EXPLORATION OF ACTIVE SITE RESIDUES OF *ESCHERICHIA*
COLI CYSTATHIONINE β -LYASE BY SITE DIRECTED
MUTAGENESIS

by

Joanna Orzechowska

A thesis submitted to the Faculty of Graduate and Postdoctoral
Affairs in partial fulfillment of the requirements
for the degree of

Master of Science

in

Biology

Carleton University
Ottawa, Ontario

© 2018

Joanna Orzechowska

Library and Archives
Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

ISBN: 978-0-494-87817-0

Our file Notre référence

ISBN: 978-0-494-87817-0

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

ABSTRACT

Cystathionine β -lyase (CBL) is the second enzyme of the bacterial transsulfuration pathway, which catalyzes the reaction of L-cystathionine to pyruvate, ammonia and L-homocysteine. CBL is dependant on the cofactor pyridoxal-5'-phosphate (PLP), which forms an internal aldimine with the catalytic lysine (eCBL-K210) in its inactive state. This work focuses on residues having a direct impact on the positioning of PLP and/or K210. Steady-state kinetics data from five site-directed D185 variants suggests that eCBL may be the only one of fold type I enzymes that performs the β -elimination reaction using E1 mechanism. The change in positioning of PLP caused by the substitutions of A207 and T209 causes enzymatic instability and inactivity, which does not occur with the Y211F variant. The R262K variant shows positive cooperation; a result that was not observed in our previously constructed eCBL variants. Finally, the importance of maintaining the aromatic properties of W340 was highlighted through its substitution.

ACKNOWLEDGEMENTS

Many thanks are due to so many people, in part for helping make this thesis happen, but also for keeping me sane during it all. My deepest gratitude to my supervisor, Professor Susan Aitken. I am sincerely grateful for her continuous support and patience. There is no better teacher and mentor. I cannot express how grateful I am to have had the opportunity to work with her.

I would like to thank my committee members, Dr. Amanda MacFarlane, and Dr. Cory Harris for the critical feedback that helped me to finish this thesis.

Much credit to Amanda Leclair, without whom the figures in this thesis would look, at best, like a child's drawings. She and Darien Klein were involved in the purification of several of the variants described in here, and for their help, I am appreciative. To Micaela Belleperche, whose thirst for knowledge constantly inspires me.

Furthermore, I am thankful to share the office with people who provide constant support and inspiration. Katie, Maryam, Narges, Natalia, Rowida, Sarah, Shana, Taylor, Dan, Duale, Hemanta, Houman, Matt and Ryan. Thank you all for having and continuing to provide me with a healthy mix of motivation and distraction.

Then there's the person who motivated me to face my challenges and pushed me forward every day, Jay Sethi. Finally, biggest thanks to my parents are necessary. Mamo, stalas i ciagle stoisz za mna murem, choc nie zawsze bylo latwo. Tato, wspierales i ciagle wspierasz kazda moja decyzje, choc niektore sa kosmicze.

DEDICATION

Joanna

Jasvardan

Miroslaw

Thank you for letting me make my own mistakes.

Thank you for helping me.

TABLE OF CONTENTS

ABSTRACT.....	iii
ACKNOWLEDGEMENTS.....	iv
DEDICATION.....	v
TABLE OF CONTENTS.....	vi
LIST OF TABLES.....	viii
LIST OF FIGURES	ix
ABBREVIATIONS	xii
1. INTRODUCTION	1
1.1 Enzymes.....	1
1.2 Importance of enzyme studies	1
1.3 Pyridoxal 5'-phosphate	3
1.3.1 Mechanism.....	5
1.3.2 Diversity of PLP dependent enzymes	5
1.4 Metabolism of sulfur containing amino acids in bacteria	7
1.5 Transsulfuration and reverse transsulfuration pathways.....	12
1.5. Cystathionine β -lyase.....	15
1.5.1 Structure.....	15
1.5.2 Mechanism.....	19
1.5.3 Active site layout	21
1.6 Project objectives	26
2. METHODS	28

2.1 Reagents	28
2.2 Construction of site-directed mutants	28
2.3 Expression and purification of wild-type and site-directed mutants of eCBL.....	31
2.4 Steady-state kinetics.....	32
2.4 Affinity chromatography under denaturing conditions.....	35
2.5 PyMOL structure analysis.....	35
3. RESULTS	37
3.1 Steady state kinetic analysis.....	42
4. DISCUSSION	48
4.1. D185 variants	49
4.2. The D185-T187-A207 hydrogen-bonding network linking the PLP pyridinium nitrogen and the backbone of K210	58
4.3. Residues flanking K210	68
4.4. eCBL residues W340 and L310	75
4.5. Conclusion	78
REFERENCES	80

LIST OF TABLES

Table 2.1.	The oligonucleotide primers employed for construction of the eCBL site-directed mutants.	30
Table 3.1.	Yields of purified eCBL variants and the wild-type enzyme	37
Table 3.2.	Kinetic parameters of wild-type eCBL and site-directed variants.	42

LIST OF FIGURES

Figure 1.1.	PLP catalyzes a range of transformations of amino acid substrates.	4
Figure 1.2.	The transsulfuration pathway of plants and bacteria.	9
Figure 1.3.	Methionine biosynthesis pathways.	11
Figure 1.4.	Cartoon representation of an overlay of the structures of eCBL, in complex with N-Hydrazinocarbonylmethyl-2-Nitro-Benzamide and eCGS.	14
Figure 1.5.	The structure of the eCBL homotetramer in complex with N-Hydrazinocarbonylmethyl-2-Nitro-Benzamide inhibitor.	16
Figure 1.6.	Cartoon representation of the eCBL monomer in complex with N-Hydrazinocarbonylmethyl-2-Nitro-Benzamide inhibitor.	17
Figure 1.7.	The structures of eCBL substrate and two of the inhibitors used in presented crystal structures.	18
Figure 1.8.	The eCBL mechanism proposed by Clausen et al. (1997).	20
Figure 1.9.	Schematic of the eCBL active site showing the interactions of targeted D185, T187, A207, T209, Y211, R262, L310 and W340 residues and the PLP cofactor.	23
Figure 1.10.	PyMOL visualization of the packing interactions between W340 and L310.	25
Figure 2.1.	The continuous, DTNB-based assay used to measure the L-Cth hydrolysis activity of eCBL.	33
Figure 3.1.	SDS-PAGE gels of fractions from the purification of eCBL-W340L and wild-type eCBL.	38

Figure 3.2.	SDS-PAGE gel comparing 2 µg of wild type eCBL and purified site-directed variants.	39
Figure 3.3.	Affinity chromatography, under denaturing conditions, of a purified sample of the eCBL-A207T variant.	40
Figure 3.4.	The effect of L-cystathionine concentration on the activity of wild-type eCBL and the D185N, D185S, D185T, D185L and D185V site-directed variants.	43
Figure 3.5.	The effect of L-cystathionine concentration on the activity of wild-type eCBL and the T187S variant.	44
Figure 3.6.	The effect of L-cystathionine concentration on the activity of the eCBL R262K variant and wild type eCBL.	46
Figure 4.1.	Comparison of the interactions of D185 and the asparagine, threonine and serine substitutions, with the pyridinium nitrogen of the cofactor	51
Figure 4.2.	Comparison of the interactions of D185, and the leucine and valine substitutions, with the pyridinium nitrogen of the cofactor.	52
Figure 4.3.	Schematic representation of hydrogen-bonding network that bridges the pyridinium nitrogen of PLP and the backbone of K210, the catalytic base.	59
Figure 4.4.	PyMOL representation of the van der Waals overlap between the PLP cofactor and the A207 substitutions.	63
Figure 4.5.	The eCBL sequence with the possible protease cleavage sites.	67
Figure 4.6.	Comparison of PyMOL models of T209V and T209N in the context of the eCBL-N-hydrazinocarbonylmethyl-2-Nitro-Benzamide complex, in which the PLP cofactor and N-hydrazinocarbonylmethyl-2-Nitro-Benzamide inhibitor are shown in Schiff base linkage and unliganded eCBL structure, where the PLP cofactor is bound to K210.	70
Figure 4.7.	The hydrogen bonding interactions of Y211 and R262 and R262K and R262L variants.	72

Figure 4.8. The relative positions of K210, S339 and W340 in the structure of the eCBL-AVG complex. 77

ABBREVIATIONS

AAT	Aspartate aminotransferase
AVG	L-aminoethoxyvinylglycine
CBL	Cystathionine β -lyase
CBS	Cystathionine β -synthase
CGL	Cystathionine γ -lyase
CGS	Cystathionine γ -synthase
deazaPLP	1-deazapyridoxal 5'-phosphate
DNA	Deoxyribonucleic acid
DTNB	5,5'-dithiobis-(2-nitrobenzoic acid)
<i>E.coli</i>	<i>Escherichia coli</i>
eCBL	Cystathionine β -lyase from <i>Escherichia coli</i>
eCGS	Cystathionine γ -synthase from <i>Escherichia coli</i>
EDTA	Ethylenediaminetetraacetic acid
GSH	Glutathione
IPTG	Isopropyl β -D-1-thiogalactopyranoside
LB	Luria broth
L-Cth	L-cystathionine

L-Hcys	L-homocysteine
NaCl	Sodium chloride
Ni-NTA	Ni-nitrilo triacetic acid
OASS	<i>O</i> -acetylserine sulfhydrylase
OE-PCR	Overlap-extension polymerase chain reaction
PCR	Polymerase chain reaction
PDB	Protein databank
PLP	Pyridoxal 5'-phosphate
SAH	<i>S</i> -adenosylhomocysteine
SAM	<i>S</i> -adenosylmethionine
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
TNB⁻	2-nitro-5-mercapto-benzoic acid anion
Tris	Tris-[hydroxymethyl]aminomethane
yCGL	Cystathionine γ -lyase from <i>Saccharomyces cerevisiae</i>

1. INTRODUCTION

1.1 Enzymes

Enzymes are macromolecules that catalyze the metabolic reactions crucial for cell metabolism. Most enzymes are proteins, commonly composed of the twenty standard L-amino acids, with some ribonucleic acids possessing a catalytic role. As biological catalysts, enzymes function by binding their substrate(s) in a manner that lowers the activation energy of the desired reaction (Berg, Tymoczko and Stryer, 2008). While many enzymes rely only on the amino acid residues of the active site for catalysis, some reactions require the additional participation of a cofactor, which is a non-protein part of the enzyme, such as metal ion(s) or small organic molecules (Berg, Tymoczko and Stryer, 2008).

