction coefficient, c is the concentration of the absorbing species and l is the path length. As a result of this logarithmic dependence, the intensity is acutely sensitive to ε, c and l. ε is a characteristic of the molecule's electronic structure. For AdoCbl, ε_{280} is 22000 $M^{-1}cm^{-1}$ and ε_{340} is 13000 $M^{-1}cm^{-1}$. These absorbances are comparatively high, with solutions of > ~10 μM obscuring 99% of the input light over 1 cm. Nonetheless, the electronic transitions responsible for these absorbances are essential for good spectral overlap for efficient energy transfer and

fluorescence quenching. The path length is dependent on the cuvette used, and can be modified. To minimize the filtering effect, a short path length cuvette (1 cm x 0.2 cm) was used with the long axis in the direction of excitation. Use of this low-volume cuvette reduces the signal-to-noise of the experiment because of the reduced material used. However, the sensitivity was sufficient for the experiments conducted.

To obtain an estimate for the extent of inner filtering, the free Trp amino acid was used in place of the enzyme in the fluorescence quenching assay. In this system, there is no binding to protein, so all of the artificial "quenching" that occurs is caused by the absorbance of the cobalamin. Figure 3 shows the result of the experiment and a Hill fitting of the data.

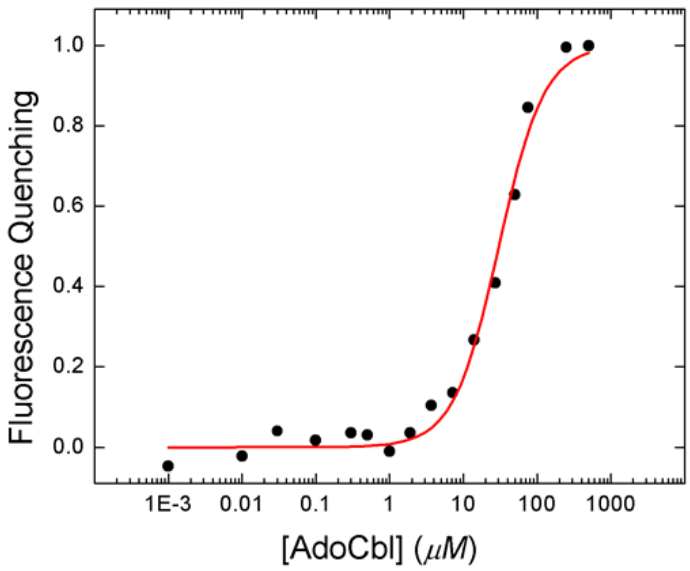


Figure 3. Inner Filter Effect of free Tryptophan by AdoCbl. Inner filter effect of free amino acid L-tryptophan in solution with AdoCbl. Data fit to Hill ligand binding model shown in red to give apparent binding constant.

The Hill equation (Equation 10) is a model for ligand binding.

$$\theta = \frac{[L]^n}{K_D^n + [L]^n} \qquad (10)$$

The fractional binding, θ, is given by the ligand concentration, [L], and the Hill constant, n. It was first used as an empirical model to fit hemoglobin binding data.[133] In Hill fits, n represents the degree of cooperativity involved in binding for cases where there are multiple ligand binding sites. A value greater than 1 is indicative of cooperativity.[134] The fitted, free Trp data shows "binding" with a Hill constant of 1.4, and an apparent K_D of 30.4 μM (Table 4). This means that at 10 μM there is 17% reduction in fluorescence intensity due to the inner filter effect. Special care must be taken to analyze data with a binding constant near this value, as a reduction in fluorescence intensity due to the inner filter effect may be erroneously attributed to binding.

Table 4. Best-fit Parameters for free Tryptophan Inner Filter Effect.*

n	***K*D (μM)**
1.41 ± 0.12	30.4 ± 9.1

3.3.4 Binding Measurements on EutB

The binding of various cobalamins to EutB was probed with fluorescence quenching. In EutB, fitting the quenching curves gave K_D values of 29.8 ± 5.8 μM for AdoCbl, cyanocobalamin and aquocobalamin. The comparable values of K_D are unexpected, because of the different structures and solution interaction properties of the different cobalamins, and this suggests that these data do not represent fluorescence quenching, but rather, the inner filter effect caused by the cobalamin. Slight variations in the binding

constants obtained are caused by the differing extinction coefficients in each cofactor analog. One can conclude that either the cofactor binds to EutB with a K_D of ≥30 μM, or, the Trp residues in EutB are too distant to be sensitive to cofactor binding.

3.3.5 Fluorescence Quenching of αF258W

The αF258W mutant introduces a Trp 9.4 Å from the binding site of the cobalamin. It replaces the aromatic residue, Phe, and is considered to be a conservative mutation. As previously stated, the larger size of Trp side chain is not expected to be disruptive to EAL structure because the phenylalanine replaced is on the surface of EutB, and can accommodate the increased size. The kinetic parameters of αF258W are not significantly changed from the wild-type values. αF258W has a k_{cat} of 10.2 s^{-1}, and a K_M of 28.8 μM for ethanolamine at 295 K. This is within a factor of two of the wild-type values (k_{cat} = 22.7 s^{-1}, K_M = 31.6 μM). The similarity of kinetic parameters is congruent with structural predictions, that despite αF258's close proximity to the active site, the Phe residue can be replaced by the similar Trp without significant impact on the structure or activity.

In the absence of significant structural rearrangement of αF258W compared with wild-type EAL, the next step was to assess the binding of the AdoCbl cofactor to this mutant by Trp fluorescence quenching. Figure 4 shows the normalized fluorescence quenching curve indicative of cofactor binding. Also included in Figure 4 is a fit of the data to Equation 11.

$$1 - \frac{F}{F_0} = \alpha\left(\frac{[L]}{[L]+K_D}\right) + (1-\alpha)\left(\frac{[L]^{1.41}}{[L]^{1.41}+30.4^{1.41}}\right) \qquad (11)$$

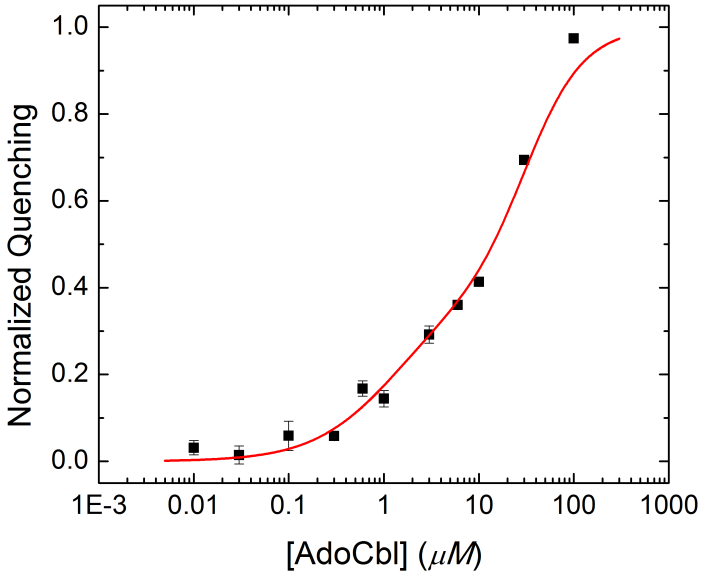


Figure 4. Fluorescence Quenching of αF258W by AdoCbl. Excitation at 280 nm, emission at 340 nm. Data shown with standard deviation of 3 experiments. Fit of data to Equation 11 shown in red.

Equation 11 is a modified form of the Hill equation in which F represents the fluorescence at each concentration of AdoCbl, [L], with the subscripted 0 referring to the condition without added AdoCbl. This normalized fluorescence quenching is dependent upon a fraction, α, of quenching caused by AdoCbl binding with a dissociation constant of K_D, and a fraction caused by the inner filter effect. The numerical values for the inner filter effect term were obtained from free Trp quenching experiments discussed earlier. The best fit parameters are given in Table 5. From the fit, 38% of the fluorescence is quenched by cobalamin binding, with the remaining fraction of fluorescence being screened by the inner filter effect. Each Trp is quenched to a different degree by the cobalamin. Recall, EutB showed no fluorescence quenching before the onset of the inner filter effect, so it is likely that the two Trp residues on EutB are too distant for quenching

by the binding cobalamin to be detected. Thus, in αF258W, the three remaining Trp residues (or a subset of them) are quenched by the cobalamin, which is responsible for the effect seen below 10 μM in Figure 4. The identity of the quenched Trp is not essential for the interpretation of the data. However, αW258 is likely to be efficiently quenched by the proximal AdoCbl. In addition, resonant energy transfer interactions may also serve to reduce the fluorescence of other nearby Trp residues. αF258W shows a red-shifted emission band (λ_m = 337 nm) compared to wild-type (λ_m = 335 nm), which implies that αW258 is the lowest energy Trp present, and can serve as a low energy trap for the other excited Trp residues. Thus, quenching of other, more distant Trp residues may also occur through αW258.

Table 5. Best-fit Parameters for αF258W Quenching.*

α	K_D (μM)
0.38 (0.28, 0.47)	1.2 (0.2, 2.2)

* 95% confidence bands shown in parentheses.

From the fitting, the binding constant of AdoCbl was found to be 1.2 μM. This is comparable to the previous estimate of 1.5 μM for K_D from K_M.[122] The congruence of K_M and K_D indicates that, according to Equation 3, k_2 is small compared with k_{-1}. Using and literature estimates for K_M (1.5 μM)[122] and k_2 (30 s^{-1}),[81] the simple model in Equation 3 predicts a k_1 of 100 $\mu M^{-1}s^{-1}$, and a k_{-1} of 120 s^{-1}. The 1.2 μM dissociation constant is weaker than the 170 nM binding constant in MCM.[117] However, MCM binds AdoCbl in the base off binding mode where a protein histidine directly ligands the central cobalt. This binding mode is not present in EAL, and can partly explain the difference in binding.

3.4 Summary

Trp fluorescence quenching has been used to make a direct measurement of the binding of AdoCbl to EAL. This fluorescence quenching technique has been used to study cofactor binding in MCM in a label-free approach.[117] In EAL, the native Trp residues are too distant from the cofactor binding site to report on cobalamin binding directly. However, a conservative Phe→Trp mutation at EutB position 258 placed an environmentally sensitive fluorescent probe within 10 Å of the cofactor binding site without significantly altering the structure of EAL. Quenching measurements on this mutant gave a dissociation constant of 1.2 μM, which is believed to be same for wild-type EAL. This approach demonstrates the power of site-directed mutagenesis and the previously developed structural model for answering specific questions about EAL, such as the binding affinity for the AdoCbl cofactor, and the cofactor release rate during catalysis.

Chapter Four

Towards a Reactive Cysteine-Free EAL

4

4.1 Overview of Site-Directed Spin Labeling

The technique of site-directed spin labeling (SDSL) allows the placement of a spin label reporter group at targeted sites on the exterior of the protein.[135] These spin label reporters are exquisitely sensitive to changes in correlation times from 10^{-7} to 10^{-11} s in cw-EPR at X-band frequencies.[66] Changes in correlation times are manifested as increases in linewidth and spectral breadth as the dipolar hyperfine coupling is less efficiently averaged with decreased motion. The changes in motion detected by the spin label may be caused by changes in the dynamics of the protein, or from conformational changes that restrict spin label movement.

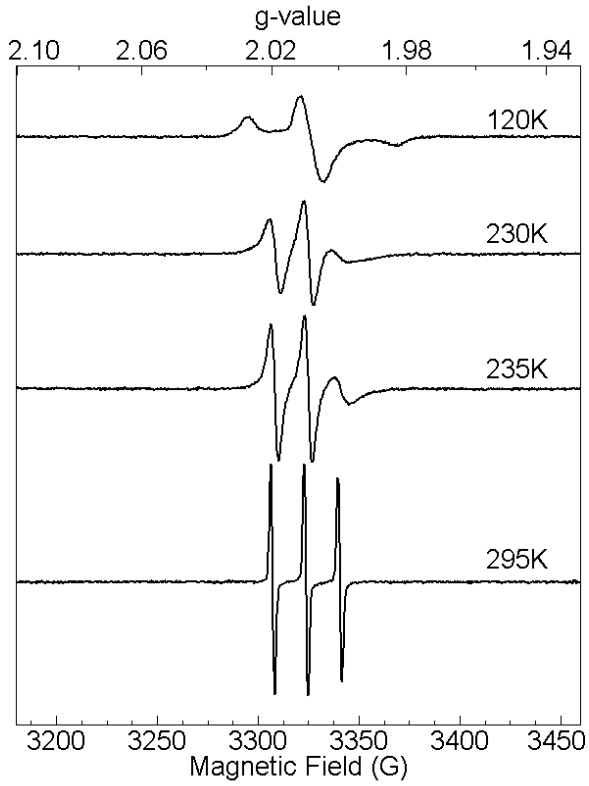


Figure 1. cw-EPR Spectra of 4MT spin label in 41% DMSO Cryosolvent at selected temperatures. 100 μM 4MT in 41% DMSO-

water cryosolvent. Temperatures as indicated. Experimental parameters: Frequency, 9.348 GHz. Power, 2 mW. Modulation frequency, 100 kHz. Modulation amplitude, 2 G. Dwell time, 10.24 ms. Time constant, 2.56 ms. 25 scans.

An example of these spectral changes is given in Figure 1, which shows the changes in lineshape of the 4MT spin label in 41% DMSO-water cryosolvent as a function of temperature. This cryosolvent has a known viscosity given by Equation 1.[136]

$$\log(\eta) = -5.3 + \frac{1600}{T} \qquad (1)$$

The viscosity, η, in cP, depends on the sample temperature, T. In the sample shown in Figure 1, the change in temperature increases the viscosity of the system from 1.3 cP at 295 K, to a value of 45 cP at 230 K. Increased viscosity reduces the correlation time of the spin label as predicted by the Stokes-Einstein relation. At 120 K, the spin label is relatively immobile, and the lineshape begins to approach its rigid limit value. Table 1 shows the correlation times (τ_c) obtained from fitting these experimental spectra to a freely tumbling nitroxide model using EasySpin.[137] At 295 K the spin label rapidly reorients with a τ_c of 89 ps. As the temperature is lowered, spin label motion is slowed yielding a τ_c of 2.4 ns at 230 K. EasySpin does not explicitly account for all spectral broadening, and some residual broadening is evident in the phenomenological linewidth parameter which increases with decreasing temperature. With such simulations, macroscopic physical information can be gleaned from spin label spectra that report on the dynamics of the system.

Table 1. Best fit parameters for simulated spectra in Figure 1.*

Temperature (K)	Correlation Time $\log(\tau_{corr})$	Linewidth (FWHM, mT)	A (MHz)
120	-8.106	0.852	11, 2, 114
230	-8.622	0.261	7, 25, 112
235	-8.679	0.240	14, 26, 102
295	-10.052	0.220	11, 37, 93

* Simulation done using EasySpin chili simulation function with the five parameters indicated floated.[138]

4.2 Site-Directed Spin Labeling of EAL

Recall, EAL is an enzyme that employs radicals for catalysis.[79] These EPR-visible radicals have been extensively studied to reveal active site structure,[37, 48, 49, 80] and time-resolved studies have elucidated rates of chemical steps within the catalytic cycle.[45-47, 53, 81] During EAL catalysis, the cob(II)alamin-substrate radical pair accumulates to over 90% during steady state turnover with *(S)*-2-aminopropanol.[39] This occurs because the next step in catalysis, radical rearrangement, is rate-limiting.[39] This radical pair state can be cryotrapped, and the lineshape examined using low temperature cw-EPR.[37, 39, 112] This line shape possesses a cob(II)alamin feature at $g_{\perp}$ =2.35 and a g=2.00 substrate radical signal.[37] Figure 2 shows this characteristic lineshape of wild-type EAL that was cryotrapped during turnover of *(S)*-2 aminopropanol.

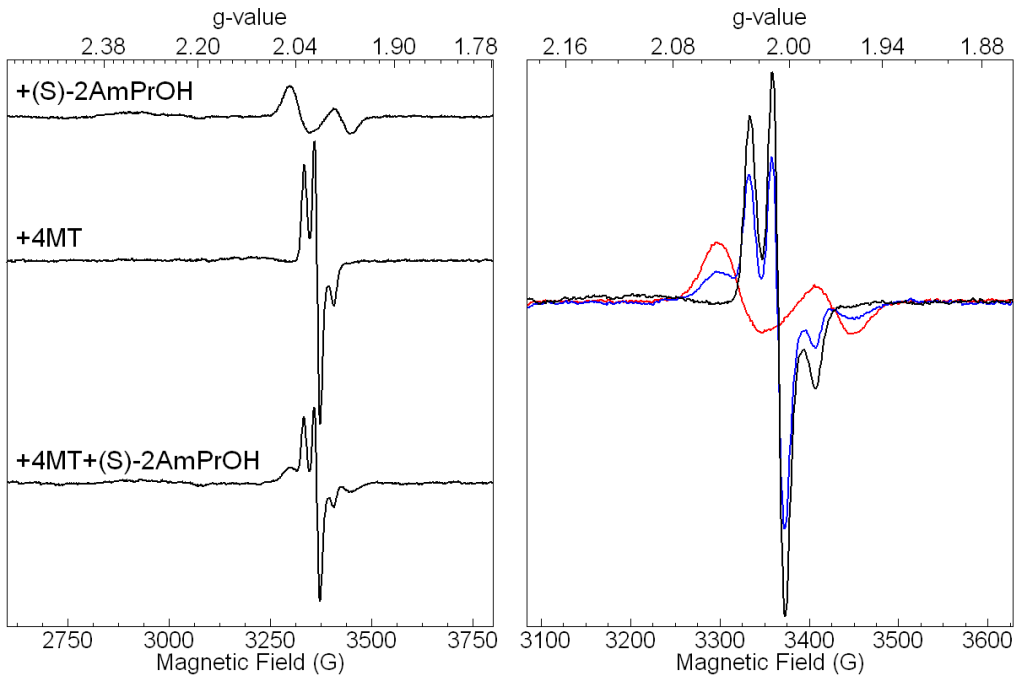


Figure 2. cw-EPR spectra of cryotrapped EAL, *(S)*-2 Aminopropanol and 4MT at 120 K. (Left) From top: (1) 60 μM EAL, 120 μM B_{12} and 20 mM *(S)*-2 aminopropanol, (2) 60 μM EAL, 120 μM B_{12} and 60 μM 4MT, (3) 60 μM EAL, 120 μM B_{12}, 60 μM 4MT and 20 mM *(S)*-2 aminopropanol. (Right) Same samples as (Left) with spectra overlaid to show *g*=2.00 spectral overlap. (1) shown in red, (2) in black, and (3) in blue. Experimental parameters: Frequency, 9.454 GHz. Power, 6.4 mW. Modulation amplitude, 12 G. Modulation Frequency, 100 kHz. Time constant, 40.96 ms. Dwell time, 0.163 s. 4 scans.

SDSL allows the concomitant monitoring of the chemical steps of EAL catalysis with motions distal to the active site by the attachment of an EPR-active spin probe to a site on the exterior of the protein. Figure 2 shows the cw-EPR spectrum of spin labeled wild-type EAL that was cryotrapped during turnover of *(S)*-2 aminopropanol. The individual radical pair spectra, and spin labeled EAL spectra are shown for comparison. In the combined sample, both the radical pair spectrum and the spin label spectrum are apparent, with the spin label overlapping the substrate radical at *g*=2.00. With this

approach it is possible to begin the study of motions that are correlated with the chemical steps of EAL catalysis.

The 4MT spin label conjugates with proteins by forming a covalent linkage with the sulfhydryl group of cysteine residues. EAL has 10 cysteines on the α-subunit, and 6 on the β-subunit. Of these, αC361, αC363, and αC364, form a highly conserved CDCC motif in the active site. Also, βC260 makes contact with the AdoCbl cofactor.[73] These four cysteines are presumed to be critical to the proper functioning of EAL due to their location within the active site, and cofactor binding site of the protein. Thus, their modification may lead to the loss of activity. In the wild-type spin labeled EAL spectra shown in Figure 2, the location of the spin label is not known. Labeling can occur on any of the 16 available cysteines that are native to *S. typhimurium* EAL. Since spin labeled EAL is still catalytically competent, and able to form the native substrate radical pair state, the site of spin label conjugation cannot be a critical cysteine residue whose modification would lead to a loss of activity. However any of the other presumably non-critical cysteines are further removed from the active site, and may undergo chemical modification without a loss of activity. The elimination of these reactive cysteines is necessary to label EAL at a specific engineered site.

4.3 Development of a Reactive Cysteine Free EAL

Armed with the homology modeled structure and protocol for high-throughput purification of EAL described in the Chapter Two, the next major goal is to develop a system for site-directed spin-labeling of EAL. Both the MTSL and the 4MT spin labels, described in Chapter One, conjugate to free sulfhydryl groups on the protein. A careful

study was undertaken to determine the reactive sulfhydryls in EAL, and to determine which of these could be removed with minimal effects on protein activity.

4.3.1 Solvent Accessibility of Cysteines in EAL

Residues with high solvent accessibility lie close to the surface of EAL. As such, it is reasonable to assume that they will react more readily with labeling reagents in solution. If labeling reagents are added in limiting quantities, the solvent exposed sites should be preferentially labeled. Thus, the first reactive cysteine is predicted to have high solvent accessibility. In addition, this consideration allows for a more focused approach to mutagenesis. Rather than the whole-scale elimination of all 16 cysteines in EAL, which is unlikely to be successful in retaining activity, all the fast-reacting solvent exposed cysteines may be removed and replaced by a single fast-reacting cysteine at a chosen site. Thus directed labeling would be achieved, while maintaining enzymatic activity.

To this end, the solvent accessibility of EAL residues was assessed using the Accessible Surface Area View web server (ASAView).[90] ASAView uses Define Secondary Structure of Proteins (DSSP)[91] to calculate the surface area of each residue that is accessible to solvent, and scales this value by the total residue surface area. Table 2 shows the solvent accessibility of each cysteine residue in wild-type EAL. Of these, αC71, αC283, βC158 and βC260 have greater than 10% solvent accessibility. αC129 has 3.1% solvent accessibility. The remaining cysteines have less than 1% solvent accessibility. βC37 and βC46 were not included in the modeled structure and were excluded from this analysis. From this accessibility analysis, a list of candidates of fast reacting cysteine residues was identified.

Table 2. Solvent Accessibility of Cysteine Residues in wild-type EAL.

Residue	Solvent Accessibility[†]
αC71	0.140
αC129	0.031
αC159	0
αC223	0.007
αC249	0
αC283	0.130
αC361	0
αC363	0
αC364	0
αC388	0.007
βC37*	--
βC46*	--
βC81	0
βC158	0.100
βC242	0
βC260	0.440

[†]Solvent accessibilities calculated by DSSP.[91]
* These residues are absent in the crystal structure of *E. coli* EAL,[43] and are also absent from the modeled structure of *S. typhimurium* EAL.

4.3.2 Single Mutant Library of EAL

After predicting the fast reacting cysteine residues based on solvent accessibility, systematic mutagenesis was done to determine how critical each one of these cysteines was to enzyme activity. Alanine was selected as the amino acid of choice for the creation of the EAL single mutant library. Alanine scanning is a molecular biology technique that helps determine the role of specific residues of a protein by the sequential replacement of each residue by alanine.[139] Alanine is an inert amino acid with a methyl side chain, which mimics the secondary structure of many other amino acids. This is in contrast to the chemically similar amino acid glycine, (hydrogen substituent as its side chain), which is also inert, but is unsuitable because it forms unique secondary structures.

Cysteine→alanine mutants are unlikely to greatly influence protein structure. In addition, the modest alanine mutation usually does not add any non-native interactions. Serine, an amino acid with a hydroxyl terminus, is also used as a target residue for SDM, but may form hydrogen bonds that were not present in the original structure.

Table 3. Primers used in the Development of the EAL Single Mutant Library. *

Primer Name	5' to 3' Primer Sequence
C71A FP	GCC TAT GAA GAG GAC GCT GTG ACG CGC CTG
C71A BP	CAG GCG CGT CAC AGC GTC CTC TTC ATA GGC
C129A FP	GGC AGT AGC GAA AAT CGC TTC CAA CGC TGA
C129A BP	GGT CAG CGT TGG AAG CGA TTT TCG CTA CTG CC
a_C283A_FP	CGC CGG GGA AAA CGC GCT GTA CTT TGA AAC CG
a_C283A_BP	CGG TTT CAA AGT ACA GCG CGT TTT CCC CGG CG
b_C37A_FP	GGC GCA AAG CCG CAG GCG GCC GCG CCG ACG G
b_C37A_BP	CCG TCG GCG CGG CCG CCT GCG GCT TTG CGC C
b_C46A_FP	GAC CGA AAG CGC CGC GCT GGA TTT AGG
b_C46A_BP	CCT AAA TCC AGC GCG GCG CTT TCG GTC
b_C158A_FP	CGC TGA AGT CAC AGG CGG TGA TGA ACC CGG
b_C158A_BP	CCG GGT TCA TCA CCG CCT GTG ACT TCA GCG
b_C261A_FP	GGC CGA CAG AAC CGC GAT TTC AAA CAT TC
b_C261A_BP	GAA TGT TTG AAA TCG CGG TTC TGT CGG CC

* Primers obtained from Integrated DNA Technologies

With candidates for the primary reacting cysteine in EAL obtained from the solvent accessibility study, a library of single cysteine→alanine mutants of EAL was created containing αC71A, αC129A, αC283A, βC37A, βC46A, βC158A and βC260A. The EALH6 pBR322 plasmid (described previously) was isolated from single colony growth of wild-type EALH6 using QIAGEN Plasmid Mini kit. Primers containing the desired mutations (Table 3) were used for mutational PCR with the Stratagene QuikChange II XL Site-Directed Mutagenesis Kit according to manufacturer's instructions. The parental EALH6 plasmid was digested with *DpnI*, and the mutant plasmid was transformed into

XL10 Gold competent cells. Single colonies of the mutant bacteria were grown, and the plasmid extracted for sequencing by Cogenics using the sequencing primers shown in Table 4. Mutation was considered successful when 100% of the desired sequence identity within the EAL reading frame was obtained.

Table 4. Sequencing Primers for EAL.*

Primer Name	5' to 3' Primer Sequence
EAL SP1	TTG ACA ATT AAT CAT CGG CTC G
EAL SP2	AAT GCC GGT GAT CAA AAA AG
EAL SP3	TGG AAG CGC GTA ACT ACG
EAL SP4	ATG GCG TCA ATG GGA CAG
EAL SP5	TGC GAT CAC CGC CAA CTA TGA AGA GAT CCT

* Primers obtained from Integrated DNA Technologies and Invitrogen

Since the sites selected for mutations in this library were all solvent exposed, they were considered to be amenable to mutagenesis since the change should not introduce significant structural rearrangements. This was confirmed as the mutant proteins could be expressed and purified successfully according to the protocol described in Chapter Two for wild-type EAL without insolubility or folding problems. The mutants were subjected to kinetic analysis using the standard ADH-NADH coupled assay,[2] and the activity parameters for the single mutant library of EAL is shown in Table 5. This shows that all but two mutants in the single-mutant library have within 25% of the wild-type activity, and K_M does not change by more than a factor of 2 for all the mutants tested. αC129A has reduced activity, which may be because of its comparatively buried nature. The solvent accessibility of αC129 is the lowest of those selected for mutation. It is unlikely that this is the primary site of reaction with 4MT because of this reduced activity. βC260A is inactive. This can be understood by a close inspection of the EAL structure. βC260

makes contact with the B_{12} cofactor in wild-type EAL. Mutating this residue may cause changes in the active site that render the mutant inactive. Similar to αC129, this site is not expected to be the primary 4MT reaction site.

Table 5. Activity parameters for selected single Cysteine to Alanine Mutants of EAL.*

Enzyme	k_{cat} (s^{-1})	K_M (μM)
Wild-type	21.8 ± 7.1	31.4 ± 5.7
αC71A	27.6 ± 4.8	15.0 ± 4.2
αC129A	4.0	
αC283A	18.6	37.9
βC37A	22.1 ± 1.1	33.2
βC46A	21.7 ± 2.9	35.8 ± 4.5
βC158A	16.1 ± 3.1	47.8 ± 4.7
βC260A	0	--

* Activity measurements done at 295 K with the NADH-ADH coupled assay.[2]

The EAL single mutants, αC71A, βC37A, βC46A and βC158A were further analyzed by 4MT spin labeling at 295 K. Figure 3 shows the resulting spin label spectra, scaled to match the intensity of the absorbance integral. This integral is proportional to the number of spins and should be equal for the samples presented here. Fitting these tethered spin label spectra to extract correlation times requires the assumption of a specific rotational model; however, several model-free parameters are evident from the data. The spectra appear similar, with the exception of βC37A which is significantly broadened, and thus shows reduced peak-to-trough amplitude on a derivative spectrum. If the high field, m_I = -1 peak at g=2.00 is examined, its intensity is reduced to 0.23 compared with the m_I = 0 peak. For comparison, all of the other spectra have a scaled high-field peak intensity of 0.36±0.02. This is indicative of the slowed rotational motion present in βC37A, as the

high-field peak is acutely sensitive to broadening.[60] The activity measurements eliminate the possibility of significant structural changes, making the likely explanation for the decreased spin label mobility a change in the labeling site of βC37A compared with wild-type and the other single mutants. This, in turn, implies that the majority of spin labels attach to βC37 in wild-type EAL. Thus, the primary spin labeling site of wild type was identified.