Understanding the relationship between an enzyme's structure and its function provides the basis for applications such as inhibitor design and modification of enzyme properties (e.g. thermal stability) or substrate or reaction specificity. While techniques based on random mutagenesis or DNA shuffling are commonly used to change enzyme properties, adjustment of specificity via rational redesign is feasible, but requires a detailed knowledge of the layout of the active site and the specific role of each active-site residue.

1.2 Importance of enzyme studies

Naturally occurring enzymes have long been used in the production of food (e.g. bread), drinks (e.g. wine and beer) and textiles (e.g. cloth finishing, leather processing). The use of enzymes, in place of standard chemical or physical industrial methods, can increase

yield and efficiency, as well as reduce costs and requirements for energy and disposal of chemical waste (Kirk, Borchert and Fuglsang, 2002). Naturally occurring enzymes are limited in stability and the range of reactions they can catalyze, as natural selection has adapted them to the specific needs of living cells. Over time, protein engineering methods have been developed to allow, not only the use of existing enzymes for industrial applications, but also the modification of enzymes to change properties, such as thermal/pH/solvent stability or alter specificity. This has expanded the range of applications, such as paper making and cosmetics, for which enzymes are available (Woodyer, Chen and Zhao, 2004).

Protein engineering methods range from targeted (i.e. rational redesign) to random (e.g. random mutagenesis and DNA shuffling) (Stemmer, 1994; Cramer et al., 1998; Tobin, Gustafsson and Huisman, 2000). Rational redesign, via site-directed mutagenesis, relies on knowledge including an enzyme's sequence and structure, conservation or differences in residues or features with related enzymes, and a detailed biochemical characterization of the active site. Since Perutz and Kendrew solved the first protein structure, that of myoglobin, in 1958, X-ray crystallography remains the primary way to determine protein structure at the atomic scale (Kendrew et al., 1958). As of March 2018, there are almost 136,000 three-dimensional structures in the Research Collaboratory for Structural Bioinformatics Protein Data Bank and recent advances in bioinformatic tools now allow comparison of structures to gain insight into features that are shared by or differ between related enzymes (Berman, 2000; Wang et al., 2013; Dong et al., 2017).

The knowledge gained from studies characterizing the role of active-site residues contribute to building a working model of a given enzyme that can provide the basis for

hypotheses guiding rational redesign experiments. An interesting set of enzymes for structure-function reaction and specificity studies are those dependent on the pyridoxal-5'-phosphate (PLP) cofactor.

1.3 Pyridoxal 5'-phosphate

Pyridoxal 5'-phosphate (PLP), a derivative of vitamin B₆ that was first identified by Gyorgy, is one of the most catalytically versatile enzyme cofactors (1934). This is illustrated by the more than 120 distinct reactions catalyzed by PLP-dependent enzymes (Percudani and Peracchi, 2003). The range of transformations of amino acid substrates catalyzed by enzymes relying on this cofactor includes decarboxylation, transamination, racemisation and side-chain elimination or replacement (Figure 1.1; Percudani and Peracchi, 2003).

As described by Limbach et al. (2011), the ability of PLP to catalyze a range of reactions can be explained by the interaction of active-site residues with the heterocyclic pyridinium ring and its phosphate and hydroxyl functional groups (Metzler and Snell, 1954). These interactions varyingly influence the ability of the cofactor to facilitate the formation of a carbanion on the α -carbon of the amino acid substrate and to stabilize this reaction intermediate via delocalization in the aromatic ring (Bach, Canepa and Glukhovtsev, 1999; Toney, 2001; Eliot and Kirsch, 2004). As such, the specific interactions between an enzyme's active-site residues and the cofactor regulate the properties of the latter. The particular layout of the active site in each of the PLP-dependent enzymes is responsible for the positioning of the cofactor, modulating its properties, and enforcing substrate

specificity to favor the desired reaction (e.g. β -elimination of cystathionine instead of transamination of aspartate).

Figure 1.1. PLP catalyzes a range of transformations of amino acid substrates.

1.3.1 Mechanism

In the resting state of PLP-dependent enzymes, a Schiff base linkage is formed between the aldehyde group of the cofactor and the ϵ -amino group of an active-site lysine (Figure 1.1). The initial step of the reaction, transaldimination, is common to all PLP-dependent enzymes catalyzing transformations of amino acids. It involves the attack of the unprotonated amino group of the substrate on the protonated Schiff base of the internal aldimine, and release of the ϵ -amino group of an active-site lysine, the catalytic base, to form the external aldimine of the substrate (Christen and Mehta, 2001; Toney, 2005). In the external aldimine the α -amino group of the substrate replaces the ϵ -amino group of the catalytic base in Schiff base linkage with the cofactor. The subsequent reaction is initiated by abstraction of the carboxylate group, the side chain or the proton attached to the α -carbon ($C\alpha$) of the amino acid substrate. The presence of the pyridine ring causes PLP to act as an electrophilic catalyst, facilitating the removal of electrons from the substrate, thus weakening the target bond enough to be broken and stabilizing the resulting carbanion (Metzler and Snell, 1954). The identity of the abstracted group is a key determinant of reaction specificity and is determined by its position relative to the aromatic ring of the cofactor (Dunathan, 1966; Toney, 2005). This is regulated by the amino acids lining the active site, as well as the surrounding environmental factors, such as pH (Metzler and Snell, 1954).

1.3.2 Diversity of PLP dependent enzymes

The diversity of the reactions catalyzed by PLP is not reflected in a similar variation in the overall protein structure of PLP-dependent enzymes. Jansonius et al. (1998) proposed

classification of PLP-dependent enzymes catalyzing transformations of amino acid substrates into four different fold types based on their tertiary structure and the active site organization. This classification excluded glycogen phosphorylase, which employs the phosphate group, rather than the heterocyclic ring of PLP, in catalysis (Jansonius, 1998).

The four fold types are as follows:

- Fold-type I – a large family of enzymes, illustrated by the extensively studied aspartate aminotransferase, that includes aminotransferases, decarboxylases, aminomutases and β - and γ -lyases. Enzymes of this group may be homodimeric or homotetrameric, with a pair of active sites situated at the subunit interface of each catalytic dimer (Percudani and Peracchi, 2003).
- Fold type II - enzymes that commonly catalyze β -elimination or replacement reactions, such as tryptophan synthase, and possess a regulatory domain or subunit (Percudani and Peracchi, 2003). In contrast with fold-type I, the active-sites of fold-type II enzymes are comprised of residues from a single subunit.
- Fold type III – a small family, also known as the alanine racemase group, characterized by an active site comprised of residues from two monomers (Jansonius, 1998; Schneider, Käck and Lindqvist, 2000).
- Fold type IV – referred to as the D-amino acid aminotransferase family, they possess an active site that is the mirror image of aminotransferases of fold-type I (Schneider, Käck and Lindqvist, 2000).

It has been proposed that fold-type I, which is the largest and most catalytically diverse of the fold types, was the first one to evolve (Christen and Mehta, 2001). The enzymes of this family provide the opportunity to explore the structure-function relationships that underlie

enzyme activity and the regulation of substrate and reaction specificity. Within the large group of structurally similar enzymes encompassed by fold-type I, the γ -subfamily is a smaller set. The enzymes of this subfamily are closely related in structure, share several conserved active-site residues, and catalyze distinct side-chain rearrangements of sulfur-containing amino acids. As such, they provide an ideal model system to characterize and compare the roles of active-site residues and to investigate determinants of substrate and reaction specificity. Our lab has investigated the active sites of three of the enzymes of this family: cystathionine β -lyase, cystathionine γ -lyase and cystathionine γ -synthase (Farsi et al., 2009; Lodha et al., 2010; Lodha and Aitken, 2011; Jaworski et al., 2012; Manders et al., 2013; Jaworski and Aitken, 2013; Hopwood, Ahmed and Aitken, 2014). Building on this work, additional studies exploring the roles of both conserved and distinct active-site residues in the context of these enzymes are required to further our understanding of the structure-function relationships underlying their specificity. This information will provide the framework for future studies aiming to engineer enzymes dependent on the catalytically versatile PLP-cofactor for a range of applications.

1.4 Metabolism of sulfur containing amino acids in bacteria

Among the amino acids that form proteins, there are only two that contain sulfur: cysteine and methionine. The transsulfuration pathway of bacteria and plants and reverse transsulfuration pathway of fungi and animals interconvert cysteine and homocysteine, the immediate precursor of methionine (Berg, Tymoczko and Stryer, 2008). In proteins cysteine and methionine play important roles based on their physical and chemical properties. Cysteine has a free thiol group that is prone to oxidation. Two of the cysteine

side chains can create disulphide bridges that can ensure the proper protein folding and stabilize their structure (Gilbert, 1995; Bigelow and Squier, 2005). Cysteine as a part of glutathione (GSH) tripeptide acts as an electron donor, regulating the redox state in cells and protecting against the oxidative stress (Meister, 1988). *S*-adenosylmethionine (SAM), derived from methionine, performs an important role in the cell metabolism as a supplier of methyl groups (Hoshi and Heinemann, 2001). Given this central role in cellular metabolism, it is curious that the animals do not possess the ability to synthesize methionine *de novo* and must to rely on bacteria and plant sources via diet (Ravanel et al., 1998).

It is important to consider enzymes within the context in which they function, as cellular metabolism comprises a vast network of pathways and enzymes. In *Escherichia coli*, cystathionine γ -gamma synthase (eCGS) catalyzes the condensation of *O*-succinyl-L-homoserine and L-cysteine, releasing succinate and forming L-cystathionine, the substrate of cystathionine β -lyase (CBL) (Figure 1.2). The products of L-cystathionine hydrolysis by *E. coli* CBL (eCBL) are ammonia, pyruvate and L-homocysteine, which is converted to L-methionine (Ravanel et al., 1998).

There are two highly conserved pathways for the biosynthesis of L-methionine from L-homocysteine. In the methH pathway, also known as the cobalamin dependant route, the methyl group of the cobalamin-dependant methionine synthase's cobalamin cofactor, obtained from 5-methyl-tetrahydrofolate, is transferred to L-homocysteine (Koutmos et al., 2009). While this is the more efficient route, bacteria have developed an additional pathway (González et al., 1996). The metE pathway, otherwise known as cobalamin- independent route, has a 50-fold lower turnover rate, but allows bacteria to synthesize L-methionine in

the absence of cobalamin, also known as vitamin B12 (Whitfield, Steers and Weissbach, 1970; Drennan, Huang and Drummond, 1994).