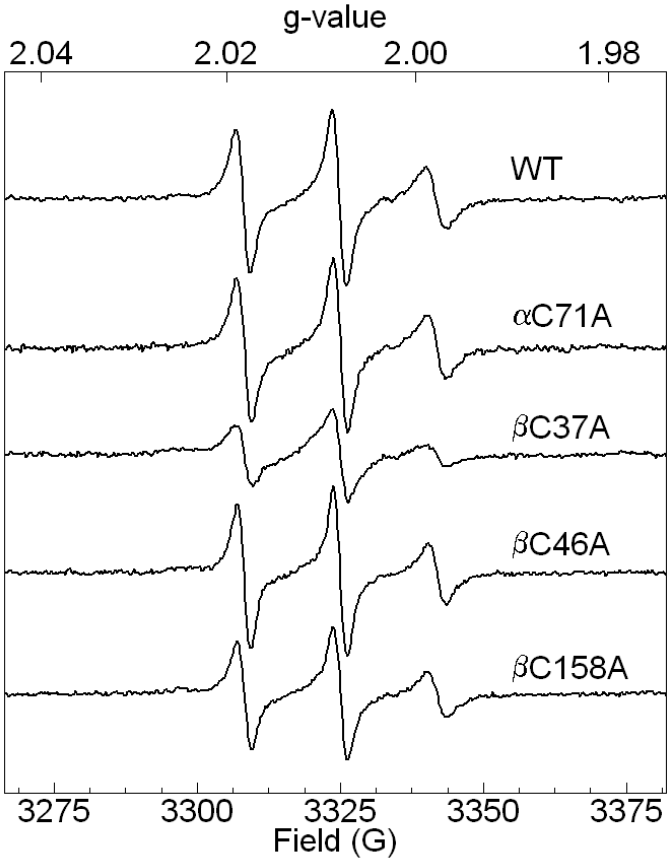


Figure 3. cw-EPR Spectra of 4MT spin labeled EAL single-mutant library at 295 K. Samples contain 60 μM protein, 120 μM B_{12}, 60 μM 4MT. Spectra have been scaled to match the absorbance integrals. Experimental parameters: Frequency, 9.344 GHz. Power, 0.2 mW. Modulation frequency, 100 kHz. Modulation amplitude, 2 G. Dwell time, 10.24 ms. Time constant, 2.56 ms. 36 or more scans.

βC37 is not visible within the modeled structure of EAL, because it is present on the cleaved leader sequence of EutC in the *E. coli* structure. Thus, only a general idea of the position of this residue is possible. If the polypeptide chain is fully extended, βC37 is 32 Å away from the first modeled residue, βA47. It is unlikely that the chain is fully extended, as even "unstructured" proteins form random coils rather than an entropically unfavorable extended chain. For comparison, if the unmodeled region was alpha-helical, it would only extend 18 Å from βA47. Figure 4 shows the possible locations of βC37, assuming that it is up to 32 Å away from βA47. From visual inspection, the primary labeled cysteine appears to lie on the surface of EutB, but not necessarily near to the active site.

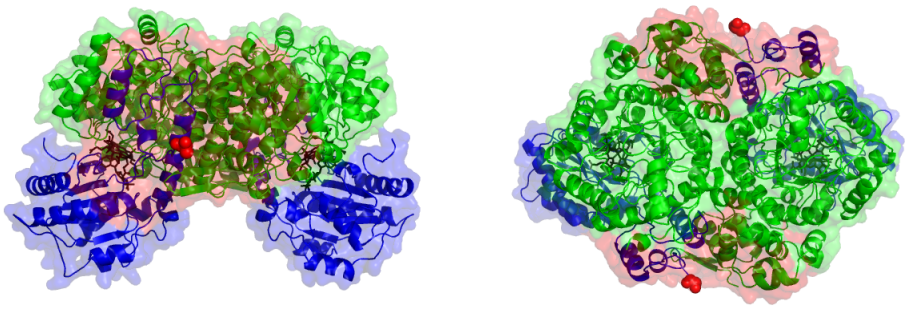


Figure 4. PyMOL structure of possible positions of βC37. PyMOL cartoon structure of EutB (green) and EutC (blue) $(\alpha\beta)_2$ dimer, with cobalamin shown as black sticks. Position of EutC N-terminus (βA47) shown as red spheres. Possible position of βC37 assuming a fully extended 32 Å polypeptide chain shown as a red surface. Unmodeled region is unlikely to be fully extended. (Right) is same as (Left) but rotated 90° along the horizontal axis.

4.3.3 Development of the Solvent-Accessible-Cysteine Free EAL mutant, EAL-5CF

Information on the activity of single cysteine→alanine mutants of EAL directed the development of an EAL with all solvent accessible cysteines removed, while retaining wild-type-like activity. Using the same protocol described for the development of the EAL single mutant library, a series of single mutations were done on βC37A using the primers in Table 3 to develop such an EAL mutant. The final construct, EAL-5CF, contained the mutations βC37A, βC46A, αC71A, βC158A, and αC283A. αC129 was omitted because it significantly affected activity, and its mostly buried nature (3.1% solvent accessibility) suggested that it belonged to a second class of more slowly reacting cysteines. βC260 was also omitted because mutation here eliminates enzymatic activity. However, this residue makes contact with the B_{12} cofactor. Thus, if B_{12} is bound prior to spin labeling, this residue may be protected from reaction by the presence of the cofactor.

Purified EAL-5CF enzyme yields from 37 °C cell growth and induction in XL10 Gold cells for this five-mutation construct were lower than wild-type purified by the same protocol, but these yields could be increased by growth at 30 °C. Kinetic analysis showed that EAL-5CF had similar parameters to wild-type EAL (Table 6). The steady state k_{cat} value was equal to that of wild-type, whereas the K_M value was a factor of 3 higher. This implies that the structure is not significantly affected by the mutations.

Table 6. Activity parameters for EAL-5CF and its mutants.*

Enzyme	k_{cat} (s^{-1})	K_M (μM)
Wild-type	21.8 ±7.1	31.4 ± 5.7
EAL-5CF	22.7 ± 0.4	79.6 ± 4.1
αD207C	10.1 ± 0.6	43.5 ± 3.7
αM434C	10.0 ± 0.1	53.2
αN62C	6.8 ± 5.1	110 ± 43.3

* Activity measurements done at 295 K with the NADH-ADH coupled assay.[2]

Figure 5 shows an EPR spectrum of the reaction of EAL-5CF with the 4MT spin label at 295 K. The spectrum is significantly different from the relatively mobile wild-type EAL spin label spectrum. There are two dynamical components evident in the EAL-5CF spectra; a mobile fraction, and a more immobile fraction. Increasing the number of equivalents of 4MT disproportionately increases the intensity of the mobile fraction, thus the spectra cannot be overlaid. This is unlike wild-type EAL, where increasing 4MT equivalents to 3:1 still yield spectra that overlap with the 1:1 labeled spectra. This suggests that in EAL-5CF the addition of the second and third equivalents of 4MT preferentially labels sites in a different motional class from the first equivalent. Interestingly, this site yields more mobile spin label spectra, rather than the immobile spectra expected from a buried site. It is possible that extensive labeling causes some denaturation of the protein, and leads to the increased mobility seen in these spectra.

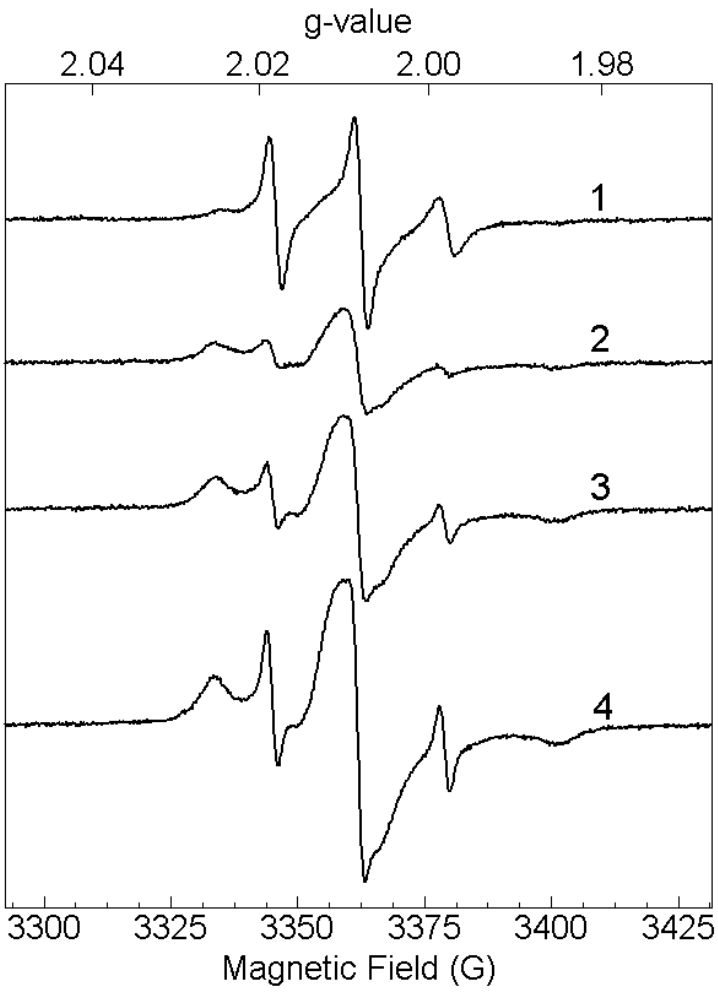


Figure 5. cw-EPR Spectra of 4MT spin labeled EAL-5CF at 295 K. (1): 60 μM EAL with 120 μM AdoCbl and 60 μM 4MT. (2) to (4): 60 μM EAL-5CF with 120 μM AdoCbl and (2) 60 μM 4MT, (3) 120 μM 4MT, (4) 180 μM. EAL-5CF spectra cannot be overlaid upon rescaling for 4MT concentration. Experimental parameters: Frequency, 9.451 GHz. Power, 2 mW. Modulation frequency, 100 kHz. Modulation amplitude, 2 G. Dwell time, 10.24 ms. Time constant, 1.28 ms. 49 or more scans.

Table 7 shows the results of fitting the 1:1 EAL-5CF : 4MT lineshape with a two component model using EasySpin.[137] It shows that 74% of the spin label has an apparent τ_c of 10 μs, compared with a faster fraction with a τ_c of 19 ns. In contrast, a τ_c of 0.27 ns was obtained for the single component present in 4MT spin labeled wild-type EAL. Thus,

both motional classes present in EAL-5CF are not as mobile as the βC37 labeling site in wild-type EAL. Since the solvent accessible cysteines were removed in EAL-5CF, this spectrum represents the background reactivity of the 11 remaining cysteines in this mutant.

Table 7. Multiple Component Fit of EAL-5CF 4MT Spin Label Spectrum.*

	Component 1	Component 2
***g*-value**	2.032, 2.030	2.035, 2.025
Linewidth (mT)	0.820	1.728
log(τ_{corr} (s))	-7.727	-4.998
A (MHz)	13.2, 112	39.3, 173
Weight	0.884	2.56

* EasySpin fit of 1:1 EAL-5CF 4MT spectrum shown in Figure 5.[138] *A* and *g* treated as axial.

4.3.4 Development of EAL-5CF single mutants

After the development of the EAL-5CF mutant, the next step was to engineer sites for spin label attachment that react more readily than the buried cysteines. The modeled structure was examined using PyMOL, to locate solvent accessible sites that were candidates for mutation. First, the temperature factors of the *E. coli* crystal structure were examined.[43] The Debye-Waller temperature factor, or β-factor, is an experimental measurement of the thermal motion present in an X-ray crystal structure. It is a measure of the exponential decay of the electron density used to locate the position of the atoms. These temperature factors cannot be directly compared across different structures, but within a single protein structure they are indicative of the relative vibrational motion of that atom.[140] Figure 6 shows the β-factor of the C-α atoms of EutB and EutC in the 3ABO

E. coli crystal structure.[43] Residues with more than 50% solvent accessibility[90] are highlighted in red. It is readily apparent that there are solvent accessible sites of both high and low mobility. From this information residue αD207 was selected, as a potentially immobile, solvent accessible site. It is present on an α-helix on the exterior of the EutB subunit as is shown in Figure 7.

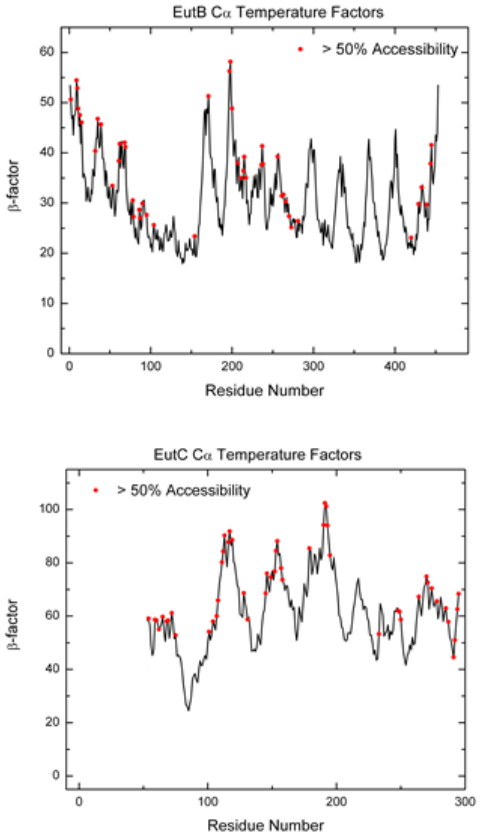


Figure 6. Temperature factors of C-α atoms of 3ABO crystal structure. Temperature factors of EutB (top) and EutC (bottom) subunits in PDB ID: 3ABO *E. coli* EAL crystal structure[43] with residues with greater than 50% solvent accessibility[90] indicated by red dots. The EutC subunit of EAL had its N-terminus truncated for crystallization, thus no temperature factor data is present for these atoms.

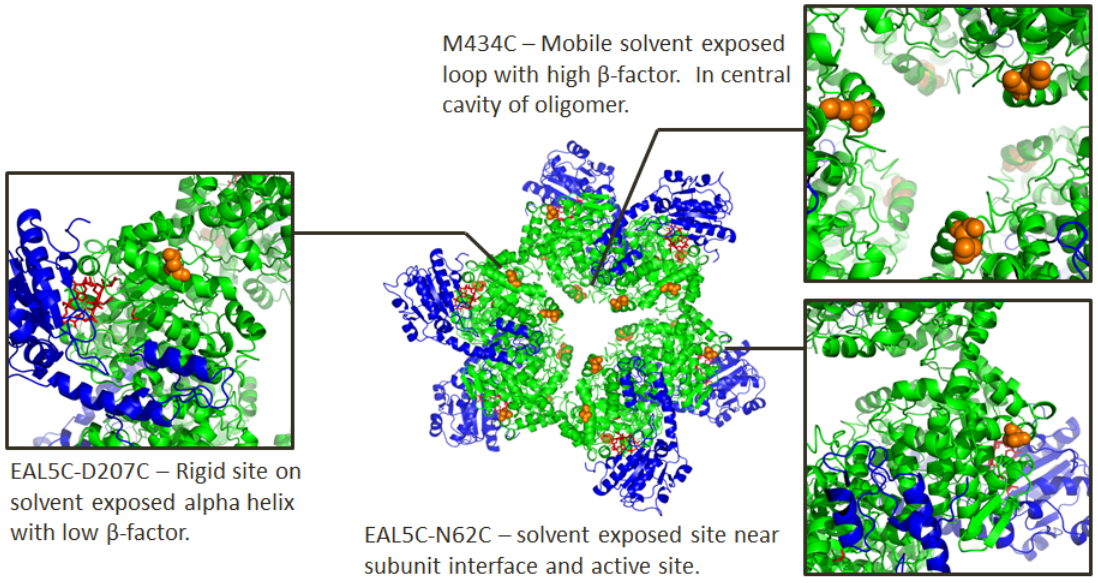


Figure 7. Location of sites selected for mutations on EAL-5CF. PyMOL cartoon depiction of EAL showing the sites selected for mutation in EAL-5CF. EutB shown in green, EutC shown in blue, cobalamin shown as red sticks. Mutation site shown as orange spheres.

Also present in the crystallographic structures is residue occupancy information. Briefly, each atom has a predicted electron density. If a protein is crystallized in multiple conformations, it is possible that some atoms appear in distinct locations. It is also possible that only one of these locations is detected (with reduced electron density) whereas the other location is below the signal to noise ratio. In 3ABO EAL, αM54, αM307, αM351, αM359 and αM434 have 75% occupancy.[43] This indicates that these residues exist in multiple conformations of which only one is visible under the experimental conditions required for crystallization. Of these residues, only αM434 is significantly solvent exposed. As such, this residue was selected as a potentially mobile

site for further study. It is located in a loop region in the central cavity of the oligomer as shown in Figure 7.

αN62 was also selected as a site for further study. It is located in a solvent exposed region near the subunit interface between EutB and EutC. Figure 7 shows its location in the EAL structure. αN62's proximity to the active site makes it an interesting candidate for study using spin label probes during catalysis.

Table 8. Primers used in the Development of targeted labeling sites for EAL-5CF.*

Primer Name	5' to 3' Primer Sequence
aN62Cfp	GGC GGA TAT CCG CAA CTG CCC GGT GAT TGC C
aN62Cbp	GGC AAT CAC CGG GCA GTT GCG GAT ATC CGC C
aF258Cfp	GGG CTT AAA AGA GTG CGG CGT CGA GCT GGC C
aF258Cbp	GGC CAG CTC GAC GCC GCA CTC TTT TAA GCC C
aD197Cfp	CCC GGT GAC CGA TTG CGT GGA GAA CCT G
aD197Cbp	CAG GTT CTC CAC GCA ATC GGT CAC CGG G
aD207C Slow FP	CCC GCG TGC TCT GCA CCG TTT ACG GCG
aD207C Slow BP	CGC CGT AAA CGG TGC AGA GCA CGC GGG
aM434C Fast FP	GGC TGG AAA CGT GCG GCA TTA TGG C
aM434C Fast BP	GCC ATA ATG CCG CAC GTT TCC AGC C
aM434C Fast FP (extend)	GAA CGC TGG CTG GAA ACG TGC GGC ATT ATG GCA AAC GGT CG
aM434C Fast BP (extend)	CGA CCG TTT GCC ATA ATG CCG CAC GTT TCC AGC CAG CGT TC

* Primers obtained from Integrated DNA Technologies

Using the same protocol described for the development of the EAL single mutant library, a library of single EAL-5CF mutants was created using the primers in Table 8. The sequences were verified by Cogenics using the sequencing primers in Table 4. These mutants could be expressed and purified in XL10 Gold cells using the protocol for wild-type expression described in Chapter Two. As with EAL-5CF, yields were improved if

grown at 30 °C. The activity of these mutants was assessed, and they all had reduced activity of approximately 50% of the wild-type value and up to a factor of 3 higher K_M (Table 6). The reduced performance of these mutants may be attributed to reduced structural integrity compared with wild-type EAL. Since the single EAL-5CF mutants contain 6 mutations, they can be considered to be extensively modified and their structures may be less stable than wild-type EAL. Nonetheless, a change in k_{cat} by a factor of 2 is modest when the exponential rate increases of enzymes are considered.

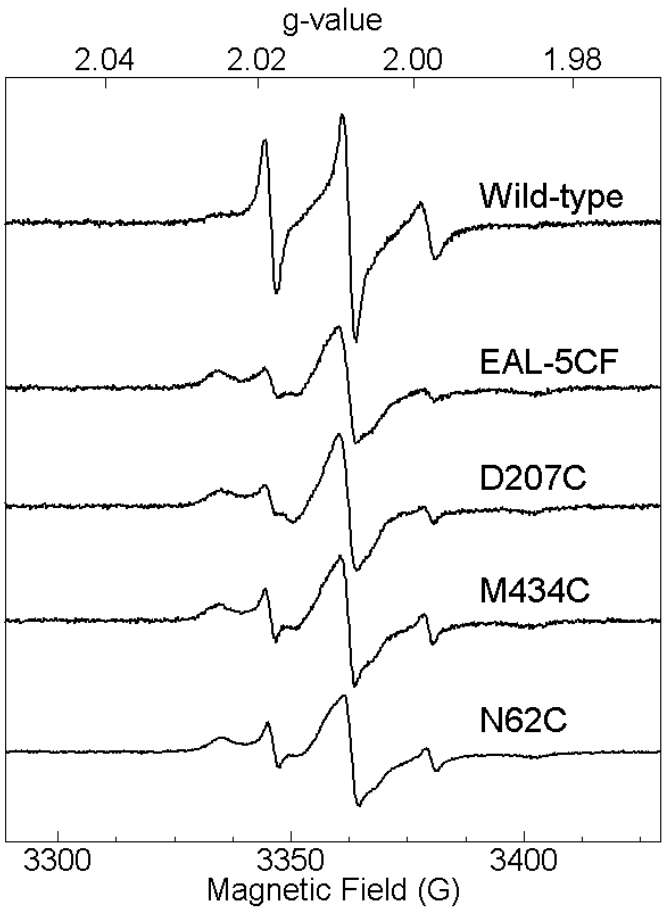


Figure 8. cw-EPR Spectra of 4MT spin labeled EAL-5CF mutants at 295 K. 1:2:1 enzyme : AdoCbl : 4MT. Spectra scaled to match the integral of absorbance. Experimental parameters: Frequency, 9.455 GHz. Power, 2 mW. Modulation frequency, 100 kHz. Modulation amplitude, 2 G. Dwell time, 10.24 ms. Time constant, 1.28 ms. 49 or more scans.

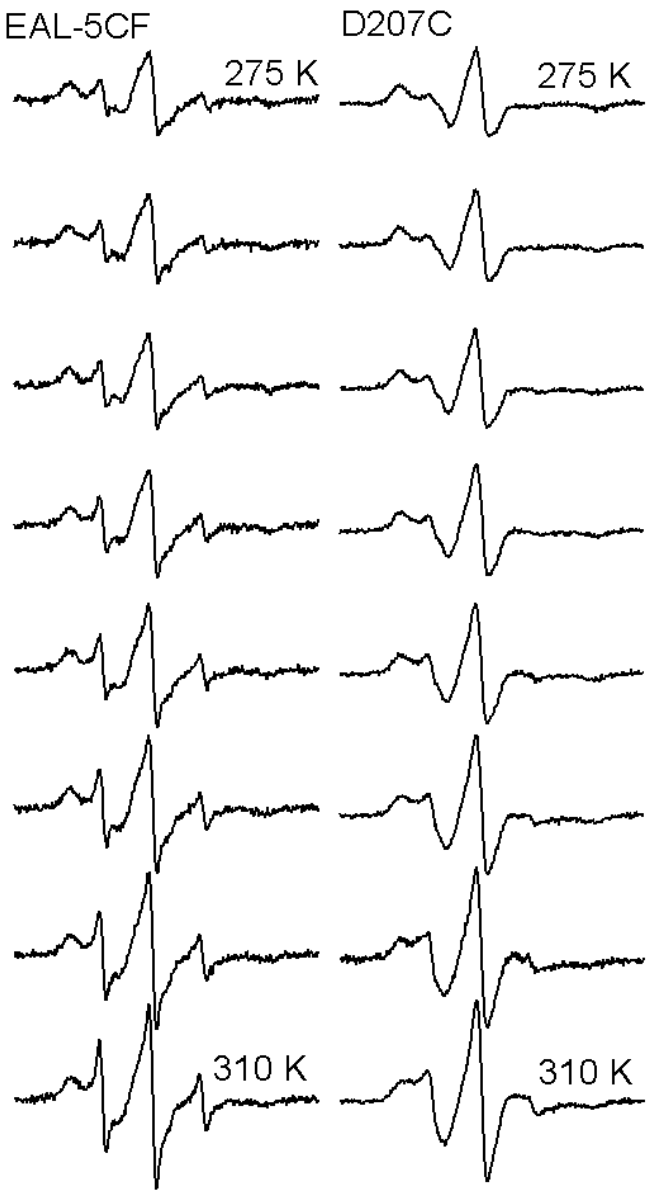


Figure 9. cw-EPR Spectra of 4MT spin labeled EAL-5CF and D207C mutants at various temperatures. (Left) 60 μM EAL-5CF, 120 μM AdoCbl, 60 μM 4MT. (Right) 60 μM D207C, 120 μM AdoCbl, 60 μM 4MT. Spectra from top to bottom are at 275 K to 310 K in 5 K steps. Spectra scaled to match integral of absorbance. Experimental parameters: Frequency, 9.452 GHz. Power, 2 mW. Modulation frequency, 100 kHz. Modulation amplitude, 2 G. Dwell time, 10.24 ms. Time constant, 2.56 ms. 49 scans.

Figure 8 shows the spin labeling spectra of the EAL-5CF mutants. The spectra share the 2-component lineshape of the EAL-5CF spectra, however, there are subtle differences, which are most apparent at *g*=1.995. These differences can be further probed by a temperature dependent study from 275 K to 310 K as shown in Figure 9. The spin-labeled αD207C spectrum shows an increase in the integral of the absorbance of 60% as the temperature is increased from 275 K to 310K. This is contrary to the expected relationship of EPR signal intensities (*I*) with temperature (*T*) given by Equation 2.

$$I_1 T_1 = I_2 T_2 \qquad (2)$$

However, this behavior can be understood in terms of a change in the conditions in the cavity under this temperature range. The signal voltage of recorded EPR spectrum (V_S) is directly proportional to the quality factor of the resonator, *Q* (Equation 3).[62]

$$V_S = \chi'' \eta Q \sqrt{P Z_0} \qquad (3)$$

V_S also depends on the magnetic susceptibility of the sample, χ'', cavity filling factor, η, the microwave power, *P*, and the impedance of the transmission line, Z_0. The quality factor is a measure of the energy stored in the resonator compared with the energy dissipated. When viewed as an RLC circuit, Q can be rewritten in terms of Equation 4.

$$Q = \frac{\omega_0 L}{R + r_{sample}} \qquad (4)$$

Here, the inductance of the resonator is given by *L*, *R* is the resonator resistance, and r_{sample} the sample resistance. ω_0 is the angular frequency. Clearly, for a lossy sample (high r_{sample}), an increase in r_{sample} will reduce *Q*, and consequently reduce signal amplitude. Biological aqueous samples have significant resistance because the complex

part of the permittivity of water is high at the microwave frequencies used in EPR, in turn leading to absorption by the water molecules.[141] The dielectric constant of water decreases from 275 K to 310 K, leading to reduction in the absorption of the sample over this temperature range. As such, the EPR signal intensity follows a trend that is opposite to what is expected. This has been observed previously in the work of Fung and Johnson who observed a 45% increase in signal intensity of spin labeled erythrocyte membranes in water from 277 K to 310 K.[142] αD207C samples with 30 mM aminopropanol show less than a 20% increase in amplitude across the temperature range, possibly due to the high concentration of substrate affecting the effective dielectric constant of the sample.

EasySpin two component fits of the EAL-5CF spectra revealed that the correlation times of both the fast and slow components were unchanged over the 275 to 310 K temperature range. However, the weight of the fast component increased from 3% at 275 K to 25% at 310 K. This suggests that spin labels in the slow motional class switch to the fast motion regime with increased temperature. This is consistent with a single site of spin label attachment with multiple conformational states. If the two components present in the spectra arose from multiple spin labeling sites, then temperature dependent interchange between each component would not be possible. Nonetheless, only 22% of the spin labels changed motional class over the temperature range examined. Thus, a majority of the spin labels may be spread across multiple sites.

Since the differences between EAL-5CF and its single mutant spectra are not as dramatic as expected, it is unclear whether the spin labels do in fact label the targeted site. A spectrum with multiple components does not necessarily indicate multiple labeling sites, but rather multiple conformations of the spin label which may be on a single or multiple

sites. Temperature-dependent analysis of the multicomponent spectra yielded mixed results, and the number of sites of spin-label attachment remains uncertain. Further experiments were done to ascertain the location of the spin label.

4.3.5 Location of the Spin Labels on EAL-5CF and its Mutants

4.3.5.1 Power Saturation Measurements

Another approach to assess solvent accessibility is through the use of cw-EPR power saturation measurements. Power saturation is a technique used to determine geometric properties of proteins and membranes that cannot be adequately studied using high-resolution structural methods.[143] It measures the depth of an attached spin label in a membrane or protein based on the relaxation caused by collision with paramagnets in the solution. Provided that these collisions occur faster than the spin-lattice relaxation rate, the solution paramagnet reduces the spin label T_1 by spin exchange, and decreases the saturation. Usually $T_2 << T_1$, and T_2 is not significantly affected. The rate of these collisions is determined by the solvent accessibility of the spin label. Two paramagnetic relaxation reagents, molecular oxygen and nickel ethylenediamine diacetate (NiEDDA) have complementary properties. Oxygen's apolarity allows it to penetrate deep within the protein, and out of the polar solvent, leading to a concentration gradient that varies with depth, with a low concentration on the surface, and a higher concentration on the interior. NiEDDA, is zwitterionic and fractionates in the opposite way, with a high concentration within the solvent, and low concentration within the membrane.[144] Fitting the saturation power of a sample with different relaxation agents using Equation 5, it is possible to determine T_1, which is usually proportional to solvent accessibility.[145]

$$A_S = I_S\sqrt{P}\left[1 + \left(2^{\frac{1}{\varepsilon}} - 1\right)\frac{P}{P_{1/2}}\right]^{-\varepsilon} \quad (5)$$

The derivative signal amplitude A_S is determined by a scaling factor, I_S, the microwave power, P, the saturation power, $P_{1/2}$ (power where the derivative amplitude is reduced to ½ its unsaturated value) and ε, a measure of the homogeneity of the resonance line, which varies from 1.5 for fully homogeneous to 0.5, for inhomogeneous. ε allows for the fitting of inhomegeneously broadened lines caused by slowed nitroxide tumbling.[146] $P_{1/2}$ is inversely proportional to T_1 and T_2 where γ is the gyromagnetic ratio as shown in Equation 6.[146, 147]

$$P_{1/2} \sim \frac{1}{\gamma T_1}\frac{1}{\gamma T_2} \quad (6)$$

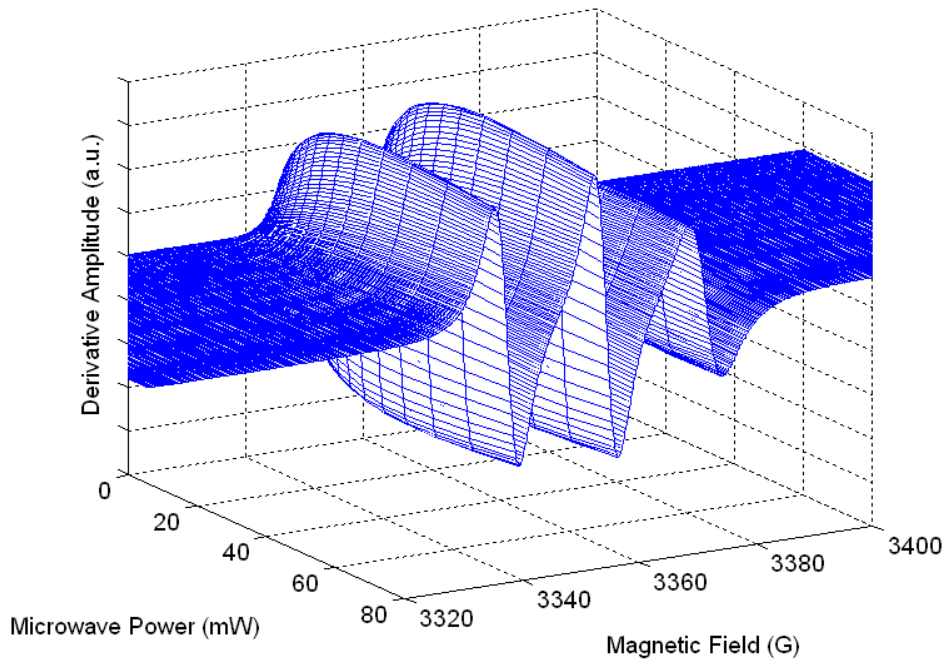


Figure 10. cw-EPR Power Saturation Spectra of TEMPO at 240 K. 100 µM TEMPO in water. Experimental parameters: Frequency, 9.456 GHz. Conversion time, 10.24 ms. Modulation Frequency, 100 kHz. Modulation Depth, 1.0 G. Time constant, 1.28 ms. 36 power points.

Figure 10 shows a power saturation curve for TEMPO at 240 K. Figure 11 shows the fit of the central peak of TEMPO to Equation 5 which determines the saturation power to be 36.6 mW, and an ε–value of 1.13 (Table 9). This indicates that the line does not have significant inhomogeneous broadening even at 240 K where the sample is frozen and its mobility is limited. When the spin labeled is attached to the protein, the saturation power decreases significantly, to 13.3 mW in wild-type protein, and to the lower value of 7.84 mW in EAL-5CF. This is believed to be caused by the spin label being buried within the protein in EAL-5CF, and on the surface in wild-type EAL. The change in spin label position causes a difference in spin label environment which in turn has an effect on T_1.

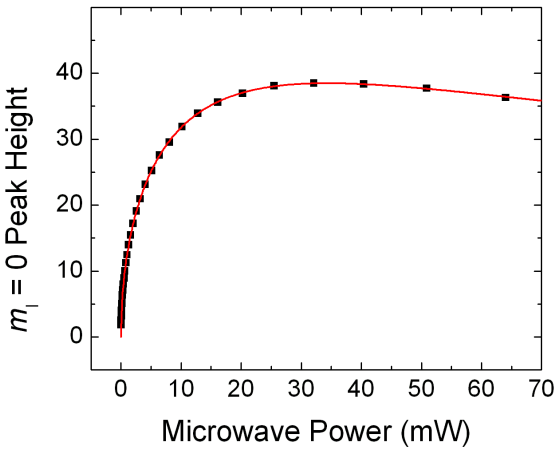


Figure 11. Fit of Power Saturation of TEMPO at 240 K. $m_I = 0$ peak-to-trough amplitude of 100 μM TEMPO fit to power saturation equation (Equation 5). Saturation power ($P_{1/2}$) found to be 36.6 mW. Full results of fit shown in Table 9.

Table 9. cw-EPR Saturation Power of EAL at 240 K.*

Sample	*I*	$P_{1/2}$ (mW)	ε
TEMPO	12.70 (12.67, 12.73)	36.63 (36.39, 36.86)	1.13 (1.10, 1.16)
Wild-type EAL	0.899 (0.875, 0.923)	13.3 (12.2, 14.3)	0.68 (0.63, 0.73)
EAL-5CF	1.099 (1.069, 1.130)	7.84 (7.23, 8.45)	0.73 (0.69, 0.77)

* 95% confidence intervals are given in parentheses. Enzyme samples labeled with 4MT. Sample concentration is 60 μM.

A more high resolution analysis of the solvent accessibility was desired, so NiEDDA was synthesized according to a published procedure,[144] with the resulting powder dried under vacuum for 18 h. EPR samples of spin labeled enzyme with molecular oxygen, and degased samples of spin labeled enzyme, and spin labeled enzyme with NiEDDA were made. Power saturation experiments were carried out at 295 K in liquid samples where the added relaxants could freely diffuse to shed light on the depth of the attached spin labels. Unfortunately, at 295 K, the samples showed no saturation at the powers that could be achieved by the instrument (up to 25 mW). CW power saturation is a promising technique for attaining more information from spin labels than just mobility. Future experiments may achieve higher microwave powers by using even smaller sample tubes than the 2 mm outer diameter tubes used here to mitigate the effects of using lossy liquid samples. This may allow saturation, and enable the calculation of T_1 which gives direct information about the spin label environment.

4.3.5.2 Mass Spectrometry Experiments

A more direct way to determine the position of the spin label is through peptide analysis by Mass Spectrometry (MS). In such an experiment, labeled samples are digested by endoproteinases which cleave the enzyme at sequence specific locations, forming peptide

fragments of known (sequence predicted) mass. Labeled peptides can then be identified by the additional mass of the attached label.[138] This technique has been used in similar experiments, such as a study that detected the 151 Da shift caused by 2-hydroxy-5-nitrobenzyl bromide modifications on model peptides.[148]

Traditionally, two forms of MS have been applied to proteins; electrospray ionization (ESI) and Matrix Assisted Laser Desorption/Ionization (MALDI). Of these two ionization techniques, MALDI is more sensitive, capable of 0.01% mass accuracy and a range of 100 Da to 1 MDa under ideal conditions.[149] In MALDI-MS, the condensed phase sample is deposited on a matrix, such as α-cyano-4-hydroxy-cinamic acid (CHCA) or sinapinic acid. This matrix, which usually contains a chromophore, absorbs a UV laser pulse (chosen to minimize damage to the sample) and the excited matrix acid (or base) goes into the gas phase where it can ionize the sample molecules through collision and the transfer of a proton. These ions are then directed to the analyzer where mass/charge ratios (*m/z*) are detected. Time-of-flight (TOF) spectrometers are the most commonly used, and can detect concentrations as low as 10^{-18} M.[149] Though MALDI ionization is very mild, causing little fragmentation, it does produce some higher energy ions which may decompose.[149]

For the detection of spin labeling sites of EAL-5CF mutants, wild-type, EAL-5CF, αD207C, αN62C and αM434C were analyzed by MS with and without labeling. For labeling, samples were treated with 5 mM TCEP on ice for 1 hour to reduce exposed sulfhydryls. The samples were then desalted with 7 kDa exclusion Zeba Spin Desalting columns (Thermo Scientific) to remove the reducing agent, and then labeled on ice with 1:1 4MT in the presence of 2:1 AdoCbl. The subunits were then separated using SDS-

PAGE, to reduce the complexity of the MS sample. The individual subunits were then excised from the gel and trypsinized using an In-Gel Trypsin Digestion Kit (Thermo Scientific) which cleaves at carboxyl side of lysine and arginine. Alternatively, to generate different peptide fragments, the Asp-N protease, which cleaves at the peptide bond at the N-terminal end of aspartic acid, was also used. The peptide fragments were then extracted from the gel using QIAquick Gel Extraction kit (QIAGEN). The samples were further desalted with C_{18} tips (Thermo Scientific) as salts are known to negatively impact MALDI-MS. A saturated solution of CHCA dissolved in 50% acetonitrile and 0.1% trifluoroacetic acid was used as the matrix for MALDI. MS was done on an Applied Biosystems 4700 Proteomics Analyzer MALDI-TOF spectrometer in reflector positive ion mode, taking the average of at least 400 laser shots. The spectrometer was calibrated using the peptides N-acetyl-RENIN substrate tetradecapeptide (1801.1 Da) and substance P (1347.63 Da). The data was analyzed using the MoverZ software package, (ProteoMetrics, NY) in conjunction with the Protein Analysis Worksheet (PAWS) program (Dr. Ronald Beavis).

Using the theoretical digestion generated by the PAWS program, peptide fragments in the samples were identified. No cysteine containing peptide fragments were found with a mass modification consistent with 4MT labeling (+ 251.3 Da). It is possible that the 4MT nitroxide label is reactive under the conditions used in MS, or that it is otherwise unstable. Another maleimide which contains an inert ethyl group in place of the nitroxide, n-ethyl-maleimide (NEM), was used for further experiments. In a further complication, cysteine containing peptides could not be identified consistently in the unlabeled control samples (in contrast to the other peptides which were reliably detected).

This may be because of the high reactivity of cysteines which can react with other compounds such as acrylamide during the sample preparation procedure.[150] Thus, to protect the cysteines, the samples were denatured and reacted with an excess of iodoacetamide (IAA) subsequent to 4MT labeling. This provides a known mass modification of +57.0 Da to cysteine containing peptide fragments, and protects the cysteines from further modification. With these changes, it was possible to detect cysteine peptides in the MS samples. However, NEM modifications (+124.1 Da) could not be found consistently at 1 equivalent maleimide. When an excess of NEM was used, labeled peptides could be identified. Figure 12 shows an example mass spectrum from EutB αN62C reacted with an excess of NEM, and Figure 13 shows the peak assignments for this result. There is extensive coverage of the EutB subunit (83% of possible coverage in a single run), including 4 of the 5 NEM labeled cysteines. A duplicate run detected the remaining NEM conjugated cysteine. One possible conclusion is that labeling at 1:1 yields MS peaks that are below the signal-to-noise, but when labeled in excess these peaks are capable of being detected. This suggests that the labeling site is not robust, and the label is spread across multiple sites in a single sample.

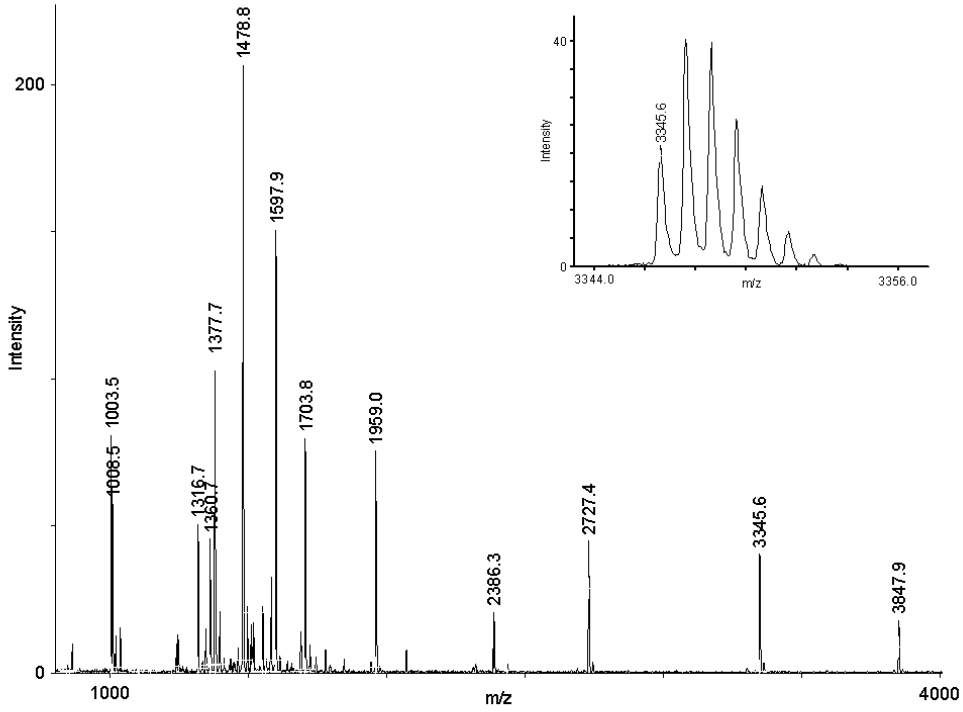
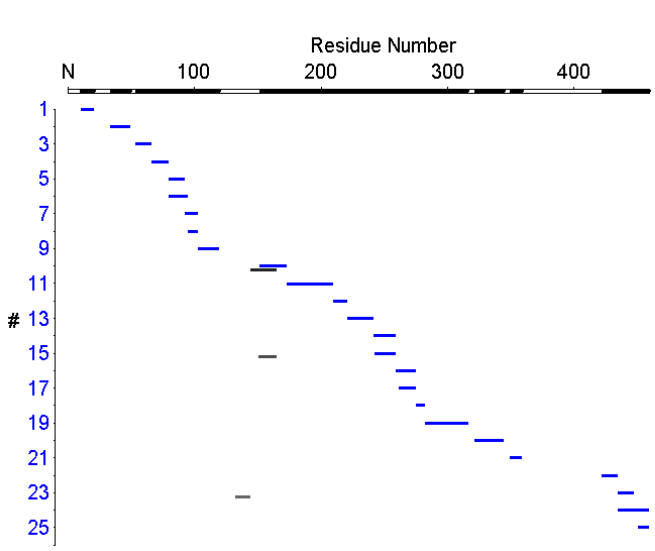


Figure 12. MALDI-MS Spectra of Trypsinized EutB subunit of N62C treated with an excess of N-Ethyl-Maleimide. Labeled monoisotopic masses of peaks with a signal to noise ratio (SNR) over 25 (for clarity). SNR of 3 used as the threshold for analysis. Inset: Zoomed in figure of single peak to show isotopic peak distribution and sensitivity.



#	Mass	Matching sequence (± error)	
1	1316.714	T[11-21]K (-0.04)	
2	1516.799	S[35-50]R (-0.05)	
3	1360.740	Q[55-66]R (-0.04)	
4	1703.839	N[67-80]R (-1.00)	
5	1549.791	L[81-93]R (-0.05)	
6	1790.948	L[81-95]K (-0.03)	
7	1244.738	I[94-103]R (-0.05)	
8	1003.540	N[96-103]R (-0.03)	
9	1958.970	E[104-120]R (-0.06)	
10	2386.282	A[153-173]R (-1.06)	K[146-166]R (-0.94)
11	3847.910	D[174-210]R (-0.04)	
12	1220.710	V[211-221]K (-0.05)	
13	2435.315	F[222-242]R (-1.00)	
14	1843.979	R[243-259]K (-1.00)	
15	1687.876	G[244-259]K (-1.00)	K[152-166]R (+1.03)
16	1776.963	G[260-275]R (-0.05)	
17	1478.764	E[263-275]R (-0.06)	
18	862.449	A[276-283]R (-0.02)	
19	3345.603	I[284-316]R (-0.05)	
20	2727.381	H[323-345]R (-0.05)	
21	1103.544	A[350-359]K (-0.04)	
22	1597.908	Q[423-435]R (-0.05)	
23	1377.697	W[436-447]R (-0.04)	I[134-145]K (-0.98)
24	2710.343	W[436-459]F (+0.03)	
25	1008.541	R[451-459]F (-0.04)	

Figure 13. Assigned Mass Fragments of the N-Ethyl-Maleimide labeled EutB subunit of N62C. The sequence coverage of the peptides identified in trypsinized EutB subunit of N62C labeled with an excess of NEM shown on left. The fragment from 360 to 422 contains no trypsin cleavage sites, and is too large (7339 kDa) to be observed by the selected experimental conditions. On right experimental masses and sequence assignments are shown. Secondary assignments are shown in gray.

In the MALDI-MS experiment, some peaks may remain unidentified, even after accounting for common modifications and missed cleavages by the proteinase used. Larger peptides can fragment at several points, preferentially along the backbone.[151] In addition, several groups may undergo less common modifications, leading to a change in the expected mass that would cause a fragment to be unassigned. Table 10 shows a partial list of peptide modifications that occur in MS samples. In an effort to account for these possibilities the labeling reagent was changed to 1:1 NEM / n-ethyl-d^5-maleimide (NEM-d^5). NEM-d^5 is a deuterium substituted NEM that is 5 Da heavier than standard

abundance NEM. Using a 50% NEM/NEM-d^5 label, the analysis could be approached from another direction. Instead of identifying peptides by masses directly, the results were analyzed for this +5 Da mass fingerprint. Two peaks separated by this mass would indicate that it was a labeled fragment. Control experiments without the labeling reagent could verify the legitimacy of the assignment. This fragment can then be identified by a comprehensive search of possible fragmentation patterns. This experimental approach has been used by Chait et al. to identify cysteine residues in peptides by alkylation with labeled and unlabeled acrylamide.[150] Unfortunately, at 1 equivalent NEM/NEM-d5, no such fingerprint spectra could be identified. A pair of peaks separated by 5 Da was detected, but this pair was predicted from the digestion of two non-cysteine containing peptides that fortuitously showed this difference in mass. Also, this pair was present in unlabeled controls, confirming its identification. This further reinforces the assessment that the labeling site is not robust, rather than the alternative explanation of labeled peptide fragmentation.

Table 10. Partial list of chemical modifications of Amino Acids during MALDI-MS.*

Modification	Amino Acid	Mass Difference (Da)
Water loss	S, T	-18.01
Ammonia loss	Q, K, R, N	-17.02
Urea Loss	c-terminal R	-60.03
Hydration	H, R	+18.01
Methylation		+14.02
Oxidation	M	+15.99
Acetylation	S	+42.01
Carbamidation	C	+57.02
Carboxylation	C	+58.00
Phosphorylation	T, S, Y	+79.97

* Selected items from http://www.weddslist.com/ms/maldi.html

These MS experiments indicate that the background cysteine reactivity in EAL-5CF mutants is significant. Thus, when labeling with a limiting concentration of 4MT, the labels are spread across multiple sites. Thus, despite the lack of solvent accessible cysteines, EAL-5CF is not an appropriate platform as a reactive-cysteine free EAL. Proteins are dynamic, and static structural information does not fully describe the accessibility of the labeling sites. Further work is needed to eliminate other reactive cysteines present in EAL.

4.3.6 Development of the EAL-11CF Construct

4.3.6.1 Identification of Mutagenesis Sites

To eliminate residual cysteine reactivity in EAL-5CF, more mutagenesis was needed. The structure of EAL was further examined to identify cysteines amenable to mutagenesis, taking into consideration steric restraints, hydrogen bonding and van der Waals interactions. Special attention was paid to residues within contact distance (< 4 Å) of the sulfhydryl moiety. In addition, the Pfam[95] EAL sequence library described in Chapter Two was used to assess the degree of conservation at each cysteine position. The residue for cysteine mutation was selected based on these considerations. Cysteine residues that were near to the active site were left unmodified. Table 11 shows information about the local interactions of the cysteine residues, and the 11 mutations selected for this construct, EAL-11CF.

Table 11. Residues selected for mutation in EAL-11CF.

Residue	Notes	Target
αC71	-SH contacts V72, E252, L262	A
αC129	-SH contacts R97, D133, V125	V
αC159	Consensus residue A (72.4 %)	A
αC223	-SH contacts V224, L225, L243, F245	V
αC249	-SH contacts G250, G254, S292, A293	S
αC283	-SH contacts G272, L284	A
αC361	Near active site	--
αC363	Near active site	--
αC364	Near active site	--
αC388	Near active site	--
βC37	Not in structure	A
βC46	Not in structure	A
βC82	-SH contacts L67, L70, V78	A
βC158	-SH contacts R124, S125, A147	A
βC243	Consensus residue L (47.8 %)	L
βC260	Near active site	--

4.3.6.2 Production and Expression of EAL-11CF

Due to the large number of mutations required for the EAL-11CF gene, it was directly synthesized by GeneArt, Invitrogen Life Technologies. The final sequence used for the construct is given in the attached appendix. A *BamHI* cutting site (GGATCC) was placed before the *eutB* start codon, and a *HindIII* site (AAGCTT) was placed after *eutC* to allow for manipulation of the EAL-11CF gene. The codon selection of the sequence was optimized by Invitrogen for bacterial growth. The final sequence did not contain any duplicate *BamHI* or *HindIII* sites, and was confirmed to yield the desired amino acid sequence. The EAL-11CF gene was subcloned into the Invitrogen pRSET A vector. pRSET A is a 2.9 kbp vector with a multiple cloning site which contains *BamHI* and *HindIII* sites in that order, allowing the sense insertion of the synthesized EAL-11CF gene. This vector also contains the T7 promoter for IPTG induced overexpression, an N-

terminal polyhistidine tag sequence, and an enterokinase linker to allow for tag removal. The vector also confers ampicillin resistance. Figure 14 shows a diagram of the arrangement of the final EAL-11CF plasmid.

Plasmid DNA Description:

The synthetic gene EAL-11CF-pRSETA was assembled from synthetic oligonucleotides and/or PCR products. The fragment was cloned into pRSET_A_A185 using BamHI and HindIII cloning sites. The plasmid DNA was purified from transformed bacteria and concentration determined by UV spectroscopy. The final construct was verified by sequencing. The sequence congruence within the used restriction sites was 100%. See the accompanying data sheets for sequences and find the original ABI trace files as well as the assembled sequences electronically on disk.
5 µg of the plasmid preparation were lyophilized for shipping.

Plasmid Map:

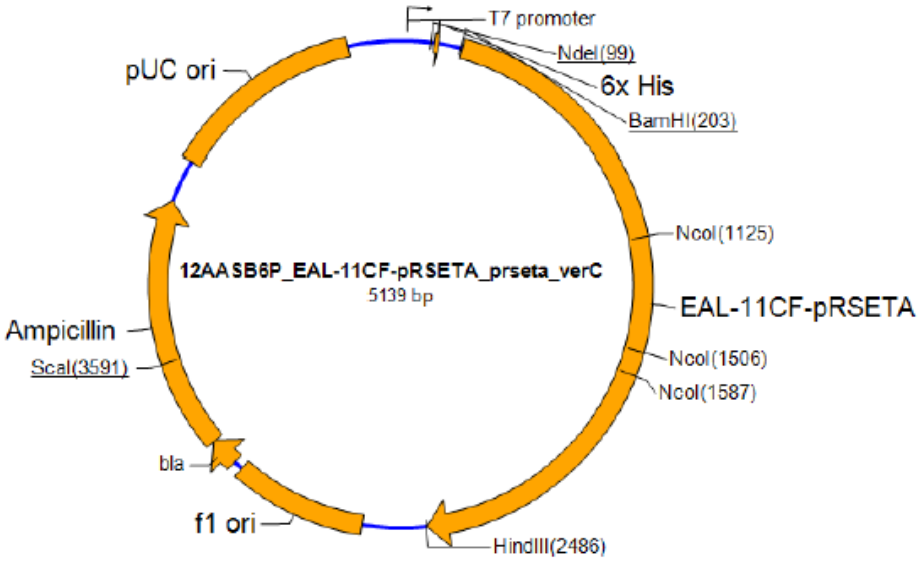


Figure 14. Description and Diagram of the EAL-11CF plasmid in pRSET A. Description of the preparation of EAL-11CF plasmid obtained from Invitrogen Life Technologies shown above. The arrangement of the plasmid, including the location of cloning sites, is shown below.

The EAL-11CF plasmid, (EAL-11CF gene subcloned into pRSET A), was transformed into BL21 (DE3) competent cells for protein expression, and XL 10 Gold cells for

cloning. *E. coli* BL21 (DE3), and its derivative strains, have been the most commonly used for protein expression for the last 15 years.[103] Sequence confirmation was performed by Beckman Coulter Genomics using the EAL-11CF specific sequencing primers listed in Table 12. Expression of EAL-11CF was attempted at 37 °C with 400 mM IPTG induction. However, SDS-PAGE analysis shows that the resulting protein was insoluble and probably present in inclusion bodies.

Table 12. Sequencing Primers for EAL11CF construct.*

Primer Name	5' to 3' Primer Sequence
EAL11CFSP1	CAC TAT AGG GAG ACC ACA ACG
EAL11CFSP2	TCT GAC CAG CGA AGT TGT TG
EAL11CFSP3	CTG GCA CGT CAT TAT GAT CC
EAL11CFSP4	CCT GGA TTG GTG TGG AAA AT
EAL11CFSP5	CAC CGT TGA AGC AGA TCG TA

* Primers obtained from Integrated DNA Technologies

In *E. coli* strains, such as BL21 (DE3), insoluble proteins accumulate as dense bodies, called inclusion bodies.[152] The tendency to form these insoluble aggregates has been partially attributed to the fast synthesis of proteins under T7 induction.[153] Genes downstream of a T7 promoter are transcribed by T7 RNA polymerase (T7RNAP), which is 8 times faster than *E. coli* RNA polymerase.[154] If the rate of protein folding is not sufficiently fast compared to translation (protein synthesis), high concentrations of the unfolded nascent polypeptide chains will be present. These chains will expose hydrophobic residues that are usually buried in the folded enzyme. This may lead to aggregation and the formation of inclusion bodies.[152, 155] Inclusion body formation may be prevented by modifying the host strain, controlling growth conditions, modifying the promoter, using fusion proteins or including a secretion vector.[152]

In an effort to mitigate inclusion body formation by lowering the rate of protein expression, EAL-11CF was expressed at low temperature (30 °C) and with lowered IPTG concentration for induction (100 mM versus 400 mM). However, in all cases, the resulting protein remained insoluble.

In *E. coli* overexpression systems, such as BL21 (DE3) cells, T7RNAP activity is controlled by the *lac*UV5 promoter. The *lac* repressor, LacI, prevents the expression of the T7RNAP protein until IPTG is added. After the addition of IPTG, LacI is displaced, which triggers production of T7RNAP. T7RNAP recognizes the T7 promoter, and initiates the overexpression of the desired recombinant protein.[154] In such a system, there is residual production of T7RNAP before induction, which causes lower levels of recombinant protein production. In fact, this basal T7RNAP production, and resulting protein expression has been used to produce soluble EAL under non-inducing conditions.[41] To prevent basal T7RNAP production, the $LacI^q$ repressor may be used. $LacI^q$ is a mutant LacI with tighter binding than the wild type variety. Alternatively, the T7Lys strain compensates for residual T7RNAP activity by producing lysozyme, a protein that inhibits T7RNAP from mRNA production. Residual T7RNAP is bound to lysozyme and temporarily rendered inactive. When induction occurs, the high level of T7RNAP saturates the lysozyme, and protein expression runs its usual course. Lysozyme also degrades bacterial cell walls which causes cells to rupture, thus, a lysozyme mutant, LysY, (with a K128Y mutation) lacking this cell wall activity, is used in its stead.[103] In the Lemo21 system, a separate pLemo plasmid (with chloramphenicol resistance) produces LysY. The LysY inhibitor concentration is in turn controlled by a separate

rhaBAD promoter that can be induced at various levels by L-rhamnose.[154] Thus tight and tunable control of T7RNAP activity is possible in the Lemo21 system.

The EAL-11CF plasmid was transformed into the Lemo21 competent cells, using both ampicillin (100 μg/mL) and chloramphenicol (30 μg/mL) antibiotics as selection agents. The protein was expressed from this cell line at 30 °C and 37 °C at 6 different L-rhamnose concentrations up to 1500 μM. However, under all tested conditions the protein remained in inclusion bodies. Similarly, expression was attempted using New England Biolabs T7 Express I^q cells (a BL21 (DE3) derivative $LacI^q$ mutant line), and T7 Express $LysY/I^q$ (contains both $LacI^q$ and LysY). The expression result was unchanged, and the protein remained insoluble. Varying the time of induction did not improve the result.

4.3.6.3 Production and Expression of Separated Subunits of EAL-11CF

Failing to express EAL-11CF in the soluble fraction, the next course of action was to purify the insoluble protein and refold it into the native state. Refolding proteins, especially those as large as the 50 kDa EutB, remains a challenging enterprise.[152] However, several commonalities have arisen in successful refolding procedures. Denaturants have been shown to improve protein refolding by suppressing aggregation.[156] However, they have a narrow range of effectiveness as they may also destabilize the native structure. Temperature also tends to impact refolding yields; however, the effect is not a simple one. Lower temperatures may minimize aggregation, but higher temperatures are needed for folding.[157] In addition, temperature may have a variety of effects on the other components used for renaturation. For proteins with disulfide bonds, the redox environment also factors into refolding.[158] Addition of small

molecule additives, such as $_L$-arginine, is also useful for refolding.[159] Other conditions such as pH, the presence of divalent cations, the presence of EDTA (a metal ion scavenger), and ionic strength also impact refolding yields. Typically, refolding experiments are done at low protein concentration (10 – 100 μg/mL), but some procedures have been successful at high protein concentration.[156]

In order to attempt refolding on EAL-11CF, several lchanges need to made to the EAL-11CF plasmid. In the EAL-11CF construct, the histidine tag is only present on the EutB subunit. During native purification of wild-type EAL, EutC associates with EutB, and a judicious choice of imidazole concentration for elution leads to the purification of both subunits. Under denaturing conditions, this association does not occur, and a separate construct containing the histidine tags on the individual subunits had to be created.