Figure 1.2. The transsulfuration pathway of plants and bacteria, comprising cystathionine γ -synthase (CGS) and cystathionine β -lyase (CBL), and cystathionine β -synthase (CBS) and cystathionine γ -lyase (CGL), of the animal and fungal reverse transsulfuration pathway, interconvert cysteine and homocysteine, the immediate precursor of methionine.

Bacteria can bypass the transsulfuration pathway by performing a direct sulfurylation of L-homoserine, activated by either succinyl- or acyl- groups with hydrogen sulfide by *O*-acetylhomoserine thiolases (Figure 1.3; Flavin and Slaughter, 1967). *O*-acetyl-L-homoserine thiolases can also use methane-thiol to bypass L-homocysteine in the direct production of methionine in the organisms in which it is a by-product of the metabolism. This reaction has so far been proven to exist in one species of bacterial plankton and in common yeast and therefore may not be relevant for *E. coli* (Yamagata et al., 1984 Kiene et al., 1999). L-homocysteine is also derived from *S*-ribosylhomocysteine in the recycling of *S*-adenosylhomocysteine (SAH), derived from methyl group donation by SAM (Taylor, 1982; Della Ragione et al., 1985).

Figure 1.3. Methionine biosynthesis pathways. Genes encode: *metB*, cystathionine γ -synthase; *metC*, cystathionine β -lyase; *metE*, cobalamin-independent methionine synthase; *metI*, *O*-acetylhomoserine sulphydrylase; *metH*, cobalamin-dependant methionine synthase; *metK*, *S*-adenosylmethionine synthetase and *mtnA*, methylthioadenosine nucleosidase (Auger, Danchin and Martin-Verstraete, 2002).

1.5 Transsulfuration and reverse transsulfuration pathways

Yeast and animals do not biosynthesize L-cysteine *de novo*, obtaining both cysteine and methionine through diet. Their reverse transsulfuration pathway irreversibly shunts excess L-homocysteine to L-cysteine production. In this pathway cystathionine β -synthase (CBS) condenses L-serine and L-homocysteine, forming the same L-cystathionine metabolite as the transsulfuration pathway. The L-cystathionine formed is hydrolyzed by cystathionine γ -lyase (CGL), yielding α -ketobutyrate, ammonia and L-cysteine (Jhee, McPhie and Miles, 2000; Yamagata et al., 1993).

The four enzymes of the transsulfuration and reverse transsulfuration pathways are dependent on the PLP cofactor and all but CBS are members of the γ -subfamily of fold type I. Cystathionine β -synthase is a fold-type II enzyme. Another unique aspect of this enzyme is that CBS from animal species possesses a heme cofactor as well as PLP. The heme is not involved in catalysis of the β -replacement reaction and is not present in CBS from yeast (Kery, Bukovska and Kraus, 1994). The purpose of the heme group has not yet been fully identified but studies suggest a structural role in assuring proper folding, assuming proper conformation and in the redox regulation (Hoshi and Heinemann, 2001; Taoka et al., 2002; Singh, Madzelan and Banerjee, 2007).

The *Escherichia coli* CGS and CBL (eCGS and eCBL) enzymes, which are coded by *metB* and *metC* genes respectively, share approximately 30% amino acid sequence identity and several conserved active-site residues (Farsi et al., 2009). Figure 1.4 illustrates the similarity of their structures, for which the root mean square deviation in placement of the C α atoms of the protein backbone is only 1.5 Å (Messerschmidt et al. 2003). This

observation is in keeping with the earlier suggestion that these enzyme share a common ancestor (Belfaiza et al., 1986).

Figure 1.4. Cartoon representation of an overlay of the structures of eCBL, in complex with N-Hydrazinocarbonylmethyl-2-Nitro-Benzamide (blue) and eCGS (green). (Clausen et al., 1998, PDB ID: 1CS1; Ejim et al. 2007, PDB ID: 2GQN)

1.5. Cystathionine β -lyase

1.5.1 Structure

The structure of eCBL was solved to 1.83 Å resolution by Clausen and colleagues (1996). The four subunits of this homotetrameric enzyme are arranged as a pair of catalytic dimers (Figure 1.5). Each monomer comprises three domains: the N-terminal domain (residues 1-60), the catalytic domain (residues 61 to 256), which includes the catalytic base (K210) that is covalently bound to the PLP cofactor via a Schiff base linkage in the resting state of the enzyme, and the C-terminal domain (residues 257 to 395) (Figure 1.6; Clausen et al., 1996). The N-terminal domain is situated at the interface between the catalytic dimers and extends into the active-site of the neighboring monomer, contributing residues that bind the cofactor and substrate (Clausen et al., 1996; Clausen et al., 1997; Lodha, Jaworski and Aitken, 2010). There is one active site per monomer, located at the interface between the catalytic and C-terminal domains.

While there is not a structure of eCBL in complex with its L-Cth substrate, structures of the enzyme have been solved with various inhibitors in bound to the active site (Figure 1.7). Clausen et al. (1997) and Ejim et al. (2007) showed that only minor structural changes in the active site occur upon binding of either L-aminoethoxyvinylglycine (AVG) and N-hydrazinocarbonylmethyl-2-Nitro-Benzamide. Contacts formed by these inhibitors have been useful in identifying amino acids to target for site-directed mutagenesis studies probing the role of amino acids lining the active site (Farsi et al., 2009; Lodha, Jaworski and Aitken, 2010; Lodha and Aitken, 2011; Manders et al., 2013).

Figure 1.5. The structure of the eCBL homotetramer in complex with N-Hydrazinocarbonylmethyl-2-Nitro-Benzamide inhibitor, showing one catalytic dimer on the right (white and dark grey monomers) and the other on the left (light grey and black monomers). The active site-site residues are coloured blue and the one inhibitor (red) is bound at each of the four active-sites (Ejim et al. 2007, PDB ID: 2GQN).

Figure 1.6. Cartoon representation of the eCBL monomer in complex with N-Hydrazinocarbonylmethyl-2-Nitro-Benzamide inhibitor (green) showing the N-terminal (blue), catalytic (pink) and C-terminal domains (orange) and catalytic domain (pink) (Ejim et al. 2007, PDB ID: 2GQN).

L-cystathionine

L-aminoethoxyvinylglycine
(AVG) N-hydrazinocarbonylmethyl-2-
Nitro-Benzamide

Figure 1.7. The structures of eCBL substrate (L-Cth) and two of the inhibitors used in presented crystal structures.

1.5.2 Mechanism

CBL catalyses the α,β -elimination of L-Cth, a reaction that produces L-homocysteine, pyruvate and ammonia (Flavin and Slaughter, 1967). The first steps of this reaction, common to all the PLP- dependant enzymes, are substrate binding and transaldimination, whereby the α -amino group of the substrate replaces the ϵ -amino group of K210 in Schiff-base linkage with the PLP, releasing the catalytic base (K210) to initiate the next step in the reaction (Figure 1.8). The catalytic base then abstracts the proton from the α -carbon of the substrate (the α -proton), leading to the formation of a carbanion. Although Clausen et al., (1997) proposed that this carbanion is resonance stabilized, via a quinonoid intermediate, in the aromatic ring of the cofactor, this is not the only possible mechanism and a quinonoid has not yet been observed for CBL. Studies on aspartate aminotransferase (AAT), the most thoroughly characterized PLP enzyme, have demonstrated that formation of the quinonoid is promoted by protonation of the pyridinium nitrogen, which is favored by the presence of a negatively charged residue within hydrogen bonding distance, at the base of the ring (Yano et al., 1992; Onuffer and Kirsch, 1994; Griswold and Toney, 2011). An acidic residue (D185) is positioned to interact with the pyridinium nitrogen of eCBL, but to date the contribution of D185 to the α,β -elimination mechanism of this enzyme has not been investigated. Following abstraction of the α -proton, the homocysteine leaving group, at the β -position is released, producing the external aldimine of aminoacrylate, which undergoes a transaldimination with the ϵ -amino group of K210 to release aminoacrylate. This product spontaneously hydrolyses to pyruvate and ammonia in solution (Clausen et al., 1997).

Figure 1.8. The eCBL mechanism proposed by Clausen et al. (1997), in which the α,β -elimination proceeds via an E1 mechanism. The (A) internal aldimine complex with the L-Cth substrate, (B) external aldimine of the L-Cth substrate, (C) quinonoid, (D) external aldimine of aminoacrylate and (E) internal aldimine resting state of the eCBL reaction are shown.

1.5.3 Active site layout

Several of the residues forming the active site are conserved among the members of the γ -subfamily of fold type I. Previous studies from our lab have characterized the role of several active-site residues in the context of eCBL, eCGS and yeast CGL, including those responsible for substrate binding and positioning of the ϵ -amino group of the catalytic base. Of the 22 residues that are conserved in the γ -subfamily enzymes, 7 are situated in the active site and have a role in cofactor and/or substrate positioning (Lodha and Aitken, 2011). Among these is the catalytic base (eCBL K210), which is conserved among all PLP dependent enzymes (Christen and Mehta, 2001) and flanked by T209 and Y211 in eCBL. While the threonine in position 209, which forms a hydrogen bond to the phosphate group of the PLP cofactor, is conserved among the members of the γ -subfamily, the tyrosine at position 211 is sometimes replaced by phenylalanine. The crystal structure of the eCBL-AVG complex (Clausen et al., 1997) suggests a role for an arginine residue (eCBL-R262) in positioning Y211 through hydrogen bond formation between the arginine side chain and the tyrosine backbone.

When released from its Schiff base linkage with the cofactor, the ϵ -amino group of the catalytic base forms hydrogen bonds with Y56 and S339, which are conserved in the majority of the γ family (Clausen et al., 1996). These interactions tether the catalytic base and regulate the substrate and cofactor positions accessible to it for proton transfer. As such they are primary determinants of reaction specificity (Lodha and Aitken, 2011). The positioning of S339 may in turn be influenced by other active site residues, such as S32, S33 and K42 of eCBL, which were explored by Manders et al. (2013).