PCR was conducted on the EAL-11CF plasmid to isolate the coding regions for the individual subunits of the gene, and to insert *NdeI* (CATATG) and *XhoI* (CTCGAG) cutting sites at the start and end of both *eutB* and *eutC*. The primers listed in Table 13 were used. The DNA from the *eutB* construct described in Chapter Two, containing the coding sequence for an N-terminal histidine tagged EutB between *NdeI* and *XhoI* cutting sites in an ampicillin resistant pET28a plasmid, was purified. The DNA resulting from the PCR reaction, and the purified pET28a plasmid were digested with *NdeI* and *XhoI* to produce complementary sticky-ended DNA. These digestions were then purified on a 1% agarose gel. The *eutB-11CF* and *eutC-11CF* genes were ligated separately with the pET28a plasmid and transformed into XL10 Gold competent cells. They were then transferred into T7 Express LysY/I^q competent cells. The sequence was verified by Beckman Coulter Genomics.

Table 13. Primers used for Development of the Separated Subunit EAL11CF construct

Primer Name	5' to 3' Primer Sequence
EutB/EAL11CF/pET28a/NdeI FP	CGC GGC AGC CAT ATG AAA CTG AAA ACC
EutB/EAL11CF/XhoI BP	GGT GTT ATC CCC CTC GAG TTA GAA AAA CAG
EutC/EAL11CF/NdeI FP	GGG GAT AAC CAT ATG GAC CAG AAA
EutC/EAL11CF/pET28a/XhoI	GTG GTG CTC GAG TTA ACG GGT CAT ATT AAT

* Primers obtained from Integrated DNA Technologies

Both EutB-11CF and EutC-11CF were expressed as insoluble inclusion bodies in T7 Express LysY/I^q. The insoluble pellet was washed with 50 mM Tris, 10 mM EDTA, 5 mM DTT, and 500 mM NaCl at pH 7.5, and then solubilized in either 6 M guanidinium chloride, or 8 M urea, using a buffer supplemented with 500 mM NaCl, 50 mM Tris, and 100 mM NaH_2PO_4 at pH 8.0. The solubilized protein was then applied to a 5 mL His Trap FF crude column (GE Healthcare) in the denatured state. Refolding was attempted for both EutB-11CF and EutC-11CF while bound to the column by slowly removing the denaturant over the course of 2 h. This was unsuccessful, as the protein precipitated on the column in both cases. Thus, in a separate purification, the protein was eluted in the denatured state with 300 mM imidazole after a 100 mM imidazole wash. It was also possible to release the column bound denatured protein by lowering the pH to 4.5, instead of imidazole addition.

Table 14. Refolding Buffer Composition.*

Buffer	Guanidine (M)	L-Arginine (M)	GSH (mM)	GSSH (mM)
1	0	0	2	0.2
2	0	0.44	2	0.4
3	0	0.88	1	1
4	0.55	0	2	0.4
5	0.55	0.44	1	1
6	0.55	0.88	2	0.2
7	1.1	0	1	1
8	1.1	0.44	2	0.2
9	1.1	0.88	2	0.4

* Base Buffer: 55 mM Tris, 21 mM NaCl, 0.88 mM KCl, pH 8.2. Base buffer is supplemented with ingredients listed in table.

The 20 mL elution from affinity purification was concentrated using a 9 kDa Pierce Protein Concentrator (Thermo Scientific). The Pierce Protein Refolding Kit (Thermo Scientific) was used as a guide in refolding. Starting with a 1.1 mg/mL concentration of protein, 50 µL was slowly pipetted into 1 mL of the buffers listed in Table 14 in 10 µL aliquots. This concentration is an estimate, as during the concentration assay, there was some aggregation. The samples were left overnight at 4 °C to refold. For EutB-11CF, in buffer 1 there was visible aggregation, suggesting that refolding was not successful. Buffers 2 to 9 showed no visible aggregation. The protein concentration in the soluble fraction of each mixture was assayed and the results are shown in Figure 15. Both guanidine and arginine improve the solubility of the protein as expected. The activity of the refolded protein was tested by adding wild-type EutC to the refolded protein. However, the resulting mixture had no detectable activity. Adding native EutB to the mixture recovered activity which suggests that the solubilized protein is either incorrectly

folded or inactive. It should be noted that activity was greatly reduced in buffers 7, 8, and 9. This may be because high concentrations of guanidine impair hexamer formation.

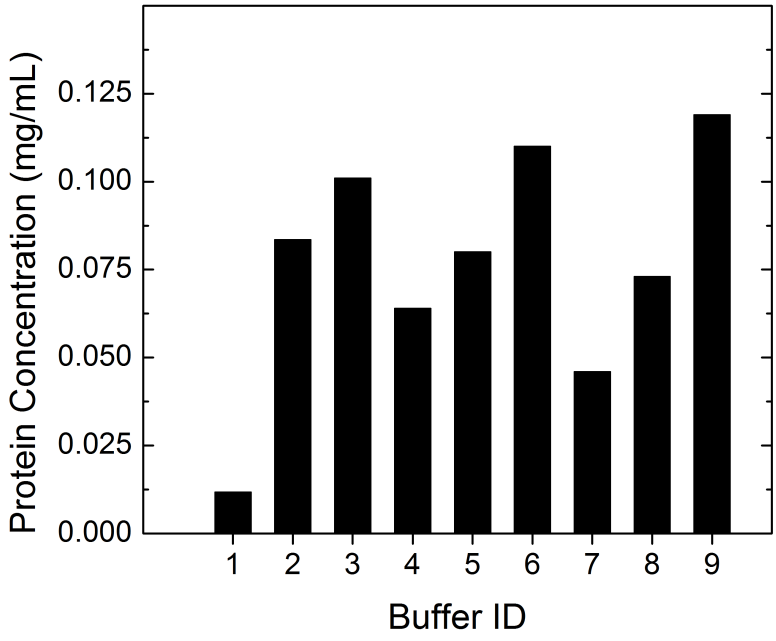


Figure 15. Concentration of Refolded EutB-11CF in Various Refolding Buffers. Concentration of EutB-11CF determined by Bradford protein assay, with a BSA standard. Concentration determined from aliquot of soluble fraction. Buffer compositions given in Table 14.

4.4 Insights from the Development of Cysteine Free EAL

In conclusion, the technique of site directed mutagenesis gives another handle to probe enzyme structure and function. In conjunction with spin labeling, many novel experiments become possible. From the library of single mutant constructs described within this chapter, it was possible to identify the primary cysteine reaction site of wild-type EAL as βC37. Unfortunately, this residue is not present within the modeled structure of EAL, so the uncertainty in its position is large, as estimated based on different

assumptions of the secondary structure of the unmodeled segment. Using the EAL single mutant library as a guide, it was possible to construct EAL-5CF which lacked solvent accessible cysteines. This construct retained wild-type-like parameters. However, it displayed background cysteine reactivity which could not be eliminated by engineered cysteines. This lead to the development of a construct in which 11 of the 16 native cysteine residues were removed, named EAL-11CF. This construct could not be expressed in soluble form under a range of experimental conditions. In addition, the individual subunits of EAL-11CF could not be refolded correctly (or were not active) in a variety of buffers. Directed labeling of EAL may not be possible by the traditional approach. Novel approaches are needed to ensure site selectivity in EAL labeling.

Chapter Five

Temperature-Dependence of Hydrogen Isotope Effects on the Coupled Radical Rearrangement-Hydrogen Transfer Reaction Sequence from 190 to 295 K

5

5.1 Introduction

EAL uses radical species to execute catalysis, and the unpaired electrons in these radical species illuminate the active site of the enzyme for the EPR technique. One of the unique features of the EAL system is that certain catalytic steps can be isolated and studied at cryogenic temperatures, where the enzyme remains functional.[45-47] The rate of formation of the Co(II)-radical pair state through the radical pair separation and the first hydrogen transfer steps have been studied at 234 to 248 K in a dimethyl sulfoxide-water based cryosolvent.[46] Further, experiments have been performed on EAL cryotrapped during steady-state turnover, in which the EAL reaction could be restarted from the Co(II)-substrate radical pair state by a temperature jump, allowing the observation of single turnover kinetics.[45, 47] In low temperature ($\leq$ 248 K) solutions, EAL enzymatic rates slow to a timescale suitable for full spectrum CW-EPR. Past low temperature studies have provided several insights into the mechanism of EAL catalysis. Several different kinetic regimes have been detected during the low temperature experiments on cryotrapped EAL in frozen aqueous solutions. Notably, at temperatures as low as 190 K, a reaction with an Arrhenius temperature dependence up to 217 K has been observed (Arrhenius prefactor: 10^{13} s^{-1}, activation energy: 15.0 kcal/mol), and identified as the native forward reaction of EAL through the radical rearrangement step.[45] Here, the low temperature reaction from 190 K has been connected with the reaction at ambient temperatures by a single model, which includes the radical rearrangement step, and the subsequent hydrogen transfer step. This model provides a deeper understanding of the mechanism of EAL catalysis, and the

degree of rate limitation of each of these two kinetic steps at different temperatures has been further elucidated.

5.1.1 EPR on Cryotrapped EAL

During steady-state turnover of aminoethanol by EAL at ambient temperatures in an ensemble sample, the reaction is not synchronized, and each state along the catalytic cycle is populated by different enzyme molecules. When the temperature is lowered during cryotrapping by using 140 K isopentane, the solution first becomes frozen, preventing the entry of additional substrate molecules into the enzyme active site (and possibly the egress of products). Other internal steps of the reaction cycle continue at lower temperatures. Since the radical rearrangement step is rate limiting,[47] it is the first of the internal steps to effectively cease. This leads to a rise in the population of the preceding Co(II)-substrate radical pair state, as the only paramagnetic intermediate detected in cryotrapped EAL.[45, 53]

Since the enzyme accumulates the Co(II)-substrate radical pair state during the trapping procedure, with no other intermediate paramagnetic intermediates present at detectable levels, the reaction becomes effectively synchronized at this step. Annealing the sample at higher temperatures allows turnover of the substrate radical to diamagnetic product, through the radical rearrangement (RR) and second hydrogen transfer (HT2) steps. By observing the resultant decay of the EPR signal, the kinetics of the microscopic reaction steps subsequent to the substrate radical state can be directly assessed. Above 214 K, the disappearance of the substrate radical pair signal is monoexponential, whereas, below 214 K, it is biexponential. The monoexponential decay, and the fast phase of the

biexponential decay, follow a single Arrhenius dependence from 190 to 217 K, suggesting that this decay corresponds to the same mechanism in both temperature regimes.[45] In addition, a slower reaction, with increased activation energy, has also been observed at temperatures below 214 K.[45] The onset of this slow reaction has been ascribed to a protein dynamical contribution becoming thermally inaccessible over the 207 to 214 K temperature window.[45] The fast phase of the biexponential decay, and the higher temperature monoexponential decay are further probed with isotope labeled substrates, from 190 to 295 K, as described later in this chapter.

5.1.2 Isotope Effects in EAL Catalysis

Radical rearrangement has been identified as a step causing significant rate limitation in the EAL reaction.[53, 56] In this step, the amine on C2 of the substrate radical migrates to C1, with the concomitant migration of the radical originally on C1, to C2. The kinetics of this transfer have been probed using $^{14}N/^{15}N$ isotope effects in steady-state turnover measurements.[81] With ethanolamine, this nitrogen isotope effect is 1.0017, which is less than the predicted value in excess of 1.03, and suggests that other steps in the catalytic cycle mask the full isotope effect.[81] Thus RR (or minimally, the microscopic C-N bond breaking step of RR,) is not the single rate determining step of EAL catalysis.

Additionally, EAL exhibits a steady-state $^{1}H/^{2}H$ isotope effect of 7.5 on k_{cat}.[56] This suggests that at least one of the two hydrogen transfers in the EAL catalytic cycle is also significantly rate limiting. The hydrogen isotope effect on the individual hydrogen transfer reactions have been determined. The $^{1}H/^{3}H$ isotope effect is 4.7 for the first hydrogen transfer, compared to 160 for the second transfer.[57] The corresponding $^{1}H/^{2}H$

isotope effect for the second transfer was calculated to be 7.3.[57] Thus, it is the second hydrogen transfer step, in which a C5' methyl hydrogen is transferred to the product radical, that contributes to rate determination. At low temperatures, 190 to 207 K, the $^{1}H/^{2}H$ isotope effect lowers to 1.4 ± 0.1.[47] This isotope effect has been attributed to a secondary hydrogen isotope effect, that originates from the rehybridization at the C2 center during rearrangement.[47] Moreover, the second hydrogen transfer reaction is not rate determining in this low temperature range.

Here, we further examine the $^{1}H/^{2}H$ isotope effect on EAL catalysis over two orders of magnitude on the temperature scale. A single model that accounts for the temperature dependence of the kinetics and the isotope effect over this wide range is proposed, thus explaining the apparent anomalous isotope effect data previously collected.[57] Moreover, EAL catalysis is kinetically complex with both radical rearrangement and the second hydrogen transfer steps contributing to rate determination. The relative rates of each of these steps are affected differently by temperature, leading to the change in isotope effects across the temperature range examined. Furthermore, because of the large values of the $^{1}H/^{2}H$ steady-state isotope effect (7.5)[56] and the $^{1}H/^{3}H$ isotope effect on the second hydrogen transfer (160),[57] it is postulated that the hydrogen transfer step proceeds by a quantum mechanical tunneling mechanism.

5.1.3 Biological Hydrogen Transfer

5.1.3.1 Transition State Theory

Hydrogen transfer reactions are widespread in biology.[160] Traditionally, these transfers have been treated in the context of transition state theory (TST).[161] Equation 1 shows the central rate equation given by this formalism.

$$k = Ae^{-\Delta G^{\ddagger}/RT} \qquad (1)$$

The rate constant, k, is given by a prefactor, A, the free energy, $\Delta G^{\ddagger}$, the gas constant, R, and the absolute temperature, T. In this theory, the reaction proceeds over a free energy curve from a single reactant well, over an activation free energy maximum as shown in Figure 1. The difference in free energy between the reactant well and the transition state determines the activation free energy, $\Delta G^{\ddagger}$.

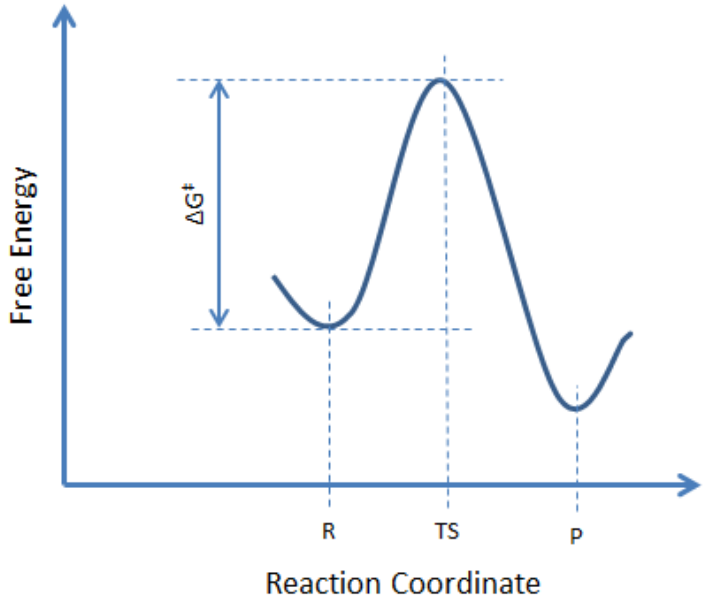


Figure 1. Free energy curve described by Transition State Theory. Reactant state denoted as R, transition state labeled TS, and product state labeled P. The activation free energy, $\Delta G^{\ddagger}$ is as shown.

In TST, catalysis can occur by two mechanisms; both of which lead to a reduction in the exponent, $\Delta G^{\ddagger}$.[162] In the first mechanism, ground state destabilization, the reactant well is destabilized, leading to an increase in its free energy and a reduction in $\Delta G^{\ddagger}$. In the other mechanism, transition state stabilization, the transition state energy is reduced by favorable interactions, again leading to a reduction in $\Delta G^{\ddagger}$. Figure 2 shows a graphical representation of the changes in the free energy curve.

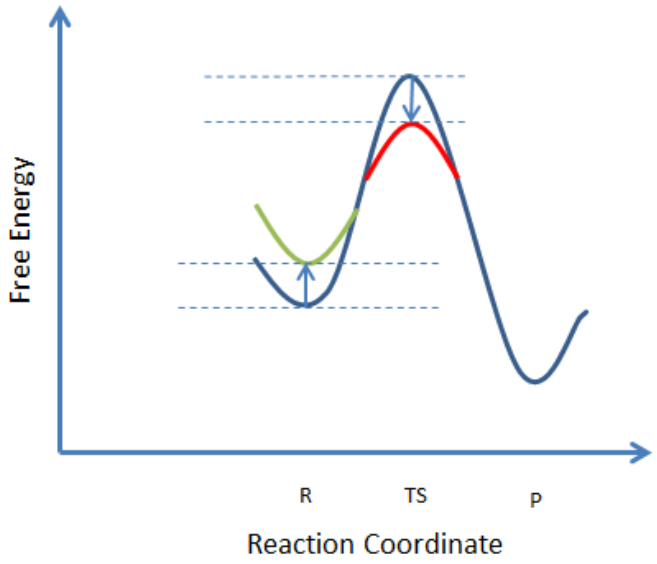


Figure 2. Mechanisms of Catalysis in Transition State Theory. Reactant state denoted as R, transition state labeled TS, and product state labeled P. The activation free energy, $\Delta G^{\ddagger}$ is reduced in ground state destabilization shown in green, and transition state stabilization shown in red.

Isotope effects seen in ^{1}H versus ^{2}H transfer can be partly explained by a ground state destabilization mechanism. In semiclassical hydrogen transfer theory, the Heisenberg uncertainty principle necessitates a zero point energy (ZPE) for hydrogen in the reactant

well. If that well is assumed to be harmonic, this energy is ½ħω which is proportional to $\sqrt{k_S/m}$ (ħ denotes Planck's constant divided by 2π, ω is the frequency of vibration within the well, k_S is the spring constant, and m is the reduced mass of the oscillator). The differences in the reduced mass of the three isotopes of hydrogen lead to ^{1}H having a higher ZPE than ^{2}H, which is, in turn, higher than ^{3}H. If the transition state free energy remains the same for all three isotopes, as is usually assumed, the change in ZPE yields a change in $\Delta G^{\ddagger}$ as shown in Figure 3, and thus there is an isotope effect on the hydrogen transfer rate.[163] The relative transfer rates due to this effect can be predicted, and these are given by the Swain-Schaad exponents.[164]

$$\frac{k_H}{k_T} = \left(\frac{k_D}{k_T}\right)^{3.26} \qquad (2)$$

$$\frac{k_H}{k_T} = \left(\frac{k_H}{k_D}\right)^{1.44} \qquad (3)$$

Equations 2 and 3 show the relative rates for ^{1}H (k_H), ^{2}H (k_D) and ^{3}H (k_T) transfer predicted by the semiclassical hydrogen transfer theory. Swain-Schaad exponents are as given in the equations.

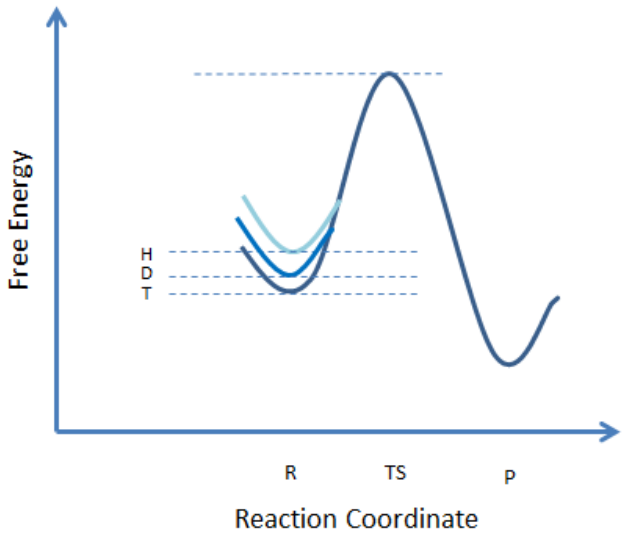


Figure 3. Origin of Hydrogen Isotope Effect in Semiclassical Hydrogen Transfer Theory. Reactant state denoted as R, transition state labeled TS, and product state labeled P. The activation free energy, $\Delta G^{\ddagger}$ is reduced by the change in zero point energy of ^{1}H (H), ^{2}H (D) and ^{3}H (T).

5.1.3.2 Tunneling Correction to Transition State Theory

It is possible for hydrogen transfer reactions to have isotope effects that deviate from those predicted by the Swain-Schaad relations and transition state theory. Notably, if the ^{1}H/^{2}H isotope effect exceeds 7, or the prefactor ratio, A_H/A_D is different from 1, this is indicative of hydrogen tunneling.[165] Typically, hydrogen tunneling occurs at short distances, less than 1 Å, from a hydrogen donor, which includes both the transferring hydrogen and the heavy atom to which it is bonded (typically a C, or N), to an acceptor (another heavy atom). The short separation between the donor and acceptor allows for tunneling through the activation barrier, rather than the thermally driven reaction over the barrier.[160] This is somewhat similar, in effect, to transition state stabilization. However,

the mechanism of stabilization is markedly different from the classical view of favorable electrostatic interactions in the transition state.

The de Broglie wavelength, which can be thought of as the extent of the hydrogen wavefunction, is given by $h/\sqrt{2mE}$. The energy, E, can be estimated by a typical value such as 5 kcal/mol.[163] This yields a wavelength of 0.63 Å for ^{1}H, and 0.45 Å for ^{2}H. The difference in wavelengths of the different isotopes of hydrogen means that they are able to tunnel at different donor-acceptor separations, with ^{1}H tunneling through wider activation barriers, whereas ^{2}H can only tunnel through sufficiently thin barriers.

In conjunction with the variation in tunneling distances of hydrogen isotopes, there is also a temperature effect due to the shape of the activation barrier. The activation energy barrier is wider at the base, and gets progressively less wide as the transition state is approached, where it has zero width. As temperature increases, the Boltzmann population of states of hydrogen in reactant wells shifts to a higher average position. ^{1}H can tunnel at lower temperatures because of its longer de Broglie wavelength. At higher temperatures, ^{2}H tunnels because of the reduced barrier width, with ^{3}H tunneling at even higher temperatures. This yields a temperature dependent tunneling effect, with the onset of tunneling occurring at higher temperatures for ^{1}H, ^{2}H and ^{3}H.[166] Figure 4 illustrates the variation in tunneling distances on a free energy curve. This is manifested as concave up deflection in a log(k) versus $1/T$ plot.[167] The TST prefactor, A, is determined from the slope of such a plot. Thus, depending on the degree of tunneling in the temperature region where A is determined, prefactor ratios lower than unity can be attained. In addition, isotope effects greater than those predicted by the Swain-Schaad relations are possible, with a maximum ^{1}H/^{2}H isotope effect of up to 12.[165] The exact degree of

tunneling depends on the specifics of the shape of the activation energy barrier.[168] The Bell correction, which assumes a parabolic barrier, is the most commonly used correction to account for tunneling in the TST formalism.[169]

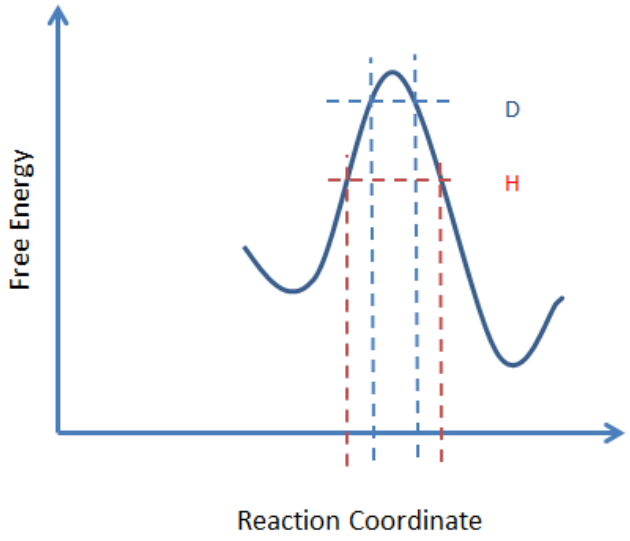


Figure 4. Variation in Tunneling Position in Bell Corrected Transition State Theory. Position of ^{1}H (H) tunneling and ^{2}H (D) tunneling shown in red and blue respectively. ^{1}H can tunnel longer distances which causes tunneling to occur at lower energy.

5.1.3.3 Full Tunneling Model

The tunneling correction models described above are insufficient to account for a variety of scenarios. They cannot adequately describe reactions that occur by "deep tunneling" where ≥99% of the reaction proceeds by a tunneling mechanism. Also, temperature dependent isotope effects, and prefactor ratios in excess of 1 also cannot be explained. To account for these phenomena, a full tunneling model is required.

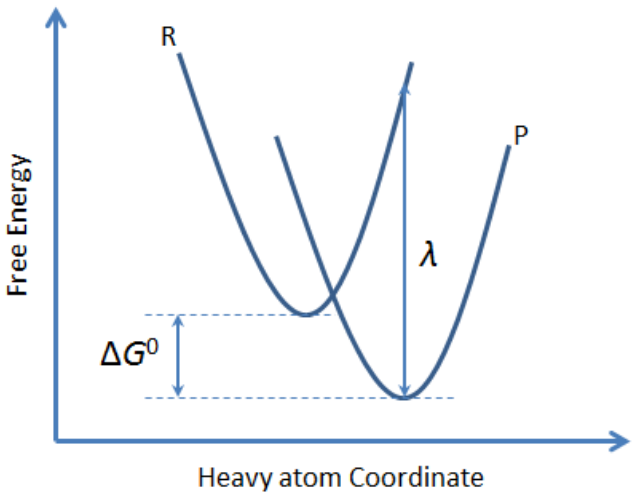


Figure 5. Free energy curve described in Marcus Theory of Electron Transfer. Reactant well denoted as R, product well labeled P. Reaction driving force, ΔG^0 and reorganization energy, λ are as shown. Transition state is at the intersection of reactant and product wells.

Full tunneling models, such as the environmentally coupled hydrogen tunneling model, are based on analogy with the Marcus theory of electron transfer.[170] Figure 5 shows a free energy curve representing the multidimensional free energy surface described by Marcus theory. In Figure 5, all heavy atom coordinates have been compressed to one dimension, along the *x*-axis. The system (which includes both the electron donor and acceptor atoms) initially oscillates in the reactant well. When the energy of the reactant state is matched to the energy of the product state (at the transition state), electron transfer is possible by conservation of energy and the Franck-Condon principle. The rate of electron transfer, k_{ET}, is also given by the coupling of electronic states, H_{AB}. A large value implies adiabatic, strong coupling in which the reactant transitions to product with a high probability when the reactant and product energies are matched. Conversely, a small

coupling is characteristic of the non-adiabatic case, where transitions are less probable. H_{AB} can be separated into a phenomenological exponential component with a decay constant β that depends on the tunneling properties of the medium separating the donor and acceptor over the distance *r*. The probability that the nuclear configuration will occupy this tunneling competent state depends on temperature, *T*, and the activation free energy, $\Delta G^{\ddagger}$, which can be separated into ΔG^{0}, the reaction driving force, and the reorganization energy, λ. λ/4 is the energy required to change bond lengths, angles, polarization and other nuclear configuration parameters. It is possible for the enzyme to reduce λ by the placement of catalytic residues. The central equation of Marcus theory is given below.

$$k_{ET} = \frac{H_{AB}^2}{\hbar}\sqrt{\frac{\pi}{\lambda RT}}e^{-\Delta G^{\ddagger}/RT} = \frac{\left(H_{AB}^0 e^{\beta(r_{AB}-r_0)}\right)^2}{\hbar}\sqrt{\frac{\pi}{\lambda RT}}exp\{-(\Delta G^{\circ}+\lambda)^2/(4\lambda RT)\} \quad (4)$$

Hydrogen tunneling can be formally treated using the Marcus model and theory of electron tunneling, but there are several differences. One of the most fundamental differences is the existence of different isotopes of hydrogen, which have the same chemical properties, but different masses. This allows for direct experimental probe of tunneling characteristics.[160] Also, hydrogen possesses a very short de Broglie wavelength relative to the electron. Thus, the range of tunneling is reduced, and there is much stronger distance dependence. β has been estimated to be around 25 $Å^{-1}$ for hydrogen transfer, compared to ~1 $Å^{-1}$ in electron transfer reactions.[171] This yields a static model for electron transfer, in which the nuclear vibrational motions do not influence the electron transfer probability, but allows for a more dynamic model for hydrogen transfer reactions in which even small changes in distance caused by nuclear vibrational motions

affect the probability of hydrogen transfer. The environmentally coupled hydrogen tunneling model takes this into account as it builds on the Marcus equation, as shown in Equation 5.[171]

$$k_{HT} = (Const.)exp\{-(\Delta G° + \lambda)^2/(4\lambda RT)\} \int_{e_0}^{e_1} \exp\{m\omega r^2/2\hbar\} \exp\left\{-\frac{E}{kT}\right\} de \quad (5)$$

In Equation 5, there is a constant term which is dependent on the number of reactive complexes, and the electronic overlap of the reactant and product. The first exponential term is the Marcus-like term which is dependent on the heavy atom coordinates. It determines the probability that the system is in the tunneling ready state (TRS). The second exponential term is analogous to the electronic coupling term. It describes the overlap of the wavefunctions, and takes into account the increased de Broglie wavelength of ^{1}H versus ^{2}H. In this term, *m* and ω are the transferring particle's mass and angular frequency, and *r* is transferring distance. The final exponential term is a gating term, which allows for a temperature dependent isotope effect. The Marcus-like term is also temperature dependent, but it is isotope insensitive and cannot generate a temperature-dependent isotope effect. Physically, the gating term enables vibrations along the axis of hydrogen transfer to modulate the reaction velocity as the donor and acceptor sample the distances between them. *E* represents the potential energy along this coordinate. The limits of the integral represent the extent of the vibrations between donor and acceptor. Gating becomes a significant effect when the equilibrium distance between donor and acceptor exceeds the distance over which the hydrogen can tunnel. The idea of these

promoting vibrations is controversial, as it is not universally agreed that vibrations can modulate reactions differently from the expected equilibrium effects.[161, 172, 173]

5.1.3.4 Tunneling in EAL and Related Enzymes

EAL has an overall $^{1}H/^{2}H$ isotope effect of 7.5,[56] which is at the limit of what can be explained by the semiclassical model.[165] However, as in other B_{12}-dependent enzymes, individual steps have been shown to have isotope effects outside of the semiclassical range, indicative of tunneling. This has been seen in methylmalonyl-CoA mutase (MCM), where the Co-C bond cleavage is kinetically coupled to the first hydrogen transfer, and shows an isotope effect of ≥20,[174, 175] or in glutamate mutase where this kinetically coupled transfer has an isotope effect of 28.[176] Similarly, both diol dehydratase and EAL show anomalous ^{3}H isotope effects for HT2, indicative of a tunneling component.[57, 177] As biological hydrogen transfers are examined more closely, it may be that tunneling is a very general mechanism of hydrogen transfer. What is still under debate is the extent to which the enzyme can modify this quantum mechanical tunneling. Finke and colleagues show that although the hydrogen transfer in MCM proceeds by a tunneling mechanism, an equivalent amount of tunneling occurs in the absence of the enzyme.[178] Thus, the enzyme may take advantage of tunneling without doing anything to optimize the tunneling behavior.