Residues T187 and A207 lie adjacent to the cofactor ring. A hydrogen bond between the peptide backbones of A207 and K210 connects these residues and may contribute to the positioning of the catalytic base with respect to the cofactor (Figure 1.9; Clausen et al., 1996). While A207 is conserved among bacterial CBL enzymes, this residue is replaced by a serine in the closely related cystathionine γ -synthase. The side chain T187, conserved among members of the γ -subfamily, is within hydrogen bonding distance from the backbone of A207 and the oxygen of the side chain carboxyl group of D185, which is positioned to interact with the pyridinium nitrogen at the base of the PLP cofactor (Clausen et al., 1996). An acidic residue in this location is a common feature of fold-type I enzymes. In contrast, fold-type II enzymes, such as *O*-Acetylserine sulfhydratase (OASS), have serine in place of eCBL-D185. Like CBL, OASS catalyzes a α,β -elimination, but unlike AAT, this reaction proceeds via an E2 mechanism that does not require the formation of a quinonoid intermediate (Tai and Cook, 2001).

The structure of eCBL in complex with the inhibitor AVG in Schiff-base linkage with the PLP cofactor illustrates some of the binding interactions of the L-Cth substrate (Figure 1.7; Clausen et al., 1997). Like L-Cth, AVG possesses α -amino and α -carboxylate moieties, as well as a distal amino group, but the sulfur atom is replaced with oxygen, and it lacks a distal carboxyl group. Binding interactions identified in the eCBL-AVG structure include hydrogen bonds between the α -carboxylate and residues R372 and W340 and between the distal carboxylate and R58 (Clausen et al., 1997). Our lab subsequently confirmed that these residues, particularly the two arginines, which also conserved in eCGS and γ CGL, are crucial for substrate binding (Messerschmidt et al., 2003; Lodha, Jaworski and Aitken,

Figure 1.9. Schematic of the eCBL active site showing the interactions of targeted D185, T187, A207, T209, Y211, R262, L310 and W340 residues and the PLP cofactor. The wavy lines represent possible hydrogen bonds.

2010; Jaworski et al., 2012). While residue W340 also contributes to substrate binding, substitution of this residue with phenylalanine reduces the K_m^{L-Cth} by only an order of magnitude. The position of W340, adjacent to S339, which tethers the ϵ -amino group of the catalytic base, and within hydrogen-bonding distance of the O3' position of the cofactor, suggest that this residue, conserved as tryptophan in bacterial CBL and as leucine in bacterial CGS and fungal and animal CGL, may act as a determinant of reaction specificity. The positioning of W340 within the active site is influenced by L310, the side chain of which packs against the tryptophan's indole group (Figure 1.10; Ejim et al. 2007).

Two other residues that interact with the substrate, although via packing interactions, and are conserved distinctly in bacterial CBL enzymes versus bacterial CGS and yeast and fungal CGL are F55 and Y338 (eCBL numbering). These two aromatic residues are replaced by an acidic pair in CGS and CGL. Their interconversion in the context of eCBL, eCGS and yCGL diminished enzyme activity but did not alter reaction specificity and this work provided the first indication that specificity is likely regulated in a concerted manner by several active-site residues (Farsi et al., 2009).

To date, our lab has characterized the role of 13 residues in the active site of eCBL, and several of the corresponding positions in eCGS and yCGL (Farsi et al. 2009; Lodha, Jaworski and Aitken 2010; Lodha and Aitken 2011; Jaworski et al. 2012; Jaworski and Aitken 2013; Manders et al. 2013; Hopwood, Ahmed and Aitken 2014). The majority of these are involved in positioning of the substrate or the catalytic base. While residues interacting with the PLP or positioning the catalytic base with respect to the cofactor, have

been characterized in the context of other PLP enzymes, to date the role of these residues have not been reported for eCBL.

Figure 1.10. PyMOL visualization of the packing interactions between W340 (orange) and L310 (blue) (Ejim et al. 2007, PDB ID: 2GQN). The dotted surface corresponds to the van der Waals surfaces of the two residues.

1.6 Project objectives

Our lab has characterized several of the residues, including those interacting with substrate(s), in the active sites of eCBL, eCGS and yeast cystathionine γ -synthase (yCGL). Based on the previous studies we hypothesize that the active site residues that interact with the PLP cofactor and position it with respect to the catalytic base (eCBL-K210) influence the activity of eCBL. To test the hypothesis following variants were constructed:

1. Conservative substitutions of T209 (T209N and T209V) and Y211 (Y211F), which flank the catalytic base, and R262 (R262K and R262L), which influences the position of the backbone of Y211, were constructed to explore the contributions of the hydrogen bonds formed by the side chains of these residues to eCBL activity.
2. The alanine residue at position 207 (eCBL-A207) was substituted with serine, threonine, asparagine and aspartic acid to investigate how the possibility of hydrogen bond formation, altering residue charge, and increasing the size of the residue, influences the backbone to backbone interaction of K210 catalytic base as well as backbone to side chain interaction with T187, as evidenced by their effect on the kinetic parameters of eCBL.
3. A series of eCBL-D185 substitutions (D185L, D185N, D185S, D185T and D185V) were constructed to probe the role of the hydrogen bond between the side chain of this residue and the pyridinium nitrogen of the cofactor in the context of the eCBL-catalyzed hydrolysis of L-Cth.
4. The T187S and T187V variants were created to investigate the hydrogen bond bridge between D185 and A207, the central component of the hydrogen bonding

network linking the pyridinium nitrogen at the base of the PLP ring and the catalytic base.

5. The W340A, W340H and W340L variants extend earlier work by Lodha and colleagues (2010) in which the eCBL-W340F substitution was characterized.

2. METHODS

2.1 Reagents

5,5-dithio-bis-(2-nitrobenzoic acid) (DTNB) was obtained from ThermoFisher Scientific. Nickel-nitriloacetic acid (Ni-NTA) resin was purchased from Qiagen and restriction endonucleases, T4 DNA ligase and Antarctic phosphatase were from New England Biolabs. The cystathionine substrate for enzyme assays and PLP were obtained from Sigma Aldrich. The oligonucleotide primers used for site-directed mutagenesis were synthesized by Integrated DNA Technologies and the resulting mutants were sequenced by BioBasic prior to expression. All the other reagents were purchased from BioShop.

2.2 Construction of site-directed mutants

A series of 20 site-directed mutants were constructed using the overlap-extension polymerase chain reaction (OE-PCR) method (Ho et al., 1989). The forward and reverse mutagenic primer pairs are presented in Table 1. The pSeq0 and pSeq7r primers are designed to anneal to sequences in the vector approximately 250 base pairs 5' and 3', respectively, of the multiple cloning site, where the eCBL coding sequence is inserted. These vector primers were used for site-directed mutagenesis and for sequencing of the resulting constructs.

Following restriction endonuclease digestion, eCBL PCR amplicons were inserted between the *Nde*I and *Kpn*I sites of the pTrc-99aAF vector, a modified version of pTrc-99 encoding an N-terminal 6-His tag and a linker containing a protease cleavage site (Farsi *et al.*, 2009).

The resulting plasmids were transformed into *E. coli* strain ER1821 $\Delta metC::cat$, in which the coding sequence of the native CBL gene is replaced with that of chloramphenicol acetyl transferase to provide a selectable marker and prevent contamination of site-directed variants expressed in this strain with the wild-type eCBL enzyme (Farsi *et al.*, 2009). All site-directed mutants were sequenced prior to protein expression and purification.

Table 2.1. The oligonucleotide primers employed for construction of the eCBL site-directed mutants.

Mutation	Forward primer	Reverse primer
D185L	5'-cattatgatcctgaacacctgggcagc	5'-gctgccaggtgttcaggatcataatg
D185N	5'-cattatgatcaacaacctgggcagc	5'-gctgccaggtgttgtgatcataatg
D185S	5'-cattatgatcagcaacctgggcagc	5'-gctgccaggtgttgctgatcataatg
D185T	5'-cattatgatcaccaacctgggcagc	5'-gctgccaggtgttggtgatcataatg
D185V	5'-cattatgatcgtgaacctgggcagc	5'-gctgccaggtgttcacgatcataatg
T187S	5'-gatcgacaacagctgggcagccggtg	5'-caccggctgccagctgtgtgcgac
T187V	5'-gatcgacaacgttgggcagccggtg	5'-caccggctgcccacgtgtgtgcgac
A207D	5'-cgatgtttctattcaagaccacc	5'-ggcggcgtcttgaatagaaacatcg
A207N	5'-cgatgtttctattcaaacaccacc	5'-ggcggcgttttgaatagaaacatcg
A207S	5'-cgatgtttctattcaaacaccacc	5'-ggcggcgttttgaatagaaacatcg
A207T	5'-cgatgtttctattcaaacaccacc	5'-ggcggcgttttgaatagaaacatcg
T209N	5'-caagccgccaacaatatctggttg	5'-caaccagatatttgtggcggttg
T209V	5'-caagccgccgtgaaatatctggttg	5'-caaccagatatttcacggcggttg
Y211F	5'-ccaccaaatttctggtgggcattc	5'-gaatgccaaccagaaatttgggtg
R262K	5'-cattaggtgtgaaattgcgtcaacac	5'-gatgttgacgcaatttcacacctaag
R262L	5'-cattaggtgtgctgttgcgtcaacac	5'-gatgttgacgcaacagcacacctaag
L310A	5'-gcagcggcgcttttctttgtgc	5'-gcacaaaggaaaacgcaccgctgc
W340A	5'-ctactcggcaggcgggtatgaatc	5'-gattcatacccgctccgagtag
W340H	5'-ctactcgcattggcgggtatgaatc	5'-gattcatacccgccatgcgagtag
W340L	5'-ctactcgcggggcgggtatgaatc	5'-gattcatacccgccagcgagtag
pSEQsec0	5'-ggcgtcaggcagccatcggaagctg	N/A
pSEQsec7r	N/A	5'-gcccgccacctccggcggtgttcgc

2.3 Expression and purification of wild-type and site-directed mutants of eCBL

A 100-mL culture, of *E. coli* strain ER1821 $\Delta metC::cat$ expressing wild-type eCBL, or each of the 20 constructed variants of eCBL, from the pTrc99aAF vector was grown overnight at 30 °C and 200rpm in Luria broth (LB) media (10g NaCl, 10g tryptone, 5g yeast extract/1 L) containing 100 µg/L of ampicillin. The next day, three 1-L cultures of LB media, containing 100 µg/L of ampicillin, in 2.8-L Fernbach flasks, were each inoculated with 10 ml of the overnight culture and grown under the same conditions until the OD₆₀₀ reached 0.5. The cultures were then induced with 150 µM isopropyl β-D-1-thiogalactopyranoside (IPTG) and left to grow for a further 16 hrs, prior to harvesting. Cells were harvested by centrifugation at 5,000 rpm and 4 °C for 10 minutes, washed by resuspending in 0.85% NaCl, and centrifuged again. The resulting cell pellets were resuspended in 70 ml of buffer A (50 mM potassium phosphate buffer, pH 7.8, with 10 mM imidazole and 20 µM PLP). Lysozyme (approximately 1 mg/mL) was added and the cells were incubated for 20 min on ice prior to sonication (8 x 30sec at 20 W). The cell lysate was centrifuged at 16,000 rpm and 4 °C for an hour to separate the soluble fraction from cellular debris.

The lysate was loaded on a chromatography column containing 5 mL of nickel nitrilotriacetic acid (Ni-NTA) resin (Qiagen), equilibrated with buffer A, and the column was washed with 100 mL of buffer A. The enzyme was then eluted with a 200-mL gradient of 10-200 mM imidazole in buffer A and the purity of collected fractions was determined visually by SDS PAGE. Fractions deemed to be at least 90% pure were combined. In the case of variants exhibiting double banding, proposed to be due to truncation, purity was assessed for a combination of the 45 and 42-kDa bands combined. The fractions were then

concentrated, and dialyzed against dialysis buffer (50 mM potassium phosphate buffer, pH 7.8, 20 μ M PLP and 2 mM ethylenediaminetetraacetic acid (EDTA)) overnight at 4 °C. Glycerol was added to 25% and the enzyme was aliquoted and stored at -80 °C.

2.4 Steady-state kinetics

The L-Cth hydrolysis activity of eCBL was measured using the continuous assay described by Aitken and Kirsch (2003). Hydrolysis of L-Cth was detected by reaction of the L-Hcys with DTNB (5,5-dithio-bis-(2-nitrobenzoic acid)) (Figure 2.1). The absorbance of the TNB⁻ ion (2-nitro-5mercapto-benzoic acid) at 412nm ($\epsilon_{412} = 13,600 \text{ M}^{-1}\text{cm}^{-1}$) allows the reaction to be monitored continuously. Enzyme activity was measured in a total volume of 100 μ l at 25°C. The assay solution contained 50mM Tris buffer, pH 8.5, 20 μ M PLP, 2 mM DNTB, 0.018-22mM Cth and enzyme concentrations ranging from 14nM to 1 μ M, depending on variant activity.

Following background correction, the initial velocity data for the range of substrate concentrations were fit to the Michaelis-Menten equation (Equation 1) to determine the kinetic parameters k_{cat} and K_m^{L-Cth} . The catalytic efficiency (k_{cat}/K_m^{L-Cth}) was determined independently by fit to equation 2. The data of variants that did not demonstrate saturation within the solubility limit of the Cth substrate were fit to equation 3, which is derived from the Michaelis-Menten equation with the assumption that K_m approximates the $K_m + [S]$ denominator because the K_m is greater than the highest substrate concentration tested ($K_m \gg [S]$). The kinetic data from R262K variant demonstrates a sigmoidal shape and as such

Figure 2.1. The continuous, DTNB-based assay used to measure the L-Cth hydrolysis activity of eCBL.

was fit to the Hill equation, which allows estimation of cooperativity in enzymes with more than one active site (Equation 4).

Equation 1

$$\frac{v}{[E]} = \frac{kcat * [S]}{Km + [S]}$$

Equation 2

$$\frac{v}{[E]} = \frac{\frac{kcat}{Km} * [S]}{1 + \frac{[S]}{Km}}$$

Equation 3

$$\frac{v}{[E]} = \frac{kcat}{Km} * [S]$$

Equation 4

$$\frac{v}{[E]} = \frac{kcat * [S]^n}{Km^n + [S]^n}$$

2.4 Affinity chromatography under denaturing conditions

The purified eCBL-A207T, T209V and W340A variants, which exhibit double banding, were analyzed by Ni-NTA chromatography under denaturing conditions in order to determine whether the site of proteolytic cleavage could be toward the N-terminus of the protein, where the 6-His affinity tag is located (Bornhorst and Falke, 2011). A 2-mL aliquot of Ni-NTA agarose resin was washed twice with 6 mL of denaturing, binding buffer (20 mM sodium phosphate, pH 7.8, 500 mM NaCl and 8 M urea) by resuspending the resin and gently rocking. An 8-ml sample of the previously purified protein was added, and the mixture incubated for 30 minutes at room temperature. The resin was then washed two times with 4 mL of denaturing binding buffer, two times with 4 mL of pH 6.0 denaturing wash buffer (20 mM sodium phosphate, pH 6.0, 500 mM NaCl, 8M urea) and two times with pH 5.3 denaturing wash buffer (20 mM sodium phosphate, pH 5.3, 500 mM NaCl, 8M urea). The protein was then eluted with 5 mL of pH 4.0 denaturing elution buffer (20 mM sodium phosphate, pH 4.0, 500 mM NaCl, 8 M urea) and five 1 mL fractions of the eluate were collected from the column. The samples of wash and eluate were compared on SDS-PAGE electrophoresis to determine the size of the eluted enzyme.

2.5 PyMOL structure analysis

A series of mutations was constructed *in silico* to provide a structural context for interpretation of the kinetic parameters of the site-directed variants characterized in this study. The *in silico* mutations were made using the PyMOL Molecular Graphics System (DeLano, 2002). The mutagenesis option from the wizard menu was used and the selected

residue was changed for the chosen substitution. The output was presented as a selection of rotamers ranked on the frequency with which each rotamer is observed in protein structures. The rotamers with the lowest degree of common space occupancy (i.e. structural overlap) with neighbouring residues were chosen.

3. RESULTS

Twenty site-directed variants, targeting eight active-site residues, of eCBL were constructed, expressed and purified. The purification yields varied between 1.0-38.4 mg/L, compared to 93.6 mg/L for the wild-type enzyme (Table 3.1). The molecular weight of eCBL, including the 6-His tag and a linker is 44.95 kDa and a prominent band of this size was visible by SDS PAGE for each variant, as illustrated in purification gels of the wild-type enzyme (Figure 3.1) and comparison of each of the purified enzymes (Figure 3.2). Fourteen of the variants (D185L, D185V, T187S, T187V, A207D, A207N, A207T, T209N, T209V, Y211F, R262L, W340A, W340H and W340L) show an additional band at approximately 42 kDa (Figure 3.2). The variants showing a 42-kDa band with intensity similar to that of the full-length enzyme also demonstrated the lowest yields, compared to wild-type eCBL (Table 3.1).

To discover whether truncation occurs at the amino- or carboxy terminus of the eCBL variants displaying a ~42-kDa band (Figure 3.2) were separated by Ni-NTA chromatography under denaturing conditions. Since all three variants that were separated show the same result, only the gel for eCBL-A207T is shown (Figure 3.3). The 42-kDa band of variants, is released from the affinity column before adding the final, pH 4.0 buffer, which eluted the band corresponding to the full-length, 45-kDa enzyme.

Table 3.1. Yields, in milligrams of protein per litre of culture, of purified eCBL variants and the wild-type enzyme, and the presence of a 42-kDa band following purification.

Variant	Yield [mg/L]	42-kDa band*	Variant	Yield [mg/L]	42-kDa band*
Wild type eCBL	93.6	-	eCBL-A207T	2.4	++
eCBL-D185L	38.4	+	eCBL-T209N	6.5	++
eCBL-D185N	34.2	-	eCBL-T209V	1.0	++
eCBL-D185S	15.2	-	eCBL-Y211F	18.0	+
eCBL-D185T	24.0	-	eCBL-R262K	10.8	-
eCBL-D185V	14.6	+	eCBL-R262L	1.9	++
eCBL-T187S	20.4	+	eCBL-L310A	20.8	-
eCBL-T187V	17.6	+	eCBL-W340A	8.4	++
eCBL-A207D	10.2	++	eCBL-W340H	7.0	+
eCBL-A207N	3.9	++	eCBL-W340L	8.9	++
eCBL-A207S	29.5	-			

* - indicates the presence of single band corresponding to full-length eCBL, + indicates that the density of the 42-kDa band is less than that of the full-length eCBL band, and ++ indicates that the density of the 42-kDa band is approximately equal to that of the 45-kDa band corresponding to full-length eCBL.

Figure 3.1. SDS-PAGE gels of fractions from the purification of (A) eCBL-W340L and (B) wild-type eCBL. The lanes correspond to: P indicates resuspended cell lysate pellet, L indicates the supernatant of the centrifuged crude lysate, F is the flow through from loading of the crude lysate on the Ni-NTA affinity column, and the numbers correspond to eluted fractions.

Figure 3.2. SDS-PAGE gel comparing 2 μ g of wild type eCBL and purified site-directed variants. Upper gel: Lane 1: A207D, lane 2: A207N, lane 3: A207S, lane 4: A207T, lane 5: wild type, lane 6: T209N, lane 7: T209V, lane 8: Y211F, lane 9: R262K, lane 10: R262L. Lower gel: Lane 1: L310A, lane 2: W340A, lane 3: W340H, lane 4: W340L, lane 5: wild type, lane 6: D185L, lane 7: D185N, lane 8: D185S, lane 9: D185T, lane 10: D185V, lane 11: T187S, lane 12: T187V.

Figure 3.3. Affinity chromatography, under denaturing conditions, of a purified sample of the eCBL-A207T variant. Lanes 1-10 contain fractions eluted from the Ni-NTA resin with denaturation wash buffer pH 7.8 (lane 1) and elution buffer at pH 6.0 (lanes 2 and 3), pH 5.3 (lanes 4 and 5) and pH 4.0 (lanes 6-10).

3.1 Steady state kinetic analysis

The steady-state kinetic parameters of wild-type eCBL and the 13 site-directed variants with detectable activity are presented in table 3.2. The catalytic efficiency (k_{cat}/K_m^{L-Cth}) of the wild-type enzyme is within experimental error of that reported previously by our lab (Lodha, Jaworski and Aitken, 2010).