In this work, HT2 of EAL is examined within a tunneling framework. The low $^{1}H/^{2}H$ isotope effect of 1.4[47] at low temperatures, where tunneling is expected to be the dominant reaction mechanism due to the muted thermal reaction, is explained. Furthermore, this work shows that EAL catalysis does proceed by tunneling, but this

behavior is masked by radical rearrangement at low temperatures (190 to 207 K), and is progressively revealed as the temperature is increased to 295 K.

5.2 Methods

The wild-type EAL enzyme used in this study was purified from the cloned *S. typhimurium* EAL sequence,[41] by the method of Faust and Babior,[14] with the enzyme dialyzed against a 100 mM HEPES (pH 7.5), 10 mM potassium chloride, 5 mM dithiothreitol, and 10% glycerol buffer.[50] Enzymatic activity was measured by using the coupled alcohol dehydrogenase/NADH assay.[2] The assay mixture was preincubated at the desired temperature for temperature-dependent steady-state activity measurements.

Natural abundance aminoethanol (Aldrich Chemical Co.) and isotope labeled 1,1,2,2-$^{2}H_4$ aminoethanol (Cambridge Isotope Laboratories) were obtained commercially. Samples contained 10 – 15 mg/mL EAL (120 – 180 μM active sites), 2:1 AdoCbl, and an excess of substrate (~100 mM). Sample preparations were done on ice, in dim, red safe lighting. The cryotrapped samples were prepared by mixing the holoenzyme with the substrate, and transferring the resulting mixture to a 4 mm outer diameter EPR tube, which was then quickly inserted into ~140 K isopentane.[179] Up to 15 s elapsed from the reaction initiating mixing of holoenzyme and substrate to the insertion into isopentane.

EPR spectra were collected using a Bruker E500 ElexSys EPR spectrometer, equipped with a Bruker ER4123 SHQE cavity and a Bruker ER4131VT cryostat. Nitrogen gas evaporated from a liquid nitrogen dewar using an ER4131VT-1011 evaporator/transfer line, and a ER4121VT-1013 was used for temperature control. This setup allowed for a stable temperature (±0.5 K) across the length of the cavity, extended run times (~ 10^5 s),

and rapid temperature jumps (<60 s) compared with the substrate radical decay time (10^2 to 10^5 s).

Time-resolved EPR spectra were obtained as described previously.[47] Starting from a staging temperature of 160 or 180 K, the temperature was set to the measurement temperatures (190 to 223 K). After the temperature stabilized at the set temperature, the spectrometer was tuned, and spectral acquisition was started. The zero time point refers to the start of spectral acquisition.

EPR signal amplitude, *S*, was determined from the peak-to-trough amplitude of the $g \approx 2.00$ derivative feature. The time (t) dependent decay, was normalized by the $t = 0$ s amplitude, and was fitted to an exponential function (Equation 6).

$$\frac{S(t)}{S(0)} = \sum_{i=1}^{N} S_i e^{-k_i t} \qquad (6)$$

In Equation 6, S_i is the normalized component amplitude, and k_i is the first-order rate constant.

5.3 Results

5.3.1 Aminoethanol-Generated Co(II)-Substrate Radical Pair

Figure 6 shows the cw-EPR spectrum of the Co(II)-substrate radical pair formed with $^2H_4$-aminoethanol at 217 K. The unpaired electron spins on Co(II) and C1 of the substrate radical gives rise to an EPR signal with prominent features at 2850 G and 3250 to 3450 G respectively. To increase the time resolution, only the substrate radical signal within a 300 G window centered at 3320 G was collected. The substrate radical feature is split into a doublet that is not completely resolved at X-band frequencies, and further broadened by

the unpaired Co(II) electron spin.[180] This inhomogeneous broadening arises from *J*-coupling and dipolar coupling between the unpaired electrons.[37, 38] These features can be accounted for by spectral simulations.[48, 181]

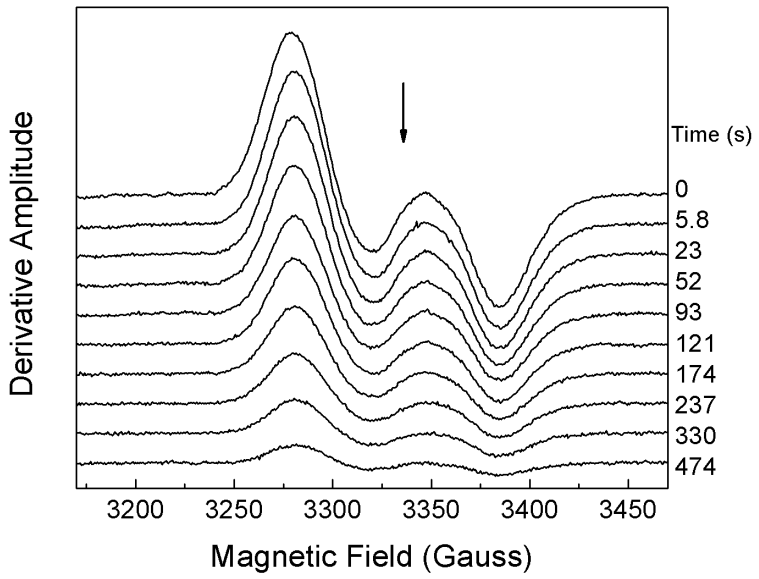


Figure 6. Time dependence of the EPR spectrum of the $^2H_4$-aminoethanol derived Co(II)-substrate radical pair in EAL at 217 K. *g* = 2.00 shown by arrow. Microwave frequency, 9.337 GHz, Power, 20.25 mW. Modulation amplitude, 1 G. Modulation frequency, 100 kHz. Conversion time, 10.24 ms. Time constant, 2.56 ms. Points, 512.

5.3.2 Time-Resolved CW-EPR of Substrate Radical Decay

Figure 6 shows the decay of the $^2H_4$-aminoethanol-generated CW-EPR signal at an annealing temperature of 217 K. Ten of the 90 spectra collected at this temperature are shown. The results show that the substrate radical decays without the rise of any other paramagnetic species. This is as reported previously for $^2H_4$-aminoethanol radical pair

decay below 207 K.[47] Radical pair decay data with $^{2}H_{4}$-aminoethanol substrate was collected at 210, 214, 217, 220 and 223K. Decay data with natural abundance aminoethanol was collected at 220, and 223K. With previously reported data,[45, 47] the temperature range for substrate radical decay data extends from 190 to 223 K with ^{1}H and ^{2}H substrates.

5.3.3 Time and Temperature Dependence of Substrate Radical Decay

Figure 7 shows the decay of the $^{2}H_{4}$-aminoethanol substrate radical signal at selected temperatures. All the data collected are well fit with an exponential function (Equation 6). For temperatures below 214 K, a biexponential function (Equation 6, $N = 2$) was necessary, whereas at and above 214 K, a monoexponential function ($N = 1$) was sufficient. This behavior is as reported previously for decay with natural abundance substrate.[45] The rate constants obtained from fitting the substrate radical decay for both ^{1}H and ^{2}H data sets are presented in Table 1.

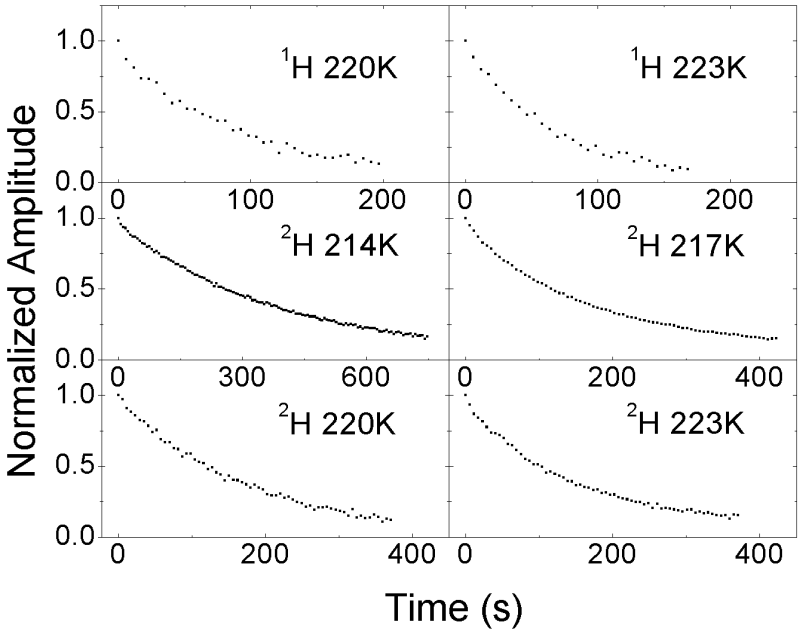


Figure 7. Decay of the amplitude of the substrate radical as a function of time at selected temperatures. Substrate and temperatures as indicated. EPR experimental conditions are as described in the legend of Figure 6.

Table 1. First-order rate constant, k_{obs}, for the fit of the exponential decays of the Co(II)-substrate radical pair, and steady-state turnover rates, k_{cat} (*), for ^{1}H and ^{2}H aminoethanol at different temperatures

Temperature (K)	$k^{‡}$		Isotope Effect
	^{1}H	^{2}H	(^{1}H / ^{2}H)
210	-	1.4 (± 0.4) x 10^{-3}	2.3 ± 0.8
214	-	2.3 (± 0.2) x 10^{-3}	1.7 ± 0.2
217	-	3.7 (± 0.3)x10^{-3}	1.8 ± 0.2
220	1.1 (± 0.1) x 10^{-2}	5.5(± 0.1) x 10^{-3}	2.0 ± 0.2
223	1.6 (± 0.1) x 10^{-2}	7.5 (± 0.4) x 10^{-3}	2.1 ± 0.2
277*	5.7 (± 0.4) x 10^{0}	1.0 (± 0.1) x 10^{0}	5.7 ± 0.7
295*	2.9 (± 2) x 10^{1}	3.9 (± 0.4) x 10^{0}	7.4 ± 0.9

‡ Data from multiple sources.[45, 47, 182]

5.3.4 Steady-State Turnover of Aminoethanol

Table 1 shows the rate of steady-state turnover (k_{cat}) of ^{1}H and ^{2}H aminoethanol at 277 and 295 K. These rates are obtained from optically following the change in absorbance of nicotinamide adenine dinucleotide upon oxidation in the aldehyde dehydrogenase coupled assay.[2] Also shown in Table 1 is the $^{1}H/^{2}H$ isotope effect (IE). At 295 K, the isotope effect is 7.4 ± 0.2 which is similar to values previously observed (7.4,[57] 7.5,[56] 8.[183])

5.3.5 Temperature Dependence of the Isotope Effect of Substrate Radical Decay

Figure 8 shows a plot of the natural logarithm of the observed first order rate constant for the decay of the Co(II)-substrate radical pair (k_{obs}) on an inverse temperature scale for both ^{1}H and ^{2}H substrates. In addition to the data collected here, the previously reported ^{1}H decay data below 217 K[45] and ^{2}H decay data below 210 K[47] are also included on this plot. For temperatures below 214 K where the decay is biexponential, only the fast phase is shown. For the ^{1}H data set, the k_{obs} values above and below 214 K transition follow the same linear relation. Previous fits from 190 to 207 K gave Arrhenius activation energy (E_a) of 13.9 ± 0.4 kcal/mol.[47] This is the same within error when data up to 223 K is included. (E_a = 13.4 ± 0.2 kcal/mol) This trend continues to higher temperatures, because the ^{1}H k_{cat} values do not deviate from this line. For the ^{2}H data set, the k_{obs} values follow a concave down trend, as temperature increases. As such, this data set does not follow the same Arrhenius relation for the whole temperature range. Below 207 K the ^{2}H data set follows a linear trend and yields an E_a of 13.9 ± 0.4 kcal/mol.

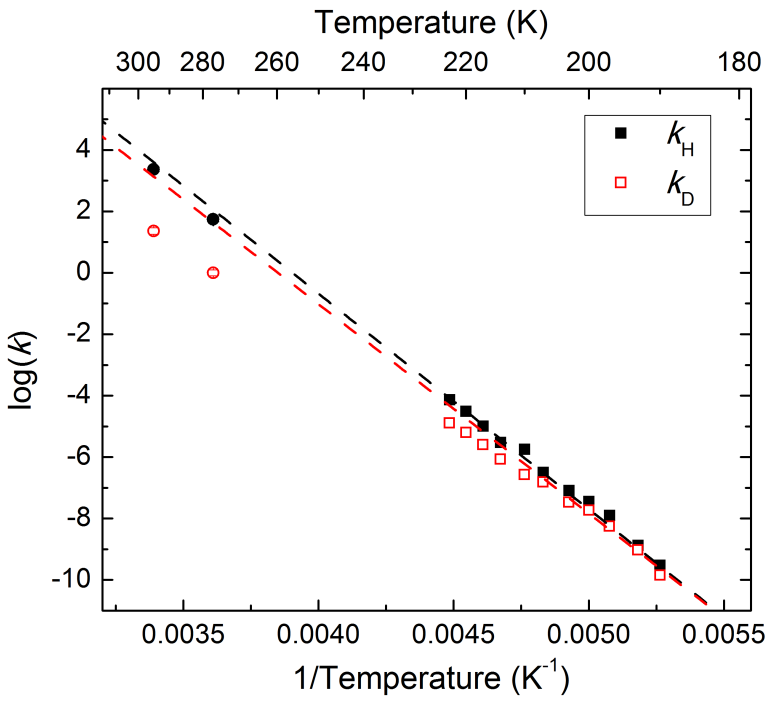


Figure 8. Arrhenius plot of the first-order rate constants of decay of the substrate radical pair state. EPR observed substrate radical Decay shown as squares. Steady-state k_{cat} values shown at 277 and 295 K as circles. At T < 214 K, only fast phase of substrate radical decay shown. ^{1}H shown in black, ^{2}H shown in red. Fit of low temperature (≤ 207 K) shown as dashed line, and extrapolated to higher temperatures.

Figure 9 shows the corresponding isotope effect (k_H/k_D) on an inverse temperature scale. The isotope effect values remain approximately constant at 1.4 ± 0.2 at low temperatures (≤ 210 K). As the temperature increases, the isotope effect also increases, attaining a value of 2.1 ± 0.2 at 223 K, and increasing to 7.4 ± 0.9 at 295 K.

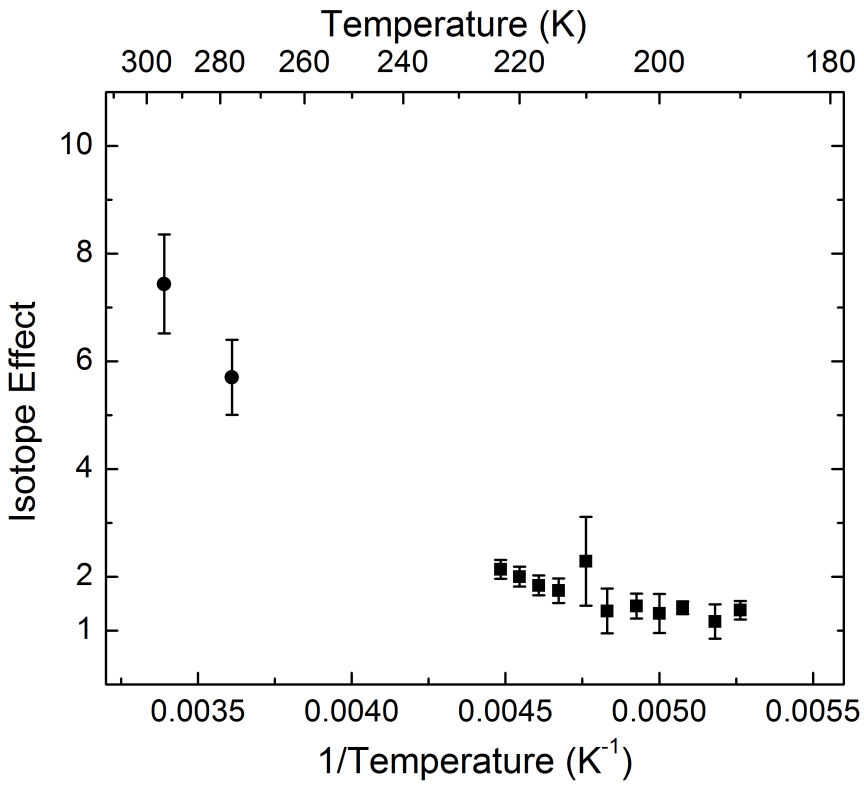


Figure 9. $^{1}H/^{2}H$ Isotope effect of substrate radical decay temperature dependence. Data at $T \leq 223$ K represents EPR observed substrate radical decay, and shown as squares. Steady-state kcat values shown as circles. At $T < 214$ K, fast phase of biexponential decay used for calculation of isotope effect.

5.4 Discussion

5.4.1 Equivalence of k_{cat} and k_{obs}

In contrast to the EPR-detected rates for substrate radical decay at lower temperatures (190 to 223K), steady-state turnover involves multiple passages through the entire catalytic cycle. However, if the rate constants for the microscopic steps preceding substrate radical formation, and the steps subsequent to the second hydrogen transfer are significantly larger, compared to the rate constants for the RR and HT2 (EPR-visible

states) steps, then the k_{cat} measurements collected in solution are equivalent to the cryotrapped k_{obs} values obtained from EPR. There are several indications that this is indeed the case, and thus, the steady-state turnover values report on the first-order decay rate constant for the radical pair state. Optical spectra, which are tuned to the disposition of the B_{12} Co-C bond, indicate that during steady-state turnover ~70% of the bound AdoCbl is in the cleaved Co(II) state.[53] Further, the rate of reformation of the Co-C bond subsequent to exhaustion of substrate is equivalent to k_{cat}, to within the standard deviation of the measurements.[56] Taken together, these results indicate that the rate determining step of the catalytic cycle at high temperature (277 and 295 K) is the radical rearrangement step, which is the step observed at low temperature (190 to 223 K). In addition, rapid-mixing, freeze-quench experiments show that the radical rearrangement step is rate limiting at high temperatures.[53] Also, the equivalence of the extrapolated ^{1}H Arrhenius relation from low temperature k_{obs} data to the 298 K k_{cat} value[45] further suggests that the same step is rate-determining from 190 to 295 K.

5.4.2 Comparison of ^{1}H Reaction Kinetics from 190 to 295 K

Figure 8 shows an Arrhenius plot of the first-order decay constant of the fast (190 to 210 K) and monoexponential (214 to 223 K) phases of the ^{1}H substrate radical decay (k_{obs}), and the steady-state turnover (k_{cat}, 277 and 295 K) of the ^{1}H reaction. The linear extrapolation of an Arrhenius fit to the 190 to 223 K data to ambient temperatures is shown as a dashed line. The 277 and 295 K data lie close to the low temperature extrapolation for the ^{1}H substrate. The agreement between k_{obs} and k_{cat} is consistent with previous conclusions, that the radical rearrangement step represents a significant barrier to the overall reaction in EAL at ambient temperature.[81] We previously concluded that

the radical rearrangement is the rate-determining step in the decay reaction at 190-207 K, for both the ^{1}H and ^{2}H substrates.[47] Figure 8 extends this conclusion for the ^{1}H substrate, by showing that the radical rearrangement step is the predominant rate-determining step for reaction of the ^{1}H-substrate radical from 190 to 295 K.

5.4.3 Comparison of ^{2}H Reaction Kinetics from 190 to 295 K

Figure 8 shows that extrapolation of the low temperature (190 to 207 K) Arrhenius relation for the ^{2}H-substrate (dashed line) leads to values for k_{obs} from 214 K to 223 K that deviate from the predicted relation. Moreover, the predicted k_{obs} values for 277 and 295 K lie far above the experimental k_{cat} values for ^{2}H substrate. The extrapolation also disagrees with the ^{1}H/^{2}H steady-state turnover isotope effect of 7.4-7.5, or 7.8[56, 57, 183] on k_{cat}, and the ^{1}H/^{2}H isotope effect of 7.3 on transfer of hydrogen from the 5'-deoxyadenosyl C5'-methyl group to the product radical.[57] Therefore, we propose that the dependence of k_{obs} on temperature for the ^{2}H substrate diverges at $T > 207$ K from the prediction of the low temperature Arrhenius relation. We test this proposal by using a minimal kinetic model, as described in the following sections.

5.4.4 Kinetic Model for Substrate Radical Reaction

The non-linear temperature dependence of the ^{2}H-substrate radical reaction kinetics (k_{obs}, k_{cat}) indicates the involvement of at least distinct two steps in the substrate radical decay. Further, the significant isotope effect at $T>210$ K indicates that at least one of these steps corresponds to a hydrogen atom transfer. This is consistent with the proposed minimal mechanism for the catalytic cycle (RR and HT2). Therefore, we develop a minimal three-state, two-step model to account for the temperature-dependent kinetics of the

cryotrapped substrate radical decay. The model is shown in Figure 10. In Figure 10, S•, P• and PH represent Co(II)-substrate radical pair state, Co(II)-product radical pair state and diamagnetic product state, respectively. The microscopic rate constants, k_{12}, k_{21} and k_{23}, are the rate constants for the rearrangement step, reverse of the rearrangement step, and the forward step of HT2, respectively. HT2 is treated as an irreversible step in this model of substrate radical decay.

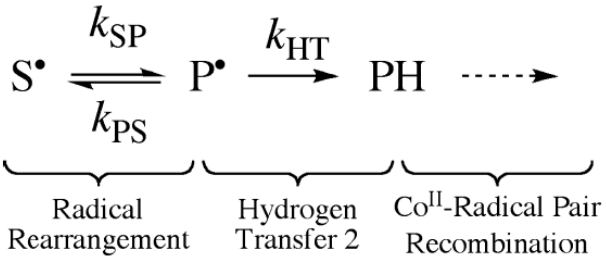


Figure 10. Model of decay of the cryotrapped substrate radical. States are $S^{\bullet}$, substrate radical, $P^{\bullet}$, product radical, PH diamagnetic product.

The observed first-order rate constant, k_{obs}, is the observed kinetic rate constant of the fast phase decay ($T < 214$ K) or the monoexponential decay ($T \geq 214$ K) of the substrate radical. From the model presented in Figure 10, k_{obs} and the microscopic first-order rate constants, k_{12}, k_{21}, and k_{23}, have the following relationship, under the steady-state approximation for the P• intermediate state:[184]

$$k_{obs} = \frac{k_{12}k_{23}}{k_{21}+k_{23}} = \frac{k_{12}}{1+{}^{k_{21}}/_{k_{23}}} \qquad (7)$$

The simulation of the temperature dependence of Figure 8 using Equation 7 requires the temperature dependence of the individual rate constants, k_{ij}, which is given by the

Arrhenius expression. Here, the Arrhenius expression is used to connect with previous treatments of substrate radical decay.[45, 47]

$$k_{ij} = A_{ij} e^{-E_{a,ij}/RT} \qquad (8)$$

In Equation 8, A_{ij} refers to the Arrhenius prefactor, which gives the value of k as the temperature, T tends to infinity. $E_{a,ij}$ refers to the activation energy. The temperature dependence of k_{12} is known, from the reported values for A_{12} and $E_{a,12}$ for the ^{1}H and ^{2}H substrates (collected in Table 2).[47] The remaining parameter in Equation 7 is given by the following relation.

$$\frac{k_{21}}{k_{23}} = \frac{A_{21}}{A_{23}} e^{(-E_{a,21}+E_{a,23})/RT} \qquad (9)$$

Table 2. Arrhenius parameters for the fit of the Temperature Dependent Decays of the Aminoethanol Derived Co(II) Substrate Radical Pair from 190 to 207 K.*

	log[A (s^{-1})]	**E_a (kcal.mol^{-1})**
^{1}H	11.9 ± 0.9	13.9 ± 0.4
^{2}H	11.4 ± 0.9	13.5 ± 0.4

* Data from Zhu and Warncke.[47]

Therefore, the substrate radical decay can be fit by a two-step, three-state model with two adjustable parameters; $a = \frac{A_{21}}{A_{23}}$, and $b = E_{a,21} - E_{a,23}$. The fitting equation is given as follows:

$$k_{obs} = k_{12} \frac{1}{1+ae^{-b/RT}} \qquad (10)$$

The ^{1}H and ^{2}H Co(II)-substrate radical pair data was simultaneously fit to Equation 10, by minimizing the 2-norm of the residuals. From this, values for the exponential prefactors, a_H and a_D, and the exponent, b were obtained. The values of b_H and b_D were assumed to be equivalent, for reasons that will be addressed later. The confidence interval of the fits was estimated using Monte Carlo simulations.[185] This was done as follows. After obtaining the best fit, the residuals were calculated. Random scatter was added to the best fit with a standard deviation equal to that of the residuals. This simulated data set was fit, and the new best fitting value was recorded. This procedure was repeated $\sim 10^4$ times, and the 2.5 and 97.5 percentile of the generated fits were used as the 95% confidence interval. Increasing the number of repetitions did not significantly change the confidence interval. The fit of the data is shown in Figure 11 and the resulting fit parameters are given in Table 3. The calculated isotope effect of the fit is shown in Figure 12.

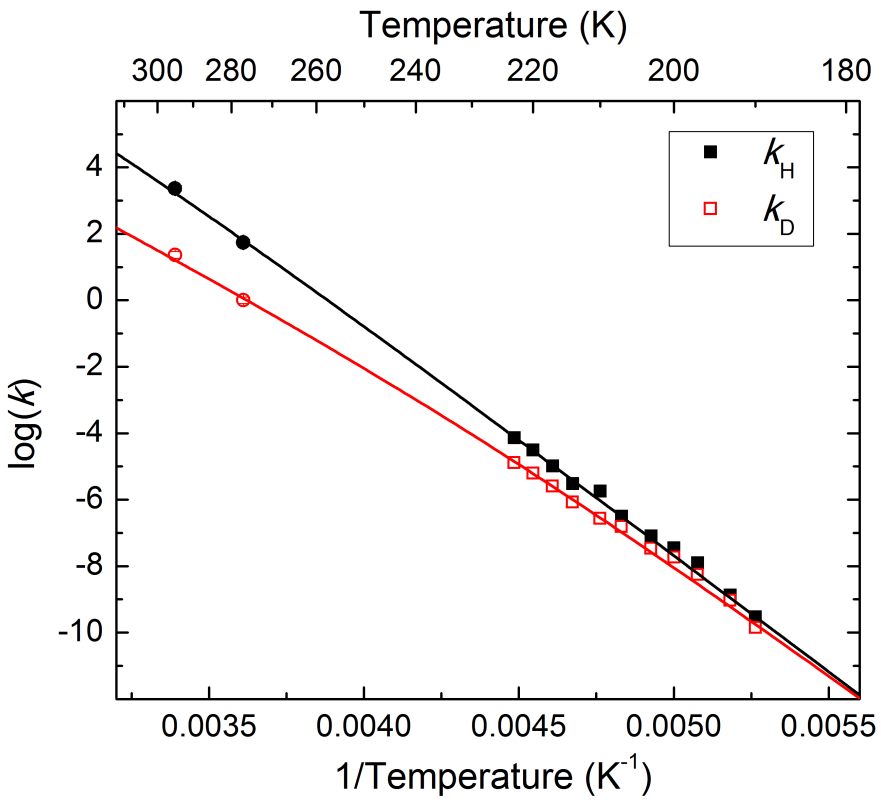


Figure 11. Fit of Arrhenius plot of the first-order rate constants of decay of the substrate radical pair state. EPR observed substrate radical Decay shown as squares. Steady-state k_{cat} values shown at 277 and 295 K as circles. At T < 214 K, only fast phase of substrate radical decay shown. ^{1}H shown in black, ^{2}H shown in red. Fit of low temperature (≤ 207 K) shown as dashed line, and extrapolated to higher temperatures. Fit to Equation 9 with parameters shown in Table 3.