The 120-10,500-fold reductions in the catalytic efficiency of the five D185 variants are dominated by decreases in their turnover rates (k_{cat}) (Table 3.2, Figure 3.4). Replacement of D185 with serine, threonine or asparagine results in 3-5 order of magnitude decreases in the catalytic efficiency of eCBL. The catalytic efficiency of D185L, which has 40% higher volume than the aspartate side chain of D185 and has no hydrogen-bonding capacity, is reduced by 1000-fold, two orders of magnitude less than D185V, which similarly lacks the ability to form hydrogen bonds but is only 15% larger than aspartate (Table 3.2; Figure 3.6; Darby and Creighton 1993).

Substitution of residue T187 with serine increases the catalytic efficiency of this enzyme by 60%, an effect caused by the 5-fold decrease in K_m compared to the wild type enzyme (Table 3.2; Figure 3.5). In contrast, replacement of T187 with valine, which removes the hydrogen-bonding capacity of the side chain, reduces catalytic efficiency by 100-fold, an effect dominated by the 20-fold reduction in k_{cat} .

Table 2.2. Kinetic parameters of wild-type eCBL and site-directed variants.

Variant	k_{cat} [s ⁻¹]	K_m^{L-Cth} [mM]	k_{cat}/K_m^{L-Cth} [M ⁻¹ s ⁻¹]
wt-eCBL	53.5 ± 1.3	0.25 ± 0.02	(2.1 ± 0.1) × 10 ⁵
wt-eCBL [†]	34.1 ± 0.6	0.18 ± 0.01	(1.9 ± 0.1) × 10 ⁵
eCBL-D185L	0.059 ± 0.002	0.26 ± 0.04	138.3 ± 19
eCBL-D185N	0.356 ± 0.008	0.19 ± 0.02	(1.8 ± 0.2) × 10 ³
eCBL-D185S	N/A	N/A	43 ± 0.1
eCBL-D185T	N/A	N/A	30 ± 0.1
eCBL-D185V	N/A	N/A	20 ± 0.8
eCBL-T187S	16.3 ± 0.4	0.05 ± 0.006	(3.3 ± 0.4) × 10 ⁵
eCBL-T187V	1.3 ± 0.08	0.4 ± 0.1	(2.8 ± 0.6) × 10 ³
eCBL-A207S	15.0 ± 0.7	0.11 ± 0.02	(1.4 ± 0.2) × 10 ⁵
eCBL-Y211F	2.6 ± 0.2	0.22 ± 0.07	(1.2 ± 0.3) × 10 ⁴
eCBL-R262K	17.5 ± 1.3	2.2 ± 0.6	(7.9 ± 1.5) × 10 ³
eCBL-L310A	32.8 ± 1.7	1.6 ± 0.2	(1.5 ± 0.2) × 10 ⁴
eCBL-W340F [‡]	79 ± 1	1.35 ± 0.05	(5.9 ± 0.2) × 10 ⁴
eCBL-W340H	7.4 ± 0.3	2.2 ± 0.2	(3.3 ± 0.2) × 10 ³
eCBL-W340L	N/A	N/A	0.19 ± 0.04

[†]kinetic parameters for wild-type eCBL reported by Lodha et al. (2010)

[‡]kinetic parameters for eCBL-W340F are from Jaworski and Aitken (2013).

N/A – not available because the kinetic parameters could not be determined from the equation 1 as K_m^{L-Cth} exceeds the solubility limit of the substrate.

Figure 3.4. The effect of L-cystathionine concentration on the activity of A: wild-type eCBL (black, primary y-axis) and the D185N (red, secondary y-axis) variant; B: wild type (black, primary y-axis), D185S (red, secondary y-axis) D185T (blue, secondary y-axis), D185L (green, secondary y-axis) and D185V (cyan, secondary y-axis) site-directed variants. The solid lines correspond to the fit of data to the Michaelis-Menten equation (equation 1).

Figure 3.5. The effect of L-cystathionine concentration on the activity of wild-type eCBL (black, primary y-axis) and the T187S (red, secondary y-axis) variant. The solid lines correspond to the fit of data to the Michaelis-Menten equation (equation 1).

The kinetic parameters of eCBL-A207S, which demonstrates no truncation, are within 3-fold of wild-type eCBL (Table 3.2). In contrast, no activity is detected for the asparagine, aspartate or threonine substitutions of A207, which all show a pronounced 42-kDa band (Table 3.1; Figure 3.2). Substitution of T209, situated adjacent to the catalytic base (K210), with asparagine or valine, also results in inactive enzyme. The purified T209N and T209V variants also exhibited low yields and a 42-kDa band. In contrast, substitution of Y211, the other flanking residue of K210, with phenylalanine results in a decrease of only 10-fold in catalytic efficiency, due to a corresponding decrease in the turnover rate of the enzyme. A 42-kDa band was observed apparent for Y211F, but to a lesser degree than the two T209 variants (Figure 3.2). While R262L exhibits no kinetic activity, kinetic data of R262K variant does not fit into Michaelis-Menten equation and fitted to Hill equation shows value of Hill coefficient of 2.38 (Figure 3.6).

Substitution of residue W340 with histidine results in a 7-fold reduction in k_{cat} , compared to the previously characterized W340F variant of eCBL, for which the turnover rate is within 1.5-fold of the wild-type enzyme (Table 3.2; Jaworski and Aitken, 2013). A similar 10-fold increase in K_m^{L-Cth} is observed for the histidine and phenylalanine substitutions. In contrast, replacement of W340 with a nonaromatic residue, such as leucine or alanine increases the degree of protein truncation, as indicated by the density of the ~42-kDa band and drastically reduces enzyme activity. No activity is detectable for eCBL-W340A and the catalytic efficiency of W340L, which is not saturated within the solubility limit of the substrate, is reduced by 6 orders of magnitude, compared to the wild type enzyme.

Figure 3.6. The effect of L-cystathionine concentration on the activity of the eCBL R262K variant (red, primary y-axis), data fit to the Hill equation (equation 4) and wild type eCBL (black, secondary y-axis), data fit to Michaelis-Menten equation (equation 1).

4. DISCUSSION

The range of amino acid transformations catalysed by enzymes dependent on the PLP cofactor highlights the role of active-site residues interacting with the substrate(s) and cofactor as determinants of enzyme activity and specificity. Cystathionine β -lyase catalyses the second reaction of the bacterial transsulfuration pathway. It is an useful model system for exploring cofactor regulation by the protein component of PLP enzymes because of its structural similarity to other enzymes catalyzing side chain rearrangements of sulfur-containing amino acids, the availability of several crystal structures and earlier work by our lab characterizing the role of several key active-site residues. These studies, and those of the related enzymes *E. coli* cystathionine γ -synthase (eCGS) and yeast cystathionine γ -lyase (yCGL), have identified the roles of several active-site residues in substrate binding and catalysis as well as part of the basis for regulation of specificity (Clausen et al., 1996; Messerschmidt et al., 2003; Ejim et al., 2007; Farsi et al., 2009; Lodha, Jaworski and Aitken, 2010; Lodha and Aitken 2011; Jaworski et al., 2012; Jaworski and Aitken, 2013; Manders et al., 2013; Hopwood, Ahmed and Aitken, 2014). This study extends the network of characterized active-site residues by characterization of 20 site-directed variants of 8 eCBL active-site residues that influence the positioning of catalytic base and/or PLP cofactor.

To investigate the potential effects of the site-directed substitutions on the structure of the eCBL active site, each of the variants was also constructed *in silico* using the ‘mutagenesis’ function of the PyMOL Molecular Graphics System (DeLano, 2002). A limitation of this method is that while PyMOL highlights overlapping van der Waals surfaces (i.e. where two atoms occupy the same space) for the most likely rotomers of the substituted side

chain, the program does not perform energy minimization to lessen these structural conflicts by allowing the structure to adapt to the change. It also does not adjust the side chain's position based on the potential to form hydrogen bond or salt bridges. As a result, the *in silico* mutagenesis is more likely to be reliable for subtle side chain substitutions (e.g. asparagine for aspartate or alanine for serine) where minimal rearrangement of neighboring residues is required to accommodate the change. The rotamers presented and discussed here are those with the lowest degree of common space occupancy with neighbouring residues.

The discussion below is organized under four headings, based on the roles of the residues investigated. Section 4.1 focuses on D185, which interacts, via a hydrogen bond, with the pyridinium nitrogen of the cofactor and has a proposed role in modulating the capacity of the cofactor to act as an electron sink during catalysis. The hydrogen-bonding network that bridges the pyridinium nitrogen of the cofactor and the catalytic base (K210) is discussed in section 4.2. The influence of the residues immediately adjacent, in sequence, to the catalytic base on the activity of eCBL is explored in section 4.3. Finally, section 4.4 builds on earlier research from our lab and discusses the role of W340, which interacts with the PLP cofactor and the L-cystathionine substrate of eCBL, and of nearby L310.

4.1. D185 variants

It is widely accepted that the ring of the PLP cofactor acts as an electron sink and protonation of the pyridinium nitrogen of the ring assists in stabilization of the carbanion following removal of one of the groups (proton, carboxyl or side chain) attached to the C α position of the substrate (Brattnshtbin and Shemyakin, 1953; Metzler, Ikawa and Snell,

1954; Snell, 1958). It is the interaction of active-site residues with the cofactor, particularly the aromatic ring, which modulates its electrophilic properties, thereby contributing to the regulation of this catalytically versatile cofactor. (Griswold and Toney, 2011). For example, while a negatively charged acidic aspartate or glutamate interacts with the pyridinium nitrogen in all fold type I enzymes, a serine or arginine can occupy this position in enzymes of other fold types (Hyde et al., 1988; Shaw, Petsko and Banerjee, 1997; Burkhard et al., 1998; Sharif et al., 2007). Typical of fold-type I enzymes, an aspartate residue (eCBL-D185) is situated within hydrogen-bonding distance of the pyridinium nitrogen in eCBL (Clausen et al., 1996). This residue has not been previously explored in the context of any of the enzymes of the γ -subfamily of fold-type I and substitution of eCBL-D185 provides the opportunity to understand the role of the conserved acidic residue in the context of a β -elimination reaction.

With the goal of investigating the position/packing, charge and hydrogen-bonding capacity of the eCBL-D185 side chain, five site directed variants were constructed. The aspartate at position 185 was substituted with asparagine (D185N), which maintains hydrogen-bonding and packing capacities but eliminates the negative charge. It was also replaced with serine (D185S) and threonine (D185T), which can form a hydrogen bond, but are uncharged and smaller than aspartate or asparagine, and with valine (D185V) and leucine (D185L), both incapable of hydrogen bonding and offering the opportunity to explore the role of packing and hydrogen-bonding interactions. All of the D185 substitution variants maintain a degree of eCBL activity (Table 3.2). With the exception of D185L, the eCBL-D185 variants did not demonstrate truncation, as evidenced by the presence a ~42-kDa band on SDS-PAGE

(Figure 3.1; Table 3.2), suggesting that these substitutions do not alter the ability of eCBL to adopt its native, homotetrameric structure.

In the crystal structure of the unliganded, resting state of eCBL, with the catalytic lysine bound in Schiff base linkage to the catalytic base (K210), the distance between the site chain oxygens of D185 and nitrogen of pyridinium ring is 2.7 Å and 3.6 Å. These distances are not changed by the presence of an inhibitor, bound in Schiff-base linkage with the cofactor, in the active site (Clausen et al., 1996; Ejim et al., 2007). As such, the distances mentioned in this discussion, unless otherwise specified, are from the structure of the resting state of eCBL without inhibitor (Clausen et al., 1996).

Replacement of the charged aspartate side chain with that of neutral, polar asparagine provides the opportunity to assess the role of the negative charge of eCBL-D185. The minor, 5-Å³, greater side-chain volume of asparagine (96 Å³), compared to aspartate (91 Å³), is unlikely to result in a shift in the position of the cofactor within the active site (Darby and Creighton, 1993). As anticipated, the substitution of D185 with asparagine *in silico* demonstrates minimal structural overlap between the N185 side chain and other active-site residues (Figure 4.1). The 120-fold decrease in catalytic efficiency of eCBL-D185N, compared to the wild-type enzyme, is due to a corresponding decrease in turnover (k_{cat}), as K_m^{L-Cth} is unchanged (Table 3.2). This provides further evidence that active-site architecture and substrate binding are not altered and the effect of substitution of D185 with the polar but uncharged side chain of asparagine is limited to modification of the electronic properties of the aromatic ring of the cofactor.

The D185S and D185T variants were also constructed to probe the contribution of the hydrogen bond from residue 185 to the pyridinium nitrogen. Although the threonine side

Figure 4.1. Comparison of the interactions of D185 and the asparagine, threonine and serine substitutions, with the pyridinium nitrogen of the cofactor. A – schematic figure of wild-type eCBL and the D185N, D185T and D185S substitution variants; B, C, D and E show the corresponding PyMOL models, respectively. All the distances shown in Å (Clausen et al. 1996, PDB ID: 1CL1)

Figure 4.2. Comparison of the interactions of D185, and the leucine and valine substitutions, with the pyridinium nitrogen of the cofactor. A – schematic figure of wild-type eCBL and the D185L and D185V substitution variants; B, C and D show the corresponding PyMOL models. (Clausen et al 1996, PDB ID: 1CL1)

chain ($93 \text{ \AA}^3$) occupies only 2% greater volume as that of aspartate ($91 \text{ \AA}^3$), the *in silico* model of eCBL-D185T indicates that the rotamer with the least structural overlap positions the side-chain hydroxyl $5.2 \text{ \AA}$ from the pyridinium nitrogen, suggesting that a hydrogen bond is not formed (Figure 4.1). The length of a hydrogen bond ranges from $2.2\text{-}4.0 \text{ \AA}$, with the strength determined by the distance between the hydrogen-bridged atoms, such that the contribution of interactions over $3.0\text{-}3.5 \text{ \AA}$ is negligible (Darby and Creighton, 1993; DeLano, 2002). The side chain oxygen of D185S variant shows the distance of $3.6 \text{ \AA}$ from the pyridinium nitrogen making the hydrogen bond interaction possible but weak (Figure 4.1). The observed 4,900 and 7,000-fold reductions in catalytic efficiency of the serine and threonine substitutions of D185, in combination with the increase in their K_m^{L-} C_{th} beyond the solubility limit of the substrate, suggests that the active sites architecture of these variants is altered in a manner that affects substrate binding (Table 3.2). This may be driven by alternative hydrogen bonds formed by the hydroxyl group of their side chains (i.e. interactions that do not occur in the native eCBL enzyme). It may also due to shifting of the PLP cofactor with respect to substrate-binding residues resulting from the packing interactions of S185 and T185 that change the shape of the base of the active site. The 40% greater catalytic efficiency of eCBL-D185S, in comparison with D185T, despite the 20% smaller volume of serine ($73 \text{ \AA}^3$), reflects a combination of structural changes that cannot be interpreted without solving the structures of these variant (Darby and Creighton, 1993).

Replacement of D185 with the hydrophobic residues leucine and valine eliminates the hydrogen bond between the side chain of the residue at position 185 and the pyridinium nitrogen of the cofactor. Unlike the asparagine substitution, the branched methyl groups of valine ($105 \text{ \AA}^3$) and leucine ($124 \text{ \AA}^3$) occupy a 15 and 36% larger volume, respectively,

than the carboxylate-containing side chain of aspartate ($91 \text{ \AA}^3$) and, as such, require movement of neighboring residues and/or the PLP cofactor in order to accommodate their greater size (Darby and Creighton, 1993). It is therefore surprising that the 1,500-fold decrease in the catalytic efficiency of the D185L variant is due entirely to a corresponding decrease in k_{cat} , while the 10,500-fold reduction in the catalytic efficiency of eCBL-D185V is due in part to an increase in K_m^{L-Cth} as the enzyme is not saturated within the solubility limit of the substrate (Table 3.2; Figure 4.2). The kinetic parameters suggest that introduction of the larger leucine side chain triggers a shift in residues neighboring D185 that allows the PLP cofactor and the architecture of the substrate binding site to maintain a catalytically productive conformation. In contrast, the D185V substitution, similar to D185T, may push the cofactor into the substrate binding site, thereby altering active-site conformation in proximity to the catalytic base. Therefore, while it likely causes some rearrangement of residues near the base of the cofactor, the eCBL-D185L variant demonstrates that loss of the hydrogen bond to the pyridinium nitrogen results in a reduction in catalytic efficiency of only one order of magnitude beyond that of the eCBL-D185N substitution, which removes the negative charge of the aspartate side chain. Without energy minimization, which is beyond the scope of this project, the *in silico* mutagenesis cannot provide insight into the structural differences resulting from these substitutions.

Aspartate aminotransferase (AAT), which catalyzes the transfer of the amino group of aspartate to the carbon backbone of the oxoglutarate, forming glutamate, is the most thoroughly studied PLP-dependent enzyme (Brattnshtbin and Shemyakin, 1953; Metzler, Ikawa and Snell, 1954; Eliot and Kirsch, 2004). Like eCBL, AAT is also a member of fold

type I (Jansonius, 1998). Onuffer and Kirsch (1994) substituted residue D222 of AAT, which corresponds to eCBL-D185, with alanine (AAT-D222A) and reported a decrease in the enzyme catalytic efficiency 5 orders of magnitude. The crystal structure of the AAT-D222A variant created by Yano et al.(1992) showed a pair of water molecules residing in place of carboxyl group of aspartic acid which allowed the enzyme to preserve some activity (Onuffer and Kirsch 1994). These findings, in combination with the observation that the k_{cat} of the AAT-D222E variant is decreased by only 10-fold, while the K_m of aspartate is unchanged, confirm the necessity of a negatively charged residue interacting with the pyridinium nitrogen of the cofactor, in the context of AAT (Yano et al. 1992). Interestingly, substitution of D222 with asparagine decreases the catalytic efficiency of AAT by a further 10-fold beyond the reduction observed for ATT-D222A, an effect dominated by the 7000-fold decrease in the turnover rate (Yano et al. 1992). This contrasts with the 120-fold decrease in catalytic efficiency, also dominated by a reduction in k_{cat} , observed for eCBL-D185N and provides insight into the role of the PLP cofactor in the mechanisms of AAT and eCBL.

Griswold and Toney (2011) complemented the earlier D222 substitution studies by substituting the PLP cofactor of AAT with 1-deazapyridoxal 5'-phosphate (deazaPLP), which replaces the pyridinium nitrogen of the cofactor with a carbon atom. This approach avoids structural changes in the active site resulting from substitution of AAT-D222. The deazaPLP-AAT enzyme did not exhibit detectable activity within the solubility limit of the L-aspartate substrate, suggesting a reduction in the turnover rate of at least nine orders of magnitude (Griswold and Toney 2011). The contrast between the inactivity of deazaPLP-AAT and the five order-of-magnitude decrease in the catalytic efficiency of AAT-D222A

indicates that, while the electronegativity of the nitrogen atom itself can be enough to support catalysis, the adjacent carboxylate group of AAT-D222 promotes protonation of the pyridinium nitrogen of the cofactor. Protonation of the pyridinium nitrogen increases the electron withdrawing nature of the cofactor, thereby enhancing stabilization of the quinonoid intermediate of the AAT reaction (Griswold and Toney, 2011).

The identity of the residue positioned to interact with the pyridinium nitrogen of the cofactor varies with the fold-type of PLP-dependent enzymes (Hyde et al. 1988; Shaw et al. 1997; Burkhard et al. 1998). For example, *O*-acetylserine sulfhydrylase (OASS), which catalyzes the final reaction in cysteine biosynthesis, exemplifies fold-type II enzymes, in which this position is occupied by a serine residue (Burkhard et al. 1998; Cook and Wedding 1974). The β -elimination reaction of eCBL (fold-type I) is shared with the enzymes of fold-type II, as the first part of the β -replacement reactions common to this group of enzymes. In contrast with the four-to-five order of magnitude decreases in catalytic efficiency of AAT-D222A and eCBL-D185V (Table 3.2), substitution of the corresponding serine (S272) residue of OASS with alanine results in a decrease of less than 4-fold (Cook, 2003).