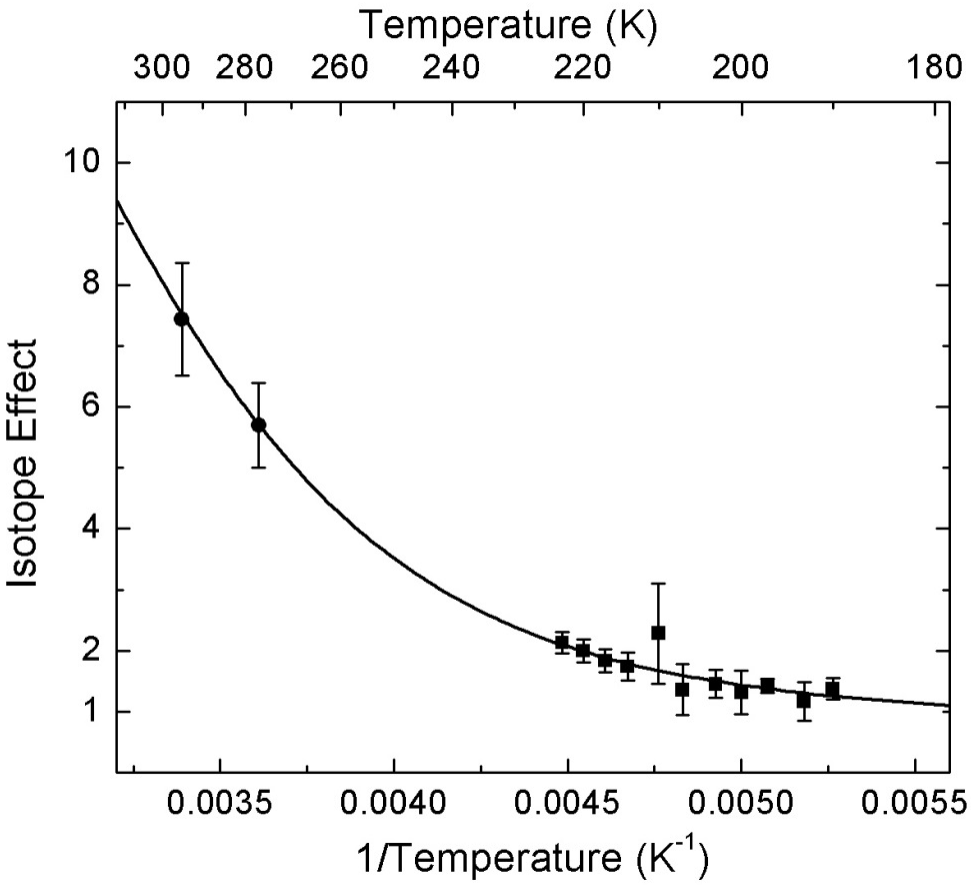


Figure 12. Fit of $^{1}H/^{2}H$ Isotope effect of substrate radical decay temperature dependence. Data at $T \leq 223$ K represents EPR observed substrate radical decay, and shown as squares. Steady-state kcat values shown as circles. At $T < 214$ K, fast phase of biexponential decay used for calculation of isotope effect. Fit to Equation 9 with parameters shown in Table 3.

Table 3. Parameters for fit of k_{cat} and k_{obs} for ^{1}H- and ^{2}H-substrate from 190 K to 295 K using Equation 10. †

	a	b* (kcal.mol^{-1})
^{1}H k_{PS}/k_{HT}	325 (± 200)	3.80 (± 0.26)
^{2}H k_{PS}/k_{HT}	3980 (± 2200)	3.80 (± 0.26)

† The value of k_{SP}, was determined from low temperature data (below 207 K) where radical rearrangement is rate limiting.[47] Fit parameters are given with 95% confidence bounds.
*The value of *b*, the difference in activation enthalpy ($\Delta H^{\ddagger}$) for P●→S● (reverse radical rearrangement) and P●→P (hydrogen transfer 2) steps, is assumed identical for different isotopes, and is simultaneously fit to a single value for ^{1}H and ^{2}H.

5.4.5 Temperature-Dependent Isotope Effect cannot be explained by Activation Entropy Differences

Another formalism for describing the individual microscopic steps partaking in substrate radical decay is the Eyring equation. This equation has the same form as the Arrhenius equation described in Equation 8. The explicit form of the Eyring equation is given in Equation 11.

$$k = \frac{k_B T}{h} e^{-\Delta G^{\ddagger}/RT} \qquad (11)$$

In contrast to the Arrhenius expression, the prefactor of the Eyring equation is temperature dependent, and the activation energy, $\Delta G^{\ddagger}$ is the free energy and can be separated into $\Delta H^{\ddagger} - T\Delta S^{\ddagger}$, where $\Delta H^{\ddagger}$ is the activation enthalpy, and $\Delta S^{\ddagger}$ is the

activation entropy. The use of the Eyring equation is required to fully account for the possibility that the temperature-dependent isotope effect is caused by a change in the entropic contribution as temperature is increased.

Using this formalism, the fitting equation becomes:

$$k_{obs} = k_{12}\frac{1}{1+\exp\left[(\Delta S^{\ddagger}_{21}-\Delta S^{\ddagger}_{23})/R\right]\exp\left[(\Delta H^{\ddagger}_{23}-\Delta H^{\ddagger}_{21})/RT\right]} \qquad (12)$$

In Equation 12, deviation from the low temperature k_{12} values are modulated by the difference in activation enthalpies and entropies. It is reasonable to assume that the change in enthalpies, $\Delta\Delta H^{\ddagger}$, should be unchanged for ^{1}H and ^{2}H substrate. When the system is in the product radical state, $\Delta\Delta H^{\ddagger}$ represents the difference between the forward (HT2) and reverse (RR) barrier heights. ^{2}H-labeled substrate destabilizes the product radical state, but should not significantly affect the transition state of either reverse radical rearrangement or hydrogen transfer 2. Thus, $\Delta\Delta H^{\ddagger}$ would be equivalent for both the ^{1}H and ^{2}H substrate. In the fit described in the previous section, this means that $b_{\mathrm{H}} = b_{\mathrm{D}}$. The deviation in the ^{2}H data above 210 K in Figure 8 may come from an entropic contribution which becomes significant at higher temperatures. This would be evident in the exponential prefactor, a. $\Delta\Delta S^{\ddagger}$ may be non-zero, as there may be differing entropic involvements between hydrogen transfer step 2 and reverse radical rearrangement. However, unless the hydrogen transfer step involves significant excited vibrational states which would be unusual,[168] the activation entropy difference is expected to be the same for ^{1}H and ^{2}H substrates. This means that if the prefactor terms are ratioed, $a_{\mathrm{H}}/a_{\mathrm{D}}$, the entropy difference terms should cancel. However, from the fit, the ratio $a_{\mathrm{H}}/a_{\mathrm{D}}$ is 0.082, and not unity as expected for the entropic hypothesis. An additional quantum mechanical

tunneling mechanism needs to be invoked to fully understand the change in the prefactor ratio.

5.4.6 Hydrogen Transfer 2 exhibits Quantum Mechanical Tunneling

As discussed before, in Equation 10, the fitting parameter, *a*, refers to the ratio of exponential prefactors, $\frac{A_{21}}{A_{23}}$ as described in Equation 9. Extending on this, the ratio of the ^{1}H and ^{2}H fitting parameters yields a ratio of the Eyring exponential prefactors as given by Equation 13.

$$\frac{a_H}{a_D} = \frac{A_{21,H}}{A_{21,D}} \frac{A_{23,D}}{A_{23,H}} \tag{13}$$

The reverse reaction out of the product radical state (reverse radical rearrangement) involves the transfer of an amine. The $^{1}H/^{2}H$ isotope effect is not expected to be significant in this step. Thus, the ratio $\frac{A_{21,H}}{A_{21,D}}$ is expected to be on the order of 1. However, the hydrogen transfer 2 step may involve a significant isotope effect. If we assign the prefactor ratio of 0.082 to this component, (the forward hydrogen transfer reaction), we obtain $\frac{A_{23,D}}{A_{23,H}} \sim 0.082$. In more familiar terms, the $^{1}H/^{2}H$ prefactor ratio is on the order of 12. Recall that prefactor ratios over 1.2 cannot be satisfactorily explained using a classical thermally activated mechanism.[165] The value of 12 attained here implies that this hydrogen transfer proceeds through tunneling. Figure 13 shows a diagram of the proposed reaction scheme from the product radical state.

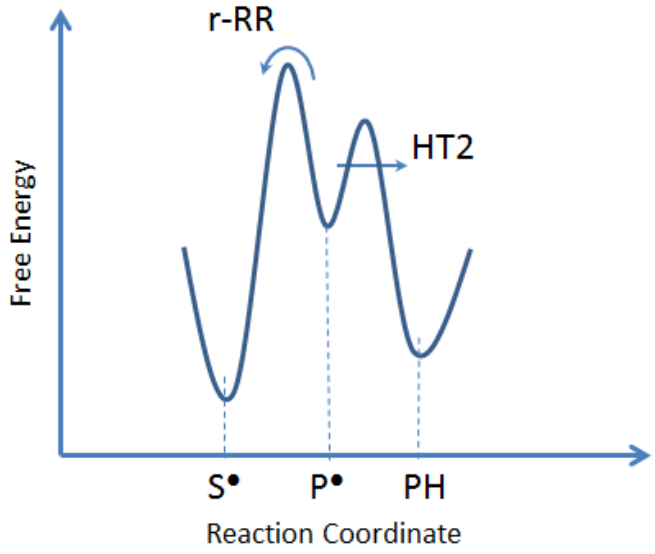


Figure 13. Free energy diagram of Co(II)-Substrate Radical Decay. Substrate radical, product radical and diamagnetic product denoted by $S^{\bullet}$, $P^{\bullet}$ and PH. Thermally activated reverse radical rearrangement (r-RR), and quantum mechanical tunneling hydrogen transfer 2 (HT2) as indicated.

The tunneling model described previously predicts that the prefactor ratio is dependent on the degree of overlap of the wavefunctions. Integrating this into the Eyring analysis of the prefactor ratio gives, to a first approximation, Equation 14.

$$\frac{a_H}{a_D} = \frac{e^{\Delta S_H^{\ddagger}/R} e^{m_H \omega_H r_H^2/2\hbar}}{e^{\Delta S_D^{\ddagger}/R} e^{m_D \omega_D r_D^2/2\hbar}} = e^{(m_H \omega_H r_H^2 - m_D \omega_D r_D^2)/2\hbar} \qquad (14)$$

In Equation 14, the relation $\frac{e^{\Delta S_H^{\ddagger}/R}}{e^{\Delta S_D^{\ddagger}/R}} = 1$ has been assumed. It is possible to extract the tunneling distance from Equation 14 if several assumptions are made. First, ω_H, the vibrational frequency along the axis of the hydrogen transfer is unknown, but a typical frequency of this vibration is 100 cm^{-1} for ^1H.[171] For ^2H, if we assume a harmonic

potential, $\omega = \sqrt{\frac{k}{m}}$, so this angular frequency becomes $\omega_D = \frac{1}{\sqrt{2}}\omega_H$. Also, if there is not significant gating which necessitates promoting vibrations to enable tunneling, the tunneling occurs over the same distance in both ^{1}H and ^{2}H samples ($r_H = r_D$). Last, if $\Delta\Delta S^{\ddagger}$ is approximately zero, as assumed and justified earlier, r_H can be calculated to be 2 Å. This is a reasonable value, within the range seen for tunneling reactions in the extensively characterized soybean lipoxygenase-1, which shows tunneling distances as long as 2.55 Å in a I553→G mutant.[171]

5.4.7 Estimation of the Activation Entropy Difference through the Product Radical State

The initial radical pair separation and hydrogen transfer 1 steps of EAL catalysis, which lead to the Co(II)-substrate radical pair state, have a significant entropic component. These steps are characterized by a large, positive entropy of 45 cal/mol/K, leading to an estimated -13.4 kcal/mol change in the activation free energy of the overall reaction at 298 K, with -11.3 kcal/mol of the entropy-derived free energy being attributed to the EAL protein for the (*S*)-2-aminopropanol substrate.[46] This compares with a -13 kcal/mol enthalpic contribution to the free energy by EAL.[46] It is unknown during which step the protein returns this positive entropy, as is required by the closed reaction cycle of EAL catalysis.

To determine whether the return from the high entropy plateau entered in the Co(II)-substrate radical pair state is effected in the radical rearrangement or hydrogen transfer 2 steps, the change in the activation entropy difference across these two steps must be examined. The activation entropy difference, $\Delta S_{21}^{\ddagger} - \Delta S_{23}^{\ddagger}$, is contained in the fitting

prefactor, a_H, mixed with the tunneling component of hydrogen transfer 2, as described above. With the determination of the tunneling component of hydrogen transfer 2, it is now possible to estimate the change in entropy through the radical rearrangement and hydrogen transfer 2 steps, which is given by Equation 15.

$$a = e^{(\Delta S_{21}^{\ddagger} - \Delta S_{23}^{\ddagger})/R} e^{m\omega r^2/2\hbar} \quad (15)$$

Using previous estimates for m and ω, the calculated value of r, and the fitted value of a, we obtain an activation entropy difference of 11 cal/mol/K for ^{1}H, and 16 cal/mol/K for ^{2}H. These values are consistent with $\frac{\Delta S_H^{\ddagger}}{\Delta S_D^{\ddagger}} \sim 1$ as previously assumed. Moreover, the values for the activation entropy difference implies that across the radical rearrangement and hydrogen transfer 2 steps the system decreases its entropy by 11 and 16 cal/mol/K for ^{1}H and ^{2}H, with the remaining entropy being released in subsequent catalytic steps.

5.5 Conclusions

The interpretation of EAL catalysis provided in this chapter proposes a three-state, two-step model for the decay of the Co(II)-substrate radical pair state. This system is kinetically complex because multiple steps, namely radical rearrangement and the second hydrogen transfer steps, factor into rate determination. For ^{1}H, radical rearrangement is dominant over the entire temperature range examined, leading to a relatively linear Arrhenius plot. For ^{2}H, isotope effects from the hydrogen transfer become more apparent, and lead to the deviation of the isotope effect data at higher temperatures. This hydrogen transfer is revealed to proceed through a tunneling mechanism, over a distance of 2 Å, subject to certain assumptions. Using this scheme, it is possible to explain the non-linear

isotope effect seen over the 190 to 295 K temperature range examined, and explains the paradox of the anomalously high $^{1}H/^{3}H$ isotope effect observed over 40 years prior.[57] The second hydrogen transfer step is partially masked by the radical rearrangement step so only a fraction of the true isotope effect is visible in the $^{1}H/^{2}H$ isotope effect measurements. The larger free energy barrier of the ^{3}H hydrogen transfer 2 step causes the ^{3}H reaction to show an anomalously large $^{1}H/^{3}H$ isotope effect because in the ^{3}H reaction hydrogen transfer 2 contributes more to rate determination (less masked by radical rearrangement). This mechanism also accounts for the comparable isotope effect anomaly observed in the related B_{12}-dependent enzyme, diol dehydratase.[177]

Further, the initial steps of EAL catalysis of radical pair separation and hydrogen transfer 1 show significant entropic contributions to the catalytic free energy reduction.[46] Here it is shown that this entropy begins to be returned in the subsequent catalytic steps of radical rearrangement and hydrogen transfer 2, with a -11 cal/mol/K change in ^{1}H-aminoethanol samples, and a -16 cal/mol/K change with ^{2}H-labeled substrate. A large entropic component to catalysis has also been seen in the related B_{12}-dependent ribonucleotide triphosphate reductase, where a 70 cal/mol/K entropy change has been observed in the Co-C bond cleavage and Co(II)-radical pair formation steps.[186] Thus, entropy driven catalysis may be a general model for enzymatic rate enhancement, in addition to the traditional enthalpy driven models usually discussed.

Chapter 6

Implications for B_{12} Enzymology

6

The results discussed in this dissertation represent a significant step forward in the field of enzymology, and provide a framework for the interpretation of past and future experimental results. In this final chapter, the main conclusions of each chapter will be summarized, and the ramifications for (a) EAL structure and function, (b) B_{12} enzyme catalysis kinetics and molecular mechanism, and (c) general mechanisms of enzyme catalysis will be examined. From this, a general understanding of the far-reaching scope of the experiments conducted herein can be appreciated.

6.1 Chapter Two: The Structural Model of *Salmonella Typhimurium* Ethanolamine Ammonia-Lyase Directs a Rational Approach to the Assembly of the Functional $[(EutB\text{-}EutC)_2]_3$ Oligomer from Isolated Subunits

Chapter Two provides a description of the structural model of EAL. This comparative model structure led to the development of a protocol for the reassembly of the individually purified EutB and EutC subunits of EAL. Combining these individually purified subunits, 70% of the wild-type ethanolamine activity could be recovered, and active-site congruence of the *(S)*-2-aminopropanol intermediate was verified with EPR spectroscopy.

In general, enzyme structures are vital for identifying critical catalytic residues, and for the determination of detailed microscopic interactions, including the hydrogen bonding network necessary for catalysis. In order to make such determinations on EAL, a comparative structural model for the enzyme was developed. The homology modeled structure, presented in Chapter Two, is the first published structure of the full EAL

$[(EutB\text{-}EutC)_2]_3$ oligomer in *S. typhimurium*.[73] Previous EAL X-ray crystal structures have been determined for the EAL enzyme of *E. coli*,[43] and *L. monocytogenes* (PDB ID: 2QEZ). Though these structures were instrumental in the development of this work, and B_{12} enzymology in general, they were not ideal, as the majority of functional experiments,[1] and the associated residue-specific, mechanistic determinations,[16] were obtained for *S. typhimurium* EAL. Despite the high, 90% sequence homology of *E. coli* EAL with that of *S. typhimurium*,[73] the *E. coli* structure is still inappropriate for comparison with the Angstrom-level EPR-spectroscopic determinations of the reactant structures in the *S. typhimurium* EAL active-site.[37, 48, 49, 80] A previous modeled EAL structure of the *S. typhimurium* enzyme contained only the EutB subunit.[42] Though the EutB subunit hosts the ethanolamine binding-site, residues within the EutC subunit also interact with the cofactor. Moreover, in the absence of EutC, the EutB subunit is unable to support catalysis or generate the catalytically relevant substrate radical species.[73] There is some, as of yet undetermined mechanism for the involvement of EutC in EAL catalysis. The full structure model presented here serves as a platform to begin the molecular-level interpretation of both structural and functional data obtained from EAL spectroscopic studies.

The first step of EAL catalysis, cobalt-carbon bond cleavage, is characterized by a large favorable entropic contribution to the activation free energy. [46] It has been posited that this large entropic contribution during the initial phases of reaction arises from protein configuration changes, and the associated changes in the interaction of protein residues with hydration layer water molecules.[136] The subsequent steps of catalysis, radical rearrangement and hydrogen transfer 2, also show evidence of a dynamical protein

contribution to the reaction.[45] The EAL structure described herein allows access to the next stage in understanding EAL catalysis; the assignment of the molecular basis for the protein contribution to catalysis. Thus, this structure is crucial in building a more complete model of EAL catalysis

The development of a protocol for the reconstitution of biologically active EAL from its individual subunits allows for high-throughput mutagenesis of EAL. Using the structure model as a guide, site-directed mutagenesis can be performed on individual subunits, and these independently modified subunits can be subsequently reconstituted. This protocol enables an extensive survey of the role of selected residues in EAL, which is crucial for the mission of understanding B_{12} catalysis.

In the course of the development of the reconstitution system, several benefits were achieved. Due to the instability of the isolated subunits, it was necessary to optimize the expression protocol for the individual subunits of EAL. The final expression system utilized *lac*UV5 promoter-controlled gene expression from a pET28a vector which conferred ampicillin resistance. With 30°C expression, high yields of the individual subunits could be attained. In addition, it was necessary to develop a robust purification protocol for wild-type EAL, the individual EutB and EutC subunits, and mutant EAL proteins. Here, an affinity purification strategy was employed by using a genetically incorporated polyhistidine tag, followed by a size-exclusion procedure to further purify the resulting protein. This is in contrast to previous EAL purifications, which have relied on the low solubility of EAL for separation of the protein from the other cellular components, by using an ammonium sulfate precipitation procedure.[14] Such a procedure has the limitation that optimization of the precipitation step is required for each protein,

as mutations may alter the solubility characteristics. The histidine tag purification strategy utilized herein avoids this complication, because of the specificity of the engineered tag. The affinity method of protein purification has recently been applied to EAL independently.[15] Here, we successfully develop and extend this strategy to the purification of the individual subunits.

Overall, the system developed here addresses a long-standing challenge of functional enzyme reconstitution, which existed from the initial cloning of EAL.[14, 41] Further, the expression and purification protocol described, in conjunction with the homology modeled structure, also serve as a platform for EAL mutagenesis, which was utilized in Chapters Three and Four of this work, and will undoubtedly serve as a platform for future mutagenesis experiments on *S. typhimurium* EAL.

6.2 Chapter Three: Cobalamin Binding to Ethanolamine Ammonia-Lyase

Chapter Three describes the creation of an EAL mutant with phenylalanine residue 258 on the EutB subunit replaced with a tryptophan (Trp) residue (αF258W). Fluorescence quenching experiments on this mutant led to the direct determination of the 5'-deoxyadenosyl cobalamin (AdoCbl) dissociation constant (K_D) for EAL as 1.2 μM.

The direct determination of K_D presented in Chapter Three is in stark contrast to the previous experimental estimate of K_D,[122] which was assumed to be equal to the cofactor Michaelis constant, K_M. This assumption is only valid if the steps subsequent to cofactor binding are sufficiently slow so that the on and off rates of the cofactor-protein interaction dominate K_M. However, since the cofactor exchange rate is unknown, the validity of using K_M as a proxy for K_D is undetermined.

Here, fluorescence quenching experiments were performed to directly determine K_D. This approach avoids obfuscating the results with kinetic effects as in previous studies.[122] Quenching experiments on EutB indicated that the Trp residues in EutB were too distant to be efficiently quenched by the binding cobalamins. This issue was overcome with the EAL mutant, αF258W, which retained activity and whose fluorescence was sensitive to cobalamin binding. The utility of this mutant serves as further verification of the reliability of the modeled structure.

This αF258W mutant opens the door to future fluorescence quenching experiments on EAL. The binding of cofactor analogs, such as methylcobalamin or methylcobinamide, can probe the specific role of individual groups, such as the 5'-deoxyadenosyl group, or the lower dimethylbenzimidazole (DMB) ligand, to cofactor binding. These studies will separate the binding free energy contribution of each cofactor moiety, and further reveal their relation to catalysis.

6.3 Chapter Four: Towards a Reactive Cysteine-Free EAL

Chapter Four describes the effort towards the site-directed spin labeling (SDSL)[67, 68] of EAL. An EAL mutant with five solvent-accessible cysteines replaced with alanine residues (EAL-5CF) was created, and this mutant was used as a starting point for the insertion of targeted spin labeling sites. During the development of the reactive-cysteine-free EAL mutant, the primary spin labeling site of wild-type EAL was identified as EutC residue, C37.

The site-specific attachment of motion-sensitive EPR-active labels to EAL presents an opportunity, wherein the radical catalysis occurring on the interior of the protein can be

followed concomitantly with the motions on the surface of the enzyme, in a single experiment. The potential for correlating the chemical steps of a reaction with solvation layer dynamics is on the forefront of enzymology research, as the role of dynamics in catalysis remains deeply controversial.[173] EAL can serve as a model system to address such questions, because of the unique ability to isolate and study individual chemical steps along the reaction cycle. A necessary antecedent for such studies is the ability to site-specifically label the EAL enzyme.

In this work, significant progress has been made towards the development of a reactive-cysteine-free EAL mutant, into which targeted labeling sites can be reengineered. EAL-5CF, a mutant which eliminated selected solvent-accessible cysteines while retaining activity, was created for this purpose. However, EAL-5CF possessed background sulfhydryl reactivity which was not eliminated with the introduction of targeted labeling sites. Nonetheless, subtle changes in the spectra indicate that the engineered sites are at least partially labeled. Further work is needed to identify and eliminate the sites responsible for background sulfhydryl reactivity to realize the goal of a reactive-cysteine-free EAL that retains biological activity.

During the development of the EAL-5CF mutant, several mutant EAL proteins were engineered. Experiments performed on a subset of these EAL mutants, the cysteine→alanine single mutant library, revealed that the primary labeling site for wild-type EAL with 4-maleimido-TEMPO (4MT) is EutC-C37. Though this residue is absent in the modeled structure, it can be localized to a region on the surface of EAL. Despite the lack of a precise position for the attached spin label, which is not ideal, important structural information can be gleaned. For example, a low temperature spin label

relaxation study revealed that interactions with the ribose ring of the binding AdoCbl cofactor leads to a change in the T_1 relaxation temperature-dependence of an attached spin label. (Sun, L., Bovell, A.M., Warncke, K., in preparation.) This is interpreted as being caused by a change in protein structure that is critical to triggering cobalt-carbon bond cleavage (Sun, L., Bovell, A.M., Warncke, K., in preparation). In that study, the distal position of the spin label from the AdoCbl binding site was enough to eliminate the trivial effect caused by a direct spin label-AdoCbl interaction.

In all, the experimental work presented in Chapter Four provides a solid foundation for future site-specific EAL labeling with spin-labels, and more generally, any cysteine-conjugating probes (such as fluorescent labels). This system has the potential to answer pressing questions on the coupling of solvent motions to the chemical steps of catalysis.

6.4 Chapter Five: Temperature-Dependence of Hydrogen Isotope Effects on the Coupled Radical Rearrangement-Hydrogen Transfer Reaction Sequence from 190 to 295 K

Chapter Five describes functional experiments on EAL catalysis with ^{1}H and ^{2}H labeled ethanolamine substrates from 190 to 295 K. It was revealed that a single kinetic model could fit the data over this 105 K temperature range. This scheme relied on multi-step rate determination by radical rearrangement and hydrogen transfer 2 steps, which lead to the observed temperature-dependent isotope effect.

The model presented in Chapter Five provides an explanation for the previously unresolved observation from over 40 years prior, wherein the EAL $^{1}H/^{3}H$ isotope effect exceeded the Swain-Schaad prediction from the $^{1}H/^{2}H$ isotope effect.[177] Moreover, this

model may also apply to other B_{12}-dependent isomerases, such as diol dehydratase, which also shows an anomalous isotope effect.[57] Thus, multi-step rate determination may be a general feature of B_{12}-dependent enzymes which can be revealed by studies over expansive temperature ranges as performed here.

Analysis of the data fit parameters revealed that the hydrogen transfer 2 reaction proceeds by quantum mechanical tunneling. It is likely that tunneling is a general mechanism of hydrogen transfer in B_{12} enzymes, as in methlmalonyl-CoA mutase,[174, 175] glutamate mutase,[176] and diol dehyratase,[57] which all show evidence of tunneling steps. Further data analysis presented in Chapter Five revealed that over the rate-limiting steps of radical rearrangement and hydrogen transfer 2 there is a decrease in entropy. This is consistent with previous data that showed that the initial steps of radical pair separation and hydrogen transfer involve a significant entropic component.[46, 136] This favorable activation entropy may be caused by changes in EAL structure, and interactions with hydration layer water molecules, as has been previously suggested.[136] This entropy-driven reaction model also necessitates that EAL must revert to its original structure at the end of its catalytic cycle by repaying the entropy of the initial steps of the reaction, as is observed in the radical rearrangement, and hydrogen transfer 2 steps studied here. Precedence for an entropy-driven catalysis model can be found in B_{12}-dependent ribonucleotide triphosphate reductase, (RTPR) where the initial steps contain a significant entropic component. However, RTPR turnover proceeds via a unique, thiyl radical based catalytic mechanism,[34] in contrast to the general, 5'-deoxyadenosyl radical based mechanism of other isomerases, including EAL. As such, EAL is a preferable model system to further probe the entropic mechanism of catalysis.

6.5 Conclusions

The findings presented here represent a significant step forward in the understanding of B_{12}-dependent EAL catalysis, and, to the extent that this system serves as a model for enzyme catalysis, enzymology in general. The ability to isolate and observe the kinetics of individual steps in the catalytic cycle positions EAL as an ideal model system. As previously stated, the groundwork provided here paves the way for the use of EAL as a model system for the study of the pertinent dynamics in enzyme catalysis. In addition, the revelation that EAL utilizes a tunneling mechanism for hydrogen transfer further increases the value of understanding EAL catalysis. EAL catalyzes the chemically difficult isomerization and deamination of ethanolamine, by using highly reactive radicals, entropy, and quantum mechanical hydrogen transfer. The design of a biomimetic catalyst using the physical and chemical mechanisms within the EAL toolkit would represent a significant advance in the field enzymology.

.

References

(1) Garsin, D. A. (2010) Ethanolamine utilization in bacterial pathogens: roles and regulation, *Nat. Rev. Microbiol. 8*, 290-295.

(2) Kaplan, B. H., and Stadtman, E. R. (1968) Ethanolamine Deaminase a Cobamide Coenzyme-Dependent Enzyme .I. Purification Assay and Properties of Enzyme, *J. Biol. Chem. 243*, 1787-1793.

(3) Babior, B. M. (1969) Mechanism of Action of Ethanolamine Deaminase .I. Studies with Isotopic Hydrogen and Oxygen, *J. Biol. Chem. 244*, 449-456.

(4) Finke, R. G., and Hay, B. P. (1984) Thermolysis of Adenosylcobalamin - a Product, Kinetic, and Co-C5' Bond-Dissociation Energy Study, *Inorg. Chem. 23*, 3041-3043.

(5) Que, L., and Tolman, W. B. (2008) Biologically inspired oxidation catalysis, *Nature 455*, 333-340.

(6) Wells, J. A., and Estell, D. A. (1988) Subtilisin - an Enzyme Designed to Be Engineered, *Trends Biochem. Sci. 13*, 291-297.

(7) Srikumar, S., and Fuchs, T. M. (2011) Ethanolamine utilization contributes to proliferation of Salmonella enterica serovar Typhimurium in food and in nematodes, *Appl. Environ. Microbiol. 77*, 281-290.

(8) Anfinsen, C. B. (1973) Principles that Govern the Folding of Protein Chains, *Science 181*, 223-230.

(9) Dyson, H. J., and Wright, P. E. (2005) Intrinsically unstructured proteins and their functions, *Nat. Rev. Mol. Cell Bio. 6*, 197-208.

(10) Pang, J. Y., Li, X., Morokuma, K., Scrutton, N. S., and Sutcliffe, M. J. (2012) Large-Scale Domain Conformational Change Is Coupled to the Activation of the Co-C

Bond in the B-12-Dependent Enzyme Ornithine 4,5-Aminomutase: A Computational Study, *J. Am. Chem. Soc. 134*, 2367-2377.

(11) Jones, A. R., Hardman, S. J. O., Hay, S., and Scrutton, N. S. (2011) Is There a Dynamic Protein Contribution to the Substrate Trigger in Coenzyme B-12-Dependent Ethanolamine Ammonia Lyase?, *Angew. Chem. Int. Edit. 50*, 10843-10846.

(12) Silva, R. G., Murkin, A. S., and Schramm, V. L. (2011) Femtosecond dynamics coupled to chemical barrier crossing in a Born-Oppenheimer enzyme, *P. Natl. Acad. Sci. U.S.A. 108*, 18661-18665.

(13) Bhabha, G., Lee, J., Ekiert, D. C., Gam, J., Wilson, I. A., Dyson, H. J., Benkovic, S. J., and Wright, P. E. (2011) A dynamic knockout reveals that conformational fluctuations influence the chemical step of enzyme catalysis, *Science 332*, 234-238.

(14) Faust, L. P., and Babior, B. M. (1992) Overexpression, Purification, and Some Properties of the Adocbl-Dependent Ethanolamine Ammonia-Lyase from Salmonella-Typhimurium, *Arch. Biochem. Biophys. 294*, 50-54.

(15) Akita, K., Hieda, N., Baba, N., Kawaguchi, S., Sakamoto, H., Nakanishi, Y., Yamanishi, M., Mori, K., and Toraya, T. (2010) Purification and some properties of wild-type and N-terminal-truncated ethanolamine ammonia-lyase of Escherichia coli, *J. Biochem. 147*, 83-93.

(16) Sun, L., Groover, O. A., Canfield, J. M., and Warncke, K. (2008) Critical role of arginine 160 of the EutB protein subunit for active site structure and radical catalysis in coenzyme B-12-dependent ethanolamine ammonia-lyase, *Biochemistry 47*, 5523-5535.

(17) Bradbeer, C. (1965) Clostridial Fermentations of Choline and Ethanolamine .I. Preparation and Properties of Cell-Free Extracts, *J. Biol. Chem. 240*, 4669-4674.

(18) Banerjee, R. (1999) *Chemistry and Biochemistry of B12*, Wiley, New York.

(19) Roof, D. M., and Roth, J. R. (1988) Ethanolamine Utilization in Salmonella-Typhimurium, *J. Bacteriol. 170*, 3855-3863.

(20) Roof, D. M., and Roth, J. R. (1989) Functions Required for Vitamin-B12-Dependent Ethanolamine Utilization in Salmonella-Typhimurium, *J. Bacteriol. 171*, 3316-3323.

(21) Jones, P. W., and Turner, J. M. (1984) Interrelationships between the Enzymes of Ethanolamine Metabolism in Escherichia-Coli, *J. Gen. Microbiol. 130*, 299-308.

(22) Stojiljkovic, I., Baumler, A. J., and Heffron, F. (1995) Ethanolamine utilization in Salmonella typhimurium: nucleotide sequence, protein expression, and mutational analysis of the cchA cchB eutE eutJ eutG eutH gene cluster, *J. Bacteriol. 177*, 1357-1366.

(23) Kofoid, E., Rappleye, C., Stojiljkovic, I., and Roth, J. (1999) The 17-gene ethanolamine (eut) operon of Salmonella typhimurium encodes five homologues of carboxysome shell proteins, *J. Bacteriol. 181*, 5317-5329.

(24) Yeates, T. O., Kerfeld, C. A., Heinhorst, S., Cannon, G. C., and Shively, J. M. (2008) Protein-based organelles in bacteria: carboxysomes and related microcompartments, *Nat. Rev. Microbiol. 6*, 681-691.

(25) Tcherkez, G. G. B., Farquhar, G. D., and Andrews, T. J. (2006) Despite slow catalysis and confused substrate specificity, all ribulose bisphosphate carboxylases may be nearly perfectly optimized, *P. Natl. Acad. Sci. U.S.A. 103*, 7246-7251.

(26) Brinsmade, S. R., Paldon, T., and Escalante-Semerena, J. C. (2005) Minimal functions and physiological conditions required for growth of Salmonella enterica on ethanolamine in the absence of the metabolosome, *J. Bacteriol. 187*, 8039-8046.

(27) Randle, C. L., Albro, P. W., and Dittmer, J. C. (1969) Phosphoglyceride Composition of Gram-Negative Bacteria and Changes in Composition during Growth, *Biochim. Biophys. Acta 187*, 214-220.

(28) Li, H., Kristensen, D. M., Coleman, M. K., and Mushegian, A. (2009) Detection of biochemical pathways by probabilistic matching of phyletic vectors, *PLoS One 4*, e5326.

(29) Kendall, M. M., Gruber, C. C., Parker, C. T., and Sperandio, V. (2012) Ethanolamine controls expression of genes encoding components involved in interkingdom signaling and virulence in enterohemorrhagic Escherichia coli O157:H7, *mBio 3*.

(30) Mead, P. S., Slutsker, L., Dietz, V., McCaig, L. F., Bresee, J. S., Shapiro, C., Griffin, P. M., and Tauxe, R. V. (1999) Food-related illness and death in the United States, *Emerg. Infect. Dis. 5*, 607-625.

(31) Braden, C. R. (2006) Salmonella enterica serotype enteritidis and eggs: A national epidemic in the United States, *Clin. Infect. Dis. 43*, 512-517.

(32) Banerjee, R., and Ragsdale, S. W. (2003) The many faces of vitamin B12: catalysis by cobalamin-dependent enzymes, *Annu. Rev. Biochem. 72*, 209-247.

(33) Marsh, E. N. (1999) Coenzyme B12 (cobalamin)-dependent enzymes, *Essays Biochem. 34*, 139-154.

(34) Toraya, T. (2003) Radical Catalysis in Coenzyme B12-Dependent Isomerization (Eliminating) Reactions, *Chem. Rev. 103*, 2095-2127.

(35) Stabler, S. P. (1999) *B12 and Nutrition*, John Wiley and Sons, New York.

(36) Bucher, D., Sandala, G. M., Durbeej, B., Radom, L., and Smith, D. M. (2012) The Elusive 5 '-Deoxyadenosyl Radical in Coenzyme-B-12-Mediated Reactions, *J. Am. Chem. Soc. 134*, 1591-1599.

(37) Canfield, J. M., and Warncke, K. (2002) Geometry of reactant centers in the Co-II-substrate radical pair state of coenzyme B-12-dependent ethanolamine deaminase determined by using orientation-selection-ESEEM spectroscopy, *J. Phys. Chem. B 106*, 8831-8841.

(38) Pilbrow, J. (1982) EPR of B12-Dependent Enzyme Reactions and Related Systems, In *B12* (Dolphin, D., Ed.), pp 431-462, Wiley, New York.

(39) Babior, B. M., Moss, T. H., Orme-Johnson, W. H., and Beinert, H. (1974) Mechanism of Action of Ethanolamine Ammonia-Lyase, a B12-Dependent Enzyme - Participation of Paramagnetic Species in Catalytic Deamination of 2-Aminopropanol, *J. Biol. Chem. 249*, 4537-4544.

(40) Retey, J., Suckling, C. J., Arigoni, D., and Babior, B. M. (1974) Stereochemistry of Reaction Catalyzed by Ethanolamine Ammonia-Lyase, an Adenosylcobalamin-Dependent Enzyme - Example of Racemization Accompanying Substitution, *J. Biol. Chem. 249*, 6359-6360.

(41) Faust, L. R. P., Connor, J. A., Roof, D. M., Hoch, J. A., and Babior, B. M. (1990) Cloning, Sequencing, and Expression of the Genes Encoding the Adenosylcobalamin-Dependent Ethanolamine Ammonia-Lyase of Salmonella-Typhimurium, *J. Biol. Chem. 265*, 12462-12466.

(42) Sun, L., and Warncke, K. (2006) Comparative model of EutB from coenzyme B12-dependent ethanolamine ammonia-lyase reveals a beta8alpha8, TIM-barrel fold and radical catalytic site structural features, *Proteins 64*, 308-319.

(43) Shibata, N., Tamagaki, H., Hieda, N., Akita, K., Komori, H., Shomura, Y., Terawaki, S., Mori, K., Yasuoka, N., Higuchi, Y., and Toraya, T. (2010) Crystal Structures of Ethanolamine Ammonia-lyase Complexed with Coenzyme B-12 Analogs and Substrates, *J. Biol. Chem. 285*, 26484-26493.

(44) Wierenga, R. K. (2001) The TIM-barrel fold: a versatile framework for efficient enzymes, *Febs Lett. 492*, 193-198.

(45) Zhu, C., and Warncke, K. (2008) Reaction of the Co(II)-Substrate Radical Pair Catalytic Intermediate in Coenzyme B(12)-Dependent Ethanolamine Ammonia-Lyase in Frozen Aqueous Solution from 190 to 217 K, *Biophys. J. 95*, 5890-5900.

(46) Wang, M., and Warncke, K. (2008) Kinetic and thermodynamic characterization of Co-II-substrate radical pair formation in coenzyme B-12-dependent ethanolamine ammonia-lyase in a cryosolvent system by using time-resolved, full-spectrum continuous-wave electron paramagnetic resonance spectroscopy, *J. Am. Chem. Soc. 130*, 4846-4858.

(47) Zhu, C., and Warncke, K. (2010) Kinetic Isolation and Characterization of the Radical Rearrangement Step in Coenzyme B-12-Dependent Ethanolamine Ammonia-lyase, *J. Am. Chem. Soc. 132*, 9610-9615.

(48) Canfield, J. M., and Warncke, K. (2005) Active site reactant center geometry in the Co(II)-product radical pair state of coenzyme B12-dependent ethanolamine deaminase

determined by using orientation-selection electron spin-echo envelope modulation spectroscopy, *J. Phys. Chem. B 109*, 3053-3064.

(49) LoBrutto, R., Bandarian, V., Magnusson, O. T., Chen, X., Schramm, V. L., and Reed, G. H. (2001) 5'-Deoxyadenosine contacts the substrate radical intermediate in the active site of ethanolamine ammonia-lyase: 2H and 13C electron nuclear double resonance studies, *Biochemistry 40*, 9-14.

(50) Harkins, T. T., and Grissom, C. B. (1995) The Magnetic-Field Dependent Step in Bit Ethanolamine Ammonia-Lyase Is Radical-Pair Recombination, *J. Am. Chem. Soc. 117*, 566-567.

(51) Robertson, W. D., and Warncke, K. (2009) Photolysis of Adenosylcobalamin and Radical Pair Recombination in Ethanolamine Ammonia-Lyase Probed on the Micro- to Millisecond Time Scale by Using Time-Resolved Optical Absorption Spectroscopy, *Biochemistry 48*, 140-147.

(52) Krouwer, J. S., Holmquist, B., Kipnes, R. S., and Babior, B. M. (1980) Mechanism of Action of Ethanolamine Ammonia-Lyase, an Adenosylcobalamin-Dependent Enzyme - Evidence That Carbon-Cobalt Bond-Cleavage Is Driven in Part by Conformational Alterations of the Corrin Ring, *Biochim. Biophys. Acta 612*, 153-159.

(53) Bender, G., Poyner, R. R., and Reed, G. H. (2008) Identification of the Substrate Radical Intermediate Derived from Ethanolamine during Catalysis by Ethanolamine Ammonia-Lyase, *Biochemistry 47*, 11360-11366.

(54) Harkins, T. T., and Grissom, C. B. (1994) Magnetic-Field Effects on B-12 Ethanolamine Ammonia-Lyase - Evidence for a Radical Mechanism, *Science 263*, 958-960.

(55) Jones, A. R., Hay, S., Woodward, J. R., and Scrutton, N. S. (2007) Magnetic field effect studies indicate reduced geminate recombination of the radical pair in substrate-bound adenosylcobalamin-dependent ethanolamine ammonia lyase, *J. Am. Chem. Soc. 129*, 15718-15727.

(56) Bandarian, V., and Reed, G. H. (2000) Isotope effects in the transient phases of the reaction catalyzed by ethanolamine ammonia-lyase: determination of the number of exchangeable hydrogens in the enzyme-cofactor complex, *Biochemistry 39*, 12069-12075.

(57) Weisblat, D. A., and Babior, B. M. (1971) Mechanism of Action of Ethanolamine Ammonia-Lyase, a B12-Dependent Enzyme .8. Further Studies with Compounds Labeled with Isotopes of Hydrogen - Identification and Some Properties of Rate-Limiting Step, *J. Biol. Chem. 246*, 6064-6071.

(58) Hay, S., Sutcliffe, M. J., and Scrutton, N. S. (2007) Promoting motions in enzyme catalysis probed by pressure studies of kinetic isotope effects, *P. Natl. Acad. Sci. U.S.A. 104*, 507-512.

(59) Ling, M. M., and Robinson, B. H. (1997) Approaches to DNA mutagenesis: An overview, *Anal. Biochem. 254*, 157-178.

(60) Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., and Pease, L. R. (1989) Site-Directed Mutagenesis by Overlap Extension Using the Polymerase Chain-Reaction, *Gene 77*, 51-59.

(61) Likhtenshtein, G. (1976) Spin Labelling Methods in Molecular Biology, Wiley, New York.

(62) Eaton, G. R., Eaton, S. S., Barr, D. P., and Weber, R. T. (2010) *Quantitative EPR*, Springer.

(63) Stoll, S., and Schweiger, A. (2007) *ESR Spectroscopy in Membrane Biophysics*, Springer.

(64) Nielsen, R. D., and Robinson, B. H. (2004) A novel relaxation equation of motion, *J. Phys. Chem. A 108*, 1589-1600.

(65) Wong, S. S. (1993) Chemistry of protein conjugation and cross-linking.

(66) Hermanson, G. T. (1996) *Bioconjugate techniques*, Academic press.

(67) Hubbell, W. L., Mchaourab, H. S., Altenbach, C., and Lietzow, M. A. (1996) Watching proteins move using site-directed spin labeling, *Structure 4*, 779-783.

(68) Mchaourab, H. S., Lietzow, M. A., Hideg, K., and Hubbell, W. L. (1996) Motion of spin-labeled side chains in T4 lysozyme, correlation with protein structure and dynamics, *Biochemistry 35*, 7692-7704.

(69) Rabenstein, M. D., and Shin, Y. K. (1995) Determination of the distance between two spin labels attached to a macromolecule, *P. Natl. Acad. Sci. U. S. A. 92*, 8239-8243.

(70) Jeschke, G. (2012) DEER Distance Measurements on Proteins, *Annu. Rev. Phys. Chem. 63*, 419-446.

(71) Farahbakhsh, Z. T., Hideg, K., and Hubbell, W. L. (1993) Photoactivated Conformational-Changes in Rhodopsin - a Time-Resolved Spin-Label Study, *Science 262*, 1416-1419.

(72) Steinhoff, H. J., Mollaaghababa, R., Altenbach, C., Hideg, K., Krebs, M., Khorana, H. G., and Hubbell, W. L. (1994) Time-Resolved Detection of Structural-Changes during the Photocycle of Spin-Labeled Bacteriorhodopsin, *Science 266*, 105-107.

(73) Bovell, A. M., and Warncke, K. (2013) The Structural Model of Salmonella typhimurium Ethanolamine Ammonia-Lyase Directs a Rational Approach to the Assembly of the Functional [(EutB–EutC)2]3 Oligomer from Isolated Subunits, *Biochemistry 52*, 1419-1428.

(74) Bandarian, V., and Reed, G. H. (1999) Ethanolamine ammonia-lyase, In *Chemistry and Biochemistry of B12* (Banerjee, R., Ed.), Wiley, New York.

(75) Frey, P. A. (2010) Cobalamin coenzymes in enzymology, In *Comprehensive Natural Products II Chemistry and Biology* (Mander, L., and Lui, H.-W., Eds.), pp 501-546, Elsevier, Oxford UK.

(76) Hubbard, B. K., Gulick, A. M., Babbitt, P. C., Rayment, I., and Gerlt, J. A. (1999) Evolution of enzymatic activities in the enolase superfamily: Mechanism, structure, and metabolic context of glucarate dehydratase from Escherichia coli, *Faseb J. 13*, A1446-A1446.

(77) Brown, K. L. (2005) Chemistry and Enzymology of Vitamin B12, *Chem. Rev. 105*, 2075-2149.

(78) Penrod, J. T., and Roth, J. R. (2006) Conserving a volatile metabolite: a role for carboxysome-like organelles in Salmonella enterica, *J. Bacteriol. 188*, 2865-2874.

(79) Stubbe, J., and van der Donk, W. (1998) Protein radicals in enzyme catalysis, *Chem. Rev. 98*, 705-762.

(80) Warncke, K., and Utada, A. S. (2001) Interaction of the substrate radical and the 5 '-deoxyadenosine-5 '-methyl group in vitamin B-12 coenzyme-dependent ethanolamine deaminase, *J. Am. Chem. Soc. 123*, 8564-8572.

(81) Poyner, R. R., Anderson, M. A., Bandarian, V., Cleland, W. W., and Reed, G. H. (2006) Probing nitrogen-sensitive steps in the free-radical-mediated deamination of amino alcohols by ethanolamine ammonia-lyase, *J. Am. Chem. Soc. 128*, 7120-7121.

(82) Gerlt, J. A., and Raushel, F. M. (2003) Evolution of function in $(\beta\alpha)_8$-barrel enzymes, *Curr. Opin. Chem. Biol. 7*, 252-264.

(83) Bandarian, V., and Reed, G. H. (1999) Hydrazine cation radical in the active site of ethanolamine ammonia-lyase: Mechanism-based inactivation by hydroxyethylhydrazine, *Biochemistry 38*, 12394-12402.

(84) Guruprasad, K., Reddy, B. V. B., and Pandit, M. W. (1990) Correlation between Stability of a Protein and Its Dipeptide Composition - a Novel-Approach for Predicting Invivo Stability of a Protein from Its Primary Sequence, *Protein Eng. 4*, 155-161.

(85) Goujon, M., McWilliam, H., Li, W. Z., Valentin, F., Squizzato, S., Paern, J., and Lopez, R. (2010) A new bioinformatics analysis tools framework at EMBL-EBI, *Nucleic Acids Res. 38*, W695-W699.

(86) Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007) Clustal W and clustal X version 2.0, *Bioinformatics 23*, 2947-2948.

(87) Eswar, N., Webb, B., Marti-Renom, M. A., Madhusudhan, M. S., Eramian, D., Shen, M. Y., Pieper, U., and Sali, A. (2007) Comparative protein structure modeling using MODELLER, *Curr. Protoc. Protein Sci. Chapter 2*, Unit 2 9.

(88) Shen, M. Y., and Sali, A. (2006) Statistical potential for assessment and prediction of protein structures, *Protein Sci. 15*, 2507-2524.

(89) Laskowski, R. A., Macarthur, M. W., Moss, D. S., and Thornton, J. M. (1993) Procheck - a Program to Check the Stereochemical Quality of Protein Structures, *J. Appl. Crystallogr. 26*, 283-291.

(90) Ahmad, S., Gromiha, M., Fawareh, H., and Sarai, A. (2004) ASAView: database and tool for solvent accessibility representation in proteins, *BMC Bioinformatics 5*, 51.

(91) Kabsch, W., and Sander, C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features, *Biopolymers 22*, 2577-2637.

(92) Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001) Electrostatics of nanosystems: Application to microtubules and the ribosome, *P. Natl. Acad. Sci. U.S.A. 98*, 10037-10041.

(93) Dolinsky, T. J., Czodrowski, P., Li, H., Nielsen, J. E., Jensen, J. H., Klebe, G., and Baker, N. A. (2007) PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations, *Nucleic Acids Res. 35*, W522-W525.

(94) Dolinsky, T. J., Nielsen, J. E., McCammon, J. A., and Baker, N. A. (2004) PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations, *Nucleic Acids Res. 32*, W665-W667.

(95) Finn, R. D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J. E., Gavin, O. L., Gunasekaran, P., Ceric, G., Forslund, K., Holm, L., Sonnhammer, E. L., Eddy, S. R., and Bateman, A. (2010) The Pfam protein families database, *Nucleic Acids Res. 38*, D211-222.

(96) Bradford, M. M. (1976) Rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Anal. Biochem. 72*, 248-254.

(97) Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., and Bairoch, A. (2005) Protein Identification and Analysis Tools on the ExPASy Server, In *The Proteomics Protocols Handbook* (Walker, J. M., Ed.), pp 571-607, Humana Press.

(98) Laemmli, U. K. (1970) Cleavage of Structural Proteins during Assembly of Head of Bacteriophage-T4, *Nature 227*, 680-685.

(99) Ornstein, L. (1964) Disc Electrophoresis .I. Background and Theory, *Ann. N.Y. Acad. Sci. 121*, 321-349.

(100) Sali, A., and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints, *J. Mol. Biol. 234*, 779-815.

(101) Engh, R. A., and Huber, R. (1991) Accurate Bond and Angle Parameters for X-Ray Protein-Structure Refinement, *Acta Crystallogr. A 47*, 392-400.

(102) Mitov, M. I., Greaser, M. L., and Campbell, K. S. (2009) GelBandFitter--a computer program for analysis of closely spaced electrophoretic and immunoblotted bands, *Electrophoresis 30*, 848-851.

(103) Samuelson, J. C. (2011) Recent developments in difficult protein expression: a guide to E. coli strains, promoters, and relevant host mutations, *Method. Mol. Biol. 705*, 195-209.

(104) Arakawa, T., and Kita, Y. (2000) Protection of bovine serum albumin from aggregation by tween 80, *J. Pharm. Sci. 89*, 646-651.

(105) Choudhary, S., Quin, M. B., Sanders, M. A., Johnson, E. T., and Schmidt-Dannert, C. (2012) Engineered protein nano-compartments for targeted enzyme localization, *PLoS One 7*, e33342.

(106) Derman, A. I., Prinz, W. A., Belin, D., and Beckwith, J. (1993) Mutations that allow disulfide bond formation in the cytoplasm of Escherichia coli, *Science 262*, 1744-1747.

(107) Cleland, W. W. (1964) Dithiothreitol New Protective Reagent for Sh Groups, *Biochemistry 3*, 480-482.

(108) Jocelyn, P. C. (1987) Chemical-Reduction of Disulfides, *Method. Enzymol. 143*, 246-256.

(109) Burns, J. A., Butler, J. C., Moran, J., and Whitesides, G. M. (1991) Selective Reduction of Disulfides by Tris(2-Carboxyethyl)Phosphine, *J. Org. Chem. 56*, 2648-2650.

(110) Getz, E. B., Xiao, M., Chakrabarty, T., Cooke, R., and Selvin, P. R. (1999) A comparison between the sulfhydryl reductants tris(2-carboxyethyl)phosphine and dithiothreitol for use in protein biochemistry, *Anal. Biochem. 273*, 73-80.

(111) Fersht, A. (1999) *Structure and mechanism in protein science*, W. H. Freeman and Company, New York.

(112) Gerfen, G. J. (1999) EPR spectroscopy of B12-dependent enzymes". In: Chemistry and Biochemistry of B12, In *Chemistry and Biochemistry of B12* (Banerjee, R., Ed.), pp 165-195, John Wiley and Sons, New York.

(113) Zelder, O., Beatrix, B., Kroll, F., and Buckel, W. (1995) Coordination of a Histidine Residue of the Protein-Component-S to the Cobalt Atom in Coenzyme B-12-Dependent Glutamate Mutase from Clostridium-Cochlearium, *Febs Lett. 369*, 252-254.

(114) Padmakumar, R., Taoka, S., Padmakumar, R., and Banerjee, R. (1995) Coenzyme B12 Is Coordinated by Histidine and Not Dimethylbenzimidazole on Methylmalonyl-CoA Mutase, *J. Am. Chem. Soc. 117*, 7033-7034.

(115) Marsh, E. N. G., and Holloway, D. E. (1992) Cloning and Sequencing of Glutamate Mutase Component-S from Clostridium-Tetanomorphum - Homologies with Other Cobalamin-Dependent Enzymes, *Febs Lett. 310*, 167-170.

(116) Mancia, F., Keep, N. H., Nakagawa, A., Leadlay, P. F., McSweeney, S., Rasmussen, B., Bosecke, P., Diat, O., and Evans, P. R. (1996) How coenzyme B-12 radicals are generated: The crystal structure of methylmalonyl-coenzyme A mutase at 2 angstrom resolution, *Structure 4*, 339-350.

(117) Chowdhury, S., and Banerjee, R. (1999) Role of the dimethylbenzimidazole tail in the reaction catalyzed by coenzyme B-12-dependent methylmalonyl-CoA mutase, *Biochemistry 38*, 15287-15294.

(118) Yamanishi, M., Yamada, S., Muguruma, H., Murakami, Y., Tobimatsu, T., Ishida, A., Yamauchi, J., and Toraya, T. (1998) Evidence for axial coordination of 5,6-dimethylbenzimidazole to the cobalt atom of adenosylcobalamin bound to diol dehydratase, *Biochemistry 37*, 4799-4803.

(119) Lawrence, C. C., Gerfen, G. J., Samano, V., Nitsche, R., Robins, M. J., Retey, J., and Stubbe, J. (1999) Binding of Cob(II)alamin to the adenosylcobalamin-dependent

ribonucleotide reductase from Lactobacillus leichmannii. Identification of dimethylbenzimidazole as the axial ligand, *J. Biol. Chem. 274*, 7039-7042.

(120) Abend, A., Bandarian, V., Nitsche, R., Stupperich, E., Retey, J., and Reed, G. H. (1999) Ethanolamine ammonia-lyase has a "base-on" binding mode for coenzyme B(12), *Arch. Biochem. Biophys. 370*, 138-141.

(121) Robertson, W. D., Wang, M., and Warncke, K. (2011) Characterization of Protein Contributions to Cobalt-Carbon Bond Cleavage Catalysis in Adenosylcobalamin-Dependent Ethanolamine Ammonia-Lyase by using Photolysis in the Ternary Complex, *J. Am. Chem. Soc. 133*, 6968-6977.

(122) Babior, B. M. (1970) Mechanism of Action of Ethanolamine Ammonia-Lyase, a B12-Dependent Enzyme .7. Mechanism of Hydrogen Transfer, *J. Biol. Chem. 245*, 6125-6133.

(123) Lakowicz, J. R. (2006) *Principles of fluorescence spectroscopy*, Vol. 1, Springer.

(124) Callis, P. R. (1991) Molecular-Orbital Theory of the 1lb and 1la States of Indole, *J. Chem. Phys. 95*, 4230-4240.

(125) Eftink, M. R., Selvidge, L. A., Callis, P. R., and Rehms, A. A. (1990) Photophysics of Indole-Derivatives - Experimental Resolution of La and Lb Transitions and Comparison with Theory, *J. Phys. Chem. 94*, 3469-3479.

(126) Callis, P. R. (1997) L-1(a) and L-1(b) transitions of tryptophan: Applications of theory and experimental observations to fluorescence of proteins, *Method. Enymol. 278*, 113-150.

(127) Alcala, J. R., Gratton, E., and Prendergast, F. G. (1987) Fluorescence Lifetime Distributions in Proteins, *Biophys. J. 51*, 597-604.

(128) Burstein, E. A., Vedenkina, N. S., and Ivkova, M. N. (1973) Fluorescence and the Location of Tryptophan Residues in Protein Molecules, *Photochem. Photobiol. 18*, 263-279.

(129) Mertens, M. L., and Kagi, J. H. R. (1979) Graphical Correction Procedure for Inner Filter Effect in Fluorescence Quenching Titrations, *Anal. Biochem. 96*, 448-455.

(130) Steiner, R. F., and Kirby, E. P. (1969) Interaction of Ground and Excited States of Indole Derivatives with Electron Scavengers, *J. Phys. Chem. 73*, 4130-4135.

(131) Chen, Y., and Barkley, M. D. (1998) Toward understanding tryptophan fluorescence in proteins, *Biochemistry 37*, 9976-9982.

(132) Yuan, T., Weljie, A. M., and Vogel, H. J. (1998) Tryptophan fluorescence quenching by methionine and selenomethionine residues of calmodulin: orientation of peptide and protein binding, *Biochemistry 37*, 3187-3195.

(133) Hill, A. (1910) The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves, *J. Physiol. 40*.

(134) Weiss, J. N. (1997) The Hill equation revisited: uses and misuses, *Faseb J. 11*, 835-841.

(135) Todd, A. P., Cong, J., Levinthal, F., Levinthal, C., and Hubbell, W. L. (1989) Site-directed mutagenesis of colicin E1 provides specific attachment sites for spin labels whose spectra are sensitive to local conformation, *Proteins 6*, 294-305.

(136) Wang, M. (2009) Entropic Origin for Catalysis of Cobalt-Carbon Bond Cleavage in Coenzyme B12 (Adenosylcobalamin) in Ethanolamine Ammonia-Lyase, In *Physics*, Emory Univesity, Atlanta, GA.

(137) Stoll, S., and Schweiger, A. (2006) EasySpin, a comprehensive software package for spectral simulation and analysis in EPR, *J. Magn. Reson. 178*, 42-55.

(138) Aebersold, R., and Mann, M. (2003) Mass spectrometry-based proteomics, *Nature 422*, 198-207.

(139) Morrison, K. L., and Weiss, G. A. (2001) Combinatorial alanine-scanning, *Curr. Opin. Chem. Biol. 5*, 302-307.

(140) Yuan, Z., Zhao, J., and Wang, Z. X. (2003) Flexibility analysis of enzyme active sites by crystallographic temperature factors, *Protein Eng. 16*, 109-114.

(141) Nesmelov, Y. E., Gopinath, A., and Thomas, D. D. (2004) Aqueous sample in an EPR cavity: sensitivity considerations, *J. Magn. Reson. 167*, 138-146.

(142) Fung, L. W., and Johnson, M. E. (1983) Temperature dependence of spin-label intensity in solutions and its implication in spin-labeled erythrocyte membrane studies, *Biophys. J. 43*, 255-257.

(143) Malmberg, N. J., and Falke, J. J. (2005) Use of EPR power saturation toanalyze the membrane-docking geometries of peripheral proteins: A applications to C2 domains, *Annu. Rev. Bioph. Biom. 34*, 71-90.

(144) Altenbach, C., Greenhalgh, D. A., Khorana, H. G., and Hubbell, W. L. (1994) A Collision Gradient-Method to Determine the Immersion Depth of Nitroxides in Lipid Bilayers - Application to Spin-Labeled Mutants of Bacteriorhodopsin, *P. Natl. Acad. Sci. U.S.A. 91*, 1667-1671.

(145) Lin, Y., Nielsen, R., Murray, D., Hubbell, W. L., Mailer, C., Robinson, B. H., and Gelb, M. H. (1998) Docking phospholipase A(2) on membranes using electrostatic potential-modulated spin relaxation magnetic resonance, *Science 279*, 1925-1929.

(146) Haas, D. A., Mailer, C., and Robinson, B. H. (1993) Using Nitroxide Spin Labels - How to Obtain T(1e) from Continuous Wave Electron-Paramagnetic Resonance-Spectra at All Rotational Rates, *Biophys. J. 64*, 594-604.

(147) Altenbach, C., Flitsch, S. L., Khorana, H. G., and Hubbell, W. L. (1989) Structural Studies on Transmembrane Proteins .2. Spin Labeling of Bacteriorhodopsin Mutants at Unique Cysteines, *Biochemistry 28*, 7806-7812.

(148) Strohalm, M., Kodicek, M., and Pechar, M. (2003) Tryptophan modification by 2-hydroxy-5-nitrobenzyl bromide studied by MALDI-TOF mass spectrometry, *Biochem. Bioph. Res. Co. 312*, 811-816.

(149) Fenselau, C. (1997) MALDI MS and strategies for protein analysis, *Anal. Chem. 69*, 661A-665A.

(150) Sechi, S., and Chait, B. T. (1998) Modification of cysteine residues by alkylation. A tool in peptide mapping and protein identification, *Anal. Chem. 70*, 5150-5158.

(151) Chaurand, P., Luetzenkirchen, F., and Spengler, B. (1999) Peptide and protein identification by matrix-assisted laser desorption ionization (MALDI) and MALDI-post-source decay time-of-flight mass spectrometry, *J. Am. Soc. Mass Spectrom. 10*, 91-103.

(152) Wingfield, P. T. (2003) Overview of the purification of recombinant proteins produced in Escherichia coli, *Curr. Protoc. Protein Sci. Chapter 6*, Unit 6 1.

(153) Georgiou, G., and Valax, P. (1999) Isolating inclusion bodies from bacteria, *Method. Enzymol. 309*, 48-58.

(154) Wagner, S., Klepsch, M. M., Schlegel, S., Appel, A., Draheim, R., Tarry, M., Hogbom, M., van Wijk, K. J., Slotboom, D. J., Persson, J. O., and de Gier, J. W. (2008)

Tuning Escherichia coli for membrane protein overexpression, *P. Natl. Acad. Sci. U.S.A. 105*, 14371-14376.

(155) Widmann, M., and Christen, P. (2000) Comparison of folding rates of homologous prokaryotic and eukaryotic proteins, *J. Biol. Chem. 275*, 18619-18622.

(156) Hevehan, D. L., and De Bernardez Clark, E. (1997) Oxidative renaturation of lysozyme at high concentrations, *Biotechnol. Bioeng. 54*, 221-230.

(157) Xie, Y., and Wetlaufer, D. B. (1996) Control of aggregation in protein refolding: the temperature-leap tactic, *Protein Sci. 5*, 517-523.

(158) Wetlaufer, D. B., Branca, P. A., and Chen, G. X. (1987) The oxidative folding of proteins by disulfide plus thiol does not correlate with redox potential, *Protein Eng. 1*, 141-146.

(159) Tsumoto, K., Umetsu, M., Kumagai, I., Ejima, D., Philo, J. S., and Arakawa, T. (2004) Role of arginine in protein refolding, solubilization, and purification, *Biotechnol. Prog. 20*, 1301-1308.

(160) Nagel, Z. D., and Klinman, J. P. (2006) Tunneling and dynamics in enzymatic hydride transfer, *Chem. Rev. 106*, 3095-3118.

(161) Ball, P. (2004) Enzymes: by chance, or by design?, *Nature 431*, 396-397.

(162) Frey, P. A., and Hegeman, A. D. (2007) *Enzymatic reaction mechanisms*, Oxford University Press, USA.

(163) Bahnson, B. J., and Klinman, J. P. (1995) Hydrogen tunneling in enzyme catalysis, *Method. Enzymol. 249*, 373-397.

(164) Swain, C. G., Stivers, E. C., Reuwer, J. F., and Schaad, L. J. (1958) Use of Hydrogen Isotope Effects to Identify the Attacking Nucleophile in the Enolization of Ketones Catalyzed by Acetic Acid1-3, *J. Am. Chem. Soc. 80*, 5885-5893.

(165) Nagel, Z. D., and Klinman, J. P. (2009) A 21(st) century revisionist's view at a turning point in enzymology, *Nat. Chem. Biol. 5*, 543-550.

(166) Knapp, M. J., and Klinman, J. P. (2002) Environmentally coupled hydrogen tunneling - Linking catalysis to dynamics, *Eur J Biochem 269*, 3113-3121.

(167) Klinman, J. P. (2006) Linking protein structure and dynamics to catalysis: the role of hydrogen tunnelling, *Phil. Trans. Roy. Soc. B 361*, 1323-1331.

(168) Romesberg, F. E., and Schowen, R. L. (2004) Isotope effects and quantum tunneling in enzyme-catalyzed hydrogen transfer. Part I. The experimental basis, *Adv. Phys. Org. Chem. 39*, 27-77.

(169) Kohen, A., and Klinman, J. P. (1999) Hydrogen tunneling in biology, *Chem. Biol. 6*, R191-R198.

(170) Stojkovic, V., Perissinotti, L. L., Willmer, D., Benkovic, S. J., and Kohen, A. (2012) Effects of the donor-acceptor distance and dynamics on hydride tunneling in the dihydrofolate reductase catalyzed reaction, *J. Am. Chem. Soc. 134*, 1738-1745.

(171) Klinman, J. P. (2009) An integrated model for enzyme catalysis emerges from studies of hydrogen tunneling, *Chem. Phys. Lett. 471*, 179-193.

(172) Glowacki, D. R., Harvey, J. N., and Mulholland, A. J. (2012) Taking Ockham's razor to enzyme dynamics and catalysis, *Nat. Chem. 4*, 169-176.

(173) Kamerlin, S. C., and Warshel, A. (2010) At the dawn of the 21st century: Is dynamics the missing link for understanding enzyme catalysis?, *Proteins 78*, 1339-1375.

(174) Padmakumar, R., Padmakumar, R., and Banerjee, R. (1997) Evidence that cobalt-carbon bond homolysis is coupled to hydrogen atom abstraction from substrate in methylmalonyl-CoA mutase, *Biochemistry 36*, 3713-3718.

(175) Chowdhury, S., and Banerjee, R. (2000) Evidence for Quantum Mechanical Tunneling in the Coupled Cobalt−Carbon Bond Homolysis−Substrate Radical Generation Reaction Catalyzed by Methylmalonyl-CoA Mutase, *J. Am. Chem. Soc. 122*, 5417-5418.

(176) Marsh, E. N., and Ballou, D. P. (1998) Coupling of cobalt-carbon bond homolysis and hydrogen atom abstraction in adenosylcobalamin-dependent glutamate mutase, *Biochemistry 37*, 11864-11872.

(177) Essenberg, M. K., Frey, P. A., and Abeles, R. H. (1971) Studies on the mechanism of hydrogen transfer in the coenzyme B12 dependent dioldehydrase reaction II, *J. Am. Chem. Soc. 93*, 1242-1251.

(178) Doll, K. M., Bender, B. R., and Finke, R. G. (2003) The first experimental test of the hypothesis that enzymes have evolved to enhance hydrogen tunneling, *J. Am. Chem. Soc. 125*, 10877-10884.

(179) Warncke, K., Schmidt, J. C., and Ke, S. C. (1999) Identification of a rearranged-substrate, product radical intermediate and the contribution of a product radical trap in vitamin B-12 coenzyme-dependent ethanolamine deaminase catalysis, *J. Am. Chem. Soc. 121*, 10522-10528.

(180) Boas, J. F., Hicks, P. R., Pilbrow, J. R., and Smith, T. D. (1978) Interpretation of Electron-Spin Resonance-Spectra Due to Some B12-Dependent Enzyme-Reactions, *J. Chem. Soc., Faraday II 74*, 417-431.

(181) Ke, S. C. (2003) Spin-spin interaction in ethanolamine deaminase, *Biochim. Biophys. Acta 1620*, 267-272.

(182) Zhu, C. (2010) Enzyme on Ice: Kinetic and EPR Spectroscopic Characterization of the CoII-Substrate Radical Decay Reaction in Coenzyme B12-Dependent Ethanolamine Ammonia-Lyase, In *Physics*, p 157, Emory University, Atlanta, GA.

(183) Gani, D., Wallis, O. C., and Young, D. W. (1983) Stereochemistry of the Rearrangement of 2-Aminoethanol by Ethanolamine Ammonia-Lyase, *Eur. J. Biochem. 136*, 303-311.

(184) Moore, J., and Pearson, R. (1981) G. Kinetics and Mechanism, Wiley: New York.

(185) Motulsky, H., and Christopoulos, A. (2004) *Fitting models to biological data using linear and nonlinear regression: a practical guide to curve fitting*, Oxford University Press, USA.

(186) Licht, S. S., Lawrence, C. C., and Stubbe, J. (1999) Thermodynamic and kinetic studies on carbon-cobalt bond homolysis by ribonucleoside triphosphate reductase: the importance of entropy in catalysis, *Biochemistry 38*, 1234-1242.

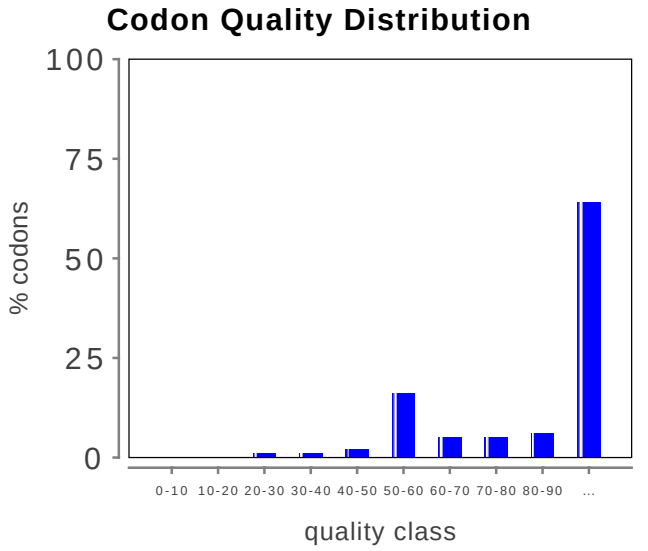
Appendix

Your GeneOptimizer® Assisted
Analysis Report

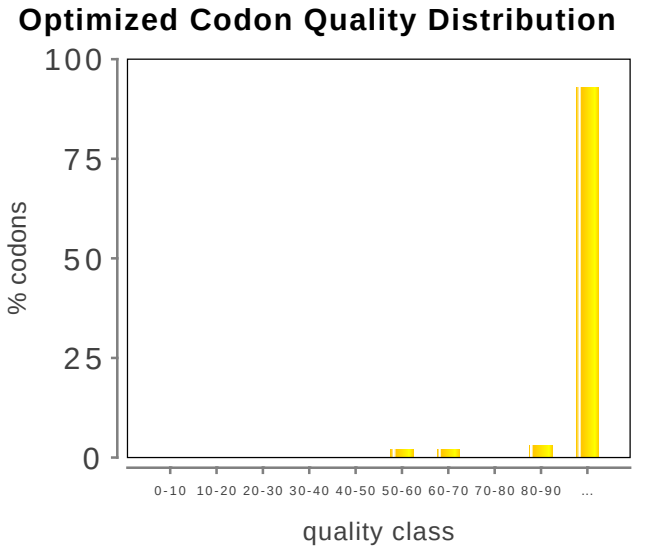
Project: Project on Mar 1, 2012
Analysis Created:03/02/2012 (10:21 AM)

Gene Name	EAL-11CF
Host Species	Escherichia coli
ORF	7-1371 [ATG...TAA] 1387-2283 [ATG...TAA]
Enzymes	1-6 BamHI [GGATCC] 2284-2289 HindIII [AAGCTT]
Protected Sequences	1372-1386 Custom [CGC...ACC]
Excluded Enzymes	

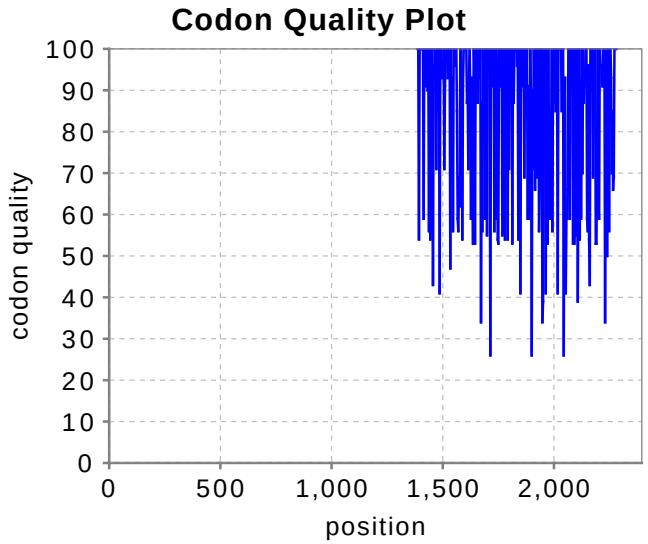
Non-Optimized

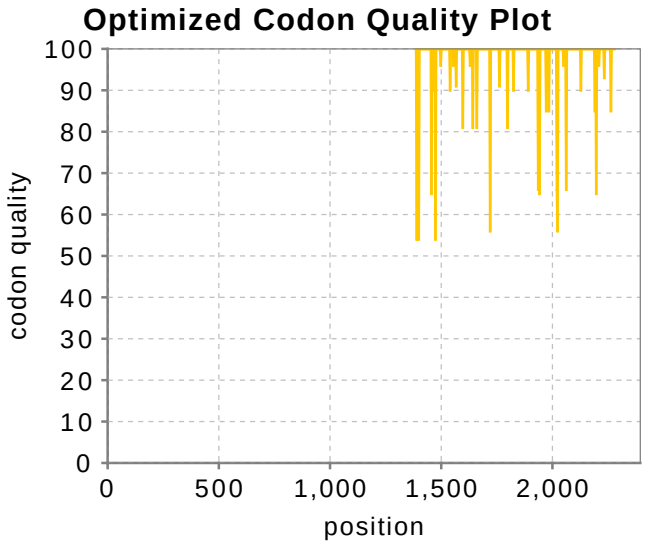


Optimized with GeneOptimizer®

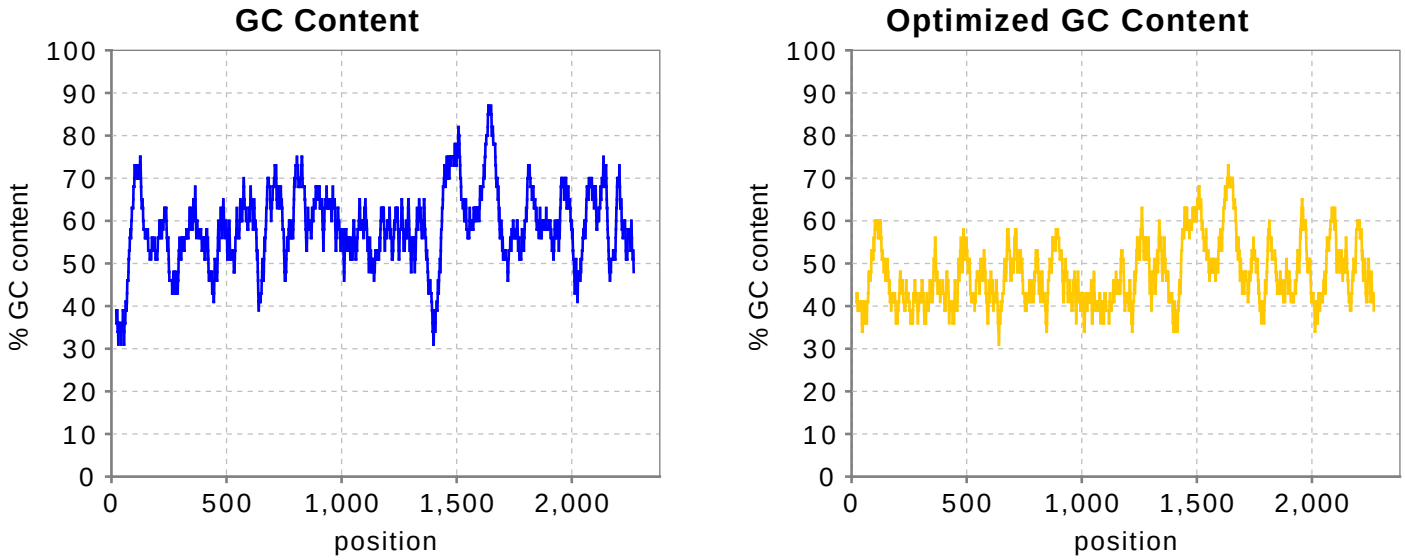


The histograms show the percentage of sequence codons which fall into a certain quality class. The quality value of the most frequently used codon for a given amino acid in the desired expression system is set to 100, the remaining codons are scaled accordingly (see also Sharp, P.M., Li, W.H., Nucleic Acids Res. 15 (3),1987).





The plots show the quality of the used codon at the indicated codon position.



The plots show the GC content in a 40 bp window centered at the indicated nucleotide position.

Your Optimized Sequence

```
        M  K  L  K  T  T  L  F  G  N  V  Y  Q  F  K  D  V  K  E  V  L
1.      GGATCCATGAAACTGAAAACCACCCTGTTTGGTAACGTGTATCAGTTCAAAGATGTGAAAGAAGTTC
          A  K  A  N  E  L  R  S  G  D  V  L  A  G  V  A  A  A  S  S  Q  E
68.     TGGCCAAAGCAAATGAACTGCGTAGCGGTGATGTTCTGGCAGGCGTTGCAGCAGCAAGCAGCCAAGA
          R  V  A  A  K  Q  V  L  S  E  M  T  V  A  D  I  R  N  N  P  V  I
135.    ACGTGTTGCAGCAAAAACAGGTTCTGAGCGAAATGACCGTTGCAGATATTCGTAATAATCCGGTGATT
          A  Y  E  E  D  A  V  T  R  L  I  Q  D  D  V  N  E  T  A  Y  N  R  I
202.    GCCTATGAAGAAGATGCAGTTACCCGTCTGATTCAGGATGATGTTAATGAAACCGCCTATAACCGCA
          K  N  W  S  I  S  E  L  R  E  Y  V  L  S  D  E  T  S  V  D  D  I
269.    TCAAAAATTGGAGCATTAGCGAACTGCGTGAATATGTTCTGAGTGATGAAACCAGCGTTGATGATAT
          A  F  T  R  K  G  L  T  S  E  V  V  A  A  V  A  K  I  V  S  N  A
336.    TGCATTTACCCGTAAAGGTCTGACCAGCGAAGTTGTTGCAGCCGTTGCAAAAATTGTTAGCAATGCC
          D  L  I  Y  G  G  K  K  M  P  V  I  K  K  A  N  T  T  I  G  I  P  G
403.    GATCTGATCTACGGTGGCAAAAAAATGCCGGTTATCAAAAAAGCCAATACCACCATTGGTATTCCGG
          T  F  S  A  R  L  Q  P  N  D  T  R  D  D  V  Q  S  I  A  A  Q  I
470.    GTACATTTAGCGCACGTCTGCAGCCGAATGATACCCGTGATGATGTTCAGAGCATTGCAGCACAGAT
          Y  E  G  L  S  F  G  A  G  D  A  V  I  G  V  N  P  V  T  D  D  V
537.    TTATGAAGGTCTGAGTTTTGGTGCCGGTGATGCCGTTATTGGTGTTAATCCGGTTACCGATGATGTT
          E  N  L  T  R  V  L  D  T  V  Y  G  V  I  D  K  F  N  I  P  T  Q  G
604.    GAAAATCTGACCCGTGTTCTGGATACCGTTTATGGTGTTATCGACAAATTTAACATTCCGACCCAGG
          V  V  L  A  H  V  T  T  Q  I  E  A  I  R  R  G  A  P  G  G  L  I
671.    GTGTTGTGCTGGCACATGTTACCACCCAGATTGAAGCAATTCGTCGTGGTGCACCGGGTGGTCTGAT
          F  Q  S  I  S  G  S  E  K  G  L  K  E  F  G  V  E  L  A  M  L  D
738.    TTTCAGAGTATTAGCGGTAGCGAAAAAGGCCTGAAAGAATTTGGTGTTGAACTGGCAATGCTGGAT
          E  A  R  A  V  G  A  E  F  N  R  I  A  G  E  N  A  L  Y  F  E  T  G
805.    GAAGCACGTGCAGTTGGTGCAGAATTTAATCGTATTGCCGGTGAAAATGCCCTGTATTTTGAAACCG
          Q  G  S  A  L  S  A  G  A  N  F  G  A  D  Q  V  T  M  E  A  R  N
872.    GTCAGGGTAGCGCACTGAGTGCCGGTGCAAATTTTGGTGCGGATCAGGTTACCATGGAAGCACGTAA
          Y  G  L  A  R  H  Y  D  P  F  L  V  N  T  V  V  G  F  I  G  P  E
939.    TTATGGTCTGGCACGTCATTATGATCCGTTTCTGGTTAATACCGTGGTGGGTTTTATTGGTCCGGAA
```

```
      Y   L   Y   N   D   R   Q   I   I   R   A   G   L   E   D   H   F   M   G   K   L   S   G
1006. TATCTGTATAATGATCGCCAGATTATTCGTGCAGGTCTGGAAGATCATTTTATGGGTAAACTGAGCG

        I   S   M   G   C   D   C   C   Y   T   N   H   A   D   A   D   Q   N   L   N   E   N
1073. GTATTAGCATGGGTTGTGATTGTTGTTATACCAATCATGCAGATGCCGATCAGAATCTGAATGAAAA

      L   M   I   L   L   A   T   A   G   C   N   Y   I   M   G   M   P   L   G   D   D   I
1140. CCTGATGATTCTGCTGGCAACCGCAGGTTGTAACTATATCATGGGTATGCCGCTGGGAGATGATATT

      M   L   N   Y   Q   T   T   A   F   H   D   T   A   T   V   R   Q   L   L   N   L   R   P
1207. ATGCTGAATTATCAGACAACCGCCTTTCATGATACCGCAACCGTTCGTCAGCTGCTGAATCTGCGTC

        S   P   E   F   E   R   W   L   E   T   M   G   I   M   A   N   G   R   L   T   K   R
1274. CGAGTCCGGAATTTGAACGTTGGCTGGAAACCATGGGTATTATGGCAAATGGTCGTCTGACCAAACG

      A   G   D   P   S   L   F   F   *   *                   M   D   Q   K   Q   I   E
1341. TGCGGGTGATCCGAGCCTGTTTTTCTAATAACGCGGGGATAACACCATGGACCAGAAACAAATTGAA

      E   I   V   R   S   V   M   A   S   M   G   Q   D   V   P   Q   P   A   A   P   S   T   Q
1408. GAAATTGTTCGTAGCGTTATGGCAAGCATGGGTCAGGATGTTCCGCAGCCTGCAGCACCGAGTACCC

        E   G   A   K   P   Q   A   A   A   P   T   V   T   E   S   A   A   L   D   L   G   S
1475. AAGAAGGTGCAAAACCGCAGGCAGCCGCACCGACCGTTACCGAAAGCGCAGCACTGGATCTGGGTAG

      A   E   A   K   A   W   I   G   V   E   N   P   H   R   A   D   V   L   T   E   L   R
1542. TGCAGAAGCAAAAGCCTGGATTGGTGTGGAAAATCCGCATCGTGCAGATGTTCTGACAGAACTGCGT

      R   S   T   A   A   R   V   A   T   G   R   A   G   P   R   P   R   T   Q   A   L   L   R
1609. CGTAGCACCGCAGCACGTGTTGCCACCGGTCGTGCTGGTCCGCGTCCGCGTACACAGGCACTGCTGC

        F   L   A   D   H   S   R   S   K   D   T   V   L   K   E   V   P   E   E   W   V   K
1676. GTTTTCTGGCAGATCATAGCCGTAGCAAAGATACCGTTCTGAAAGAGGTTCCGGAAGAATGGGTTAA

      A   Q   G   L   L   E   V   R   S   E   I   S   D   K   N   L   Y   L   T   R   P   D
1743. AGCACAGGGTCTGCTGGAAGTGCGTAGCGAAATTAGCGATAAAAATCTGTATCTGACACGTCCGGAT

      M   G   R   R   L   S   P   E   A   I   D   A   L   K   S   Q   A   V   M   N   P   D   V
1810. ATGGGTCGTCGTCTGTCACCGGAAGCAATTGATGCACTGAAAAGCCAGGCAGTTATGAATCCGGATG

        Q   V   V   V   S   D   G   L   S   T   D   A   I   T   A   N   Y   E   E   I   L   P
1877. TTCAGGTTGTTGTTAGTGATGGTCTGAGCACCGATGCAATTACCGCAAATTATGAAGAAATCCTGCC

      P   L   L   A   G   L   K   Q   A   G   L   N   V   G   T   P   F   F   V   R   Y   G
1944. TCCGCTGCTGGCAGGTCTGAAACAGGCAGGCCTGAATGTTGGCACCCCGTTTTTTGTTCGTTATGGT

      R   V   K   I   E   D   Q   I   G   E   I   L   G   A   K   V   V   I   L   L   V   G   E
2011. CGTGTTAAAATCGAGGATCAGATTGGTGAAATTCTGGGGTGCCAAAGTTGTTATCCTGCTGGTTGGTG

        R   P   G   L   G   Q   S   E   S   L   S   L   Y   A   V   Y   S   P   R   V   A   T
2078. AACGTCCGGGTCTGGGTCAGAGCGAAAGCCTGAGCCTGTATGCAGTTTATAGTCCGCGTGTTGCAAC

      T   V   E   A   D   R   T   C   I   S   N   I   H   Q   G   G   T   P   P   V   E   A
2145. CACCGTTGAAGCAGATCGTACCTGTATTAGCAATATTCATCAGGGTGGCACCCCTCCGGTTGAAGCC

      A   A   V   I   V   D   L   A   K   R   M   L   E   Q   K   A   S   G   I   N   M   T   R
2212. GCAGCAGTTATTGTTGATCTGGCGAAACGTATGCTGGAACAGAAAGCAAGCGGCATTAATATGACCC

        *
2279. GTTAAAAGCTT
```