The mechanism of AAT includes a β -elimination step in which the proton at the β -position of the substrate is abstracted. β -elimination reactions, such as those catalyzed by AAT, CBL and OASS, can occur by a two-step or a one-step process, referred to as E1 and E2 mechanisms respectively. OASS follows an E2 mechanism, which does not require delocalization of the carbanion, formed following abstraction of the C α proton, via a quinonoid intermediate, which is stabilized by the presence of a negatively charged residue at the base of the PLP ring that promotes protonation of the pyridinium nitrogen (Tai and

Cook, 2001). This is demonstrated by the minor 4-fold decrease in catalytic efficiency of OASS-S272A (Cook, 2003). Similarly, in contrast with AAT, replacement of the PLP cofactor of OASS with deazaPLP decreases k_{cat} by only 260-fold, confirming that protonation of the pyridinium nitrogen is not required for the OASS β -elimination and even the presence of the pyridinium nitrogen is not strictly required for the reaction to proceed (Griswold and Toney, 2011). In contrast, the five and at least nine order of magnitude reductions in the catalytic efficiency of AAT-D222A and deazaPLP-AAT, respectively, demonstrates that this enzyme follows an E1 mechanism that requires stabilization of the carbanion via delocalization in the cofactor ring (Carey and Sundberg, 2007). The 120 and 1,500-fold decreases in k_{cat} of the eCBL-D185N and D185L variants, and the similarly to the 260-fold decrease in the k_{cat} of deazaPLP-OASS, compared with the drastic decreases observed for AAT, suggests that protonation of the pyridinium nitrogen of the PLP cofactor is not strictly required for the eCBL β -elimination reaction. As such, eCBL may be the first example of a fold-type I PLP enzyme that follows an E1, rather than E2, mechanism.

4.2. The D185-T187-A207 hydrogen-bonding network linking the PLP pyridinium nitrogen and the backbone of K210

In the resting state of the enzyme, the side chain of the catalytic base (K210) is covalently bound to the PLP cofactor, but during catalysis the α -amino group of the substrate replaces the ϵ -amino group of K210 in Schiff base linkage with the cofactor, freeing the latter for its role in catalysis. Positioning and tethering of the ϵ -amino group of catalytic base determines the range of proton transfers possible and is central to the ability of PLP enzymes to regulate reaction specificity (Aitken, Lodha and Morneau, 2011; Lodha and

Aitken, 2011). Tethering of the K210 side chain is accomplished by conserved residues S339 and Y56 of eCBL (Lodha and Aitken, 2011). The catalytic base is also positioned relative to the cofactor by hydrogen-bonding networks that define the shape of the active site, including one that connects the K210 backbone to the pyridinium nitrogen (Figure 4.3). The residues comprising this network are D185, T187 and A207. A hydrogen bond is present between the side chains of residues T187 and D185, which also forms a hydrogen bond to the pyridinium nitrogen of the cofactor. This network is extended to the catalytic base (K210) via hydrogen bonds between the side chain of T187 and backbone of A207 and between the latter and the backbone of K210. The contribution of this network to the L-Cth hydrolysis activity of eCBL was probed by conservatively substituting the central threonine at position 187 with serine and valine. The T187S substitution maintains the hydrogen-bonding capacity of T187, but alters its packing by removing the methyl group, while valine maintains the relative shape of the T187 side chain, by replacing the hydroxyl moiety with a methyl group, but is incapable of forming a hydrogen bond (Clausen et al., 1996).

Figure 4.3. Schematic representation of hydrogen-bonding network that bridges the pyridinium nitrogen of PLP and the backbone of K210, the catalytic base. All the distances presented in the PyMOL inset are in angstroms and correspond to the wavy lines in the schematic (Clausen et al., 1996, PDB ID: 1CL1).

Residue T187 packs against the ring of the eCBL cofactor, preventing it from pivoting within the active site (Clausen et al., 1996). This position is conserved as a threonine amongst the bacterial CBL enzymes, but in CBL from some plant species, such as *Arabidopsis thaliana*, it is replaced with serine. Uniquely, among the site-directed variants characterized in this study, the catalytic efficiency of eCBL-T187S is increased by 60% as its 3.3-fold decrease in k_{cat} is complemented by a 5-fold decrease of the K_m value (Table 3.2; Figure 3.4). In contrast, replacing threonine with valine decreases the catalytic efficiency by two orders of magnitude, an effect dominated by the 80-fold decrease in the turnover rate (k_{cat}), compared to the wild type enzyme.

Substitution of eCBL-T187 with serine ($73 \text{ \AA}^3$) has the effect of removing the side-chain methyl group of threonine ($93 \text{ \AA}^3$) and results in a 22% reduction in the van der Waals volume, which may allow neighbouring residues or the cofactor to shift in compensation within the active site. The residues within $5 \text{ \AA}$ of T187 include D185, A207, K210 and W340, which all interact with the PLP cofactor (Clausen et al., 1996). A shift of one or more of these residues may slightly alter the position of the cofactor, thereby allowing the substrate to bind with greater affinity and producing the observed 5-fold decrease in K_m^{L-Cth} . This is possible because the enzyme has evolved to be responsive to changes in the range of the typical intracellular concentration of its substrate rather than to maximize binding affinity.

The van der Waals volume of valine is $12 \text{ \AA}^3$ (13%) greater than that of threonine but the lack of hydroxyl group prevents it from forming hydrogen bonds (Schien, 1990; Darby and Creighton, 1993). The PyMOL modeling of eCBL-T187V suggests that van der Waals conflict between the additional methyl group of valine and the cofactor ring could result in

repositioning of the cofactor within the active site, perhaps to a less catalytically competent conformation that results in the observed 50-fold decrease in k_{cat} . While the weak, 3.5-Å, backbone-to-backbone hydrogen bond between T187 and D185 could potentially remain upon replacement of the former by valine, the T187V substitution eliminates the strong (2.7 Å) hydrogen bond between the side chains of these two residues and between the side chain of T187 and the backbone of A207 (Clausen et al., 1996). Loss of these interactions may also contribute to the observed 75-fold reduction in catalytic efficiency for T187V.

Bridging T187 and K210, via hydrogen bonds to its backbone, A207 is situated adjacent to the phosphate group and the ring of the PLP cofactor (Clausen et al., 1996). The alanine in this position is conserved among the bacterial CBL enzymes but it is replaced by serine in the closely related cystathionine γ -synthase (CGS) and cystathionine γ -lyase (CGL) (Messerschmidt et al., 2003; Clausen et al., 1998). In contrast with amino acids with larger side chains possessing aromatic, polar or charged groups, which can be substituted to probe the role of each element (e.g. the D185 substitution series), the side chain of alanine is a single methyl group. Therefore, substitution of alanine via site-directed mutagenesis tends to be more challenging to interpret as effects observed may be secondary, due to changes in packing. With this in mind, A207 was substituted with serine, threonine, asparagine and aspartate to probe the degree to which substitution is permitted at this position. The importance of a small, but not necessarily nonpolar, residue at position 207, in the context of the eCBL active-site, is demonstrated by the discovery that of the four A207 substitutions constructed, activity was observed for only A207S (Table 3.2). *In silico* modeling of the A207 substitution variants highlights areas of van der Waals overlap resulting from each replacement, ranging from modest (green) to a considerable degree

(red) of overlap of the substituted side chain with other active-site residues or the cofactor (Figure 4.4). Modest overlap can be overcome by a localized adjustment while larger overlaps may result in other residues adopting different conformations, thereby altering active-site architecture and/or dynamics.

Replacement of alanine ($67 \text{ \AA}^3$) with a polar serine results in only $6 \text{ \AA}^3$ increase in the van der Waals volume of the residue at position 207 and adds hydrogen bonding capacity to the side chain (Darby and Creighton, 1993; Berg, Tymoczko and Stryer, 2008). The eCBL-A207S rotamer with least structural overlap nudges the phosphate group of the PLP cofactor, which may be the cause of the 33% decrease in catalytic efficiency. The same rotamer shows the possibility of hydrogen bonds as the side chain hydroxyl moiety is within $2.5 \text{ \AA}$ and $3.5 \text{ \AA}$ of the PLP phosphate group and the side chain hydroxyl of T209 (Clausen et al., 1996).

The addition of even a single additional methyl group to serine (i.e. replacing A207 with threonine instead of serine), increases the van der Waals volume by additional $20 \text{ \AA}^3$ and eliminates eCBL activity, within the solubility limits of L-Cth. Similarly, no activity was detected for the aspartate ($91 \text{ \AA}^3$) and asparagine ($96 \text{ \AA}^3$) substitutions, which increase the van der Waals volume of A207 by 36 and 43%, respectively, and overlap with both the PLP ring and phosphate group. (Figure 4.4).

Figure 4.4. PyMOL representation of the van der Waals overlap between the PLP cofactor and the A207 substitutions (Clausen et al., 1996, PDB ID: 1CL1). The disks represent gradient of increasing degrees of overlap, from green to red.

Substitution of A207 with aspartate, asparagine or threonine also reduced the stability of eCBL, resulting in truncation of the enzyme, as evidenced by the presence of two bands on SDS-PAGE following affinity purification (Figure 3.1). The wild-type eCBL enzyme presents a single band of 45 kDa (Figure 3.1), consistent with the 44.95 kDa molecular weight of the eCBL monomer and the 1.75-kDa, N-terminal 6-His tag and linker (Farsi et al. 2009). While this 45-kDa band is present for all purified variants, 14 presented a second band at approximately 42 kDa (Table 3.1; Figure 3.2). The observed truncation of 14 of the 20 variants, including A207D, A207N and A207T, is likely caused by a change in active-site dynamics, exposing a protease cleavage site that is otherwise protected in the native enzyme.

The eCBL active site is located in the cleft at the interface between the catalytic and C-terminal domains and shielded from solvent by the PLP cofactor and a segment of the C-terminus and a portion of the small N-terminal domain from the adjacent monomer of the catalytic dimer (Clausen et al., 1996). The N-terminus of the eCBL monomer also forms part of the interface between the pair of catalytic dimers that come together to form the homotetrameric enzyme structure. Therefore, while a shift in active-site dynamics could expose either end of the eCBL monomer, a change at the N-terminus may also destabilize the quaternary structure of eCBL, resulting in the observed low yields of purified enzyme, lack of activity and truncation (Tables 3.1 and 3.2; Figures 3.1 and 3.3). With the goal of detecting which end of eCBL is destabilized and cleaved in the variants displaying the 42-kDa band, affinity chromatography of the purified eCBL-A207T variant was performed under denaturing conditions (Bornhorst and Falke, 2000). Purification under denaturing conditions ensures that any monomers lacking the 6-His tag will not bind to the Ni-NTA

resin by association with a monomer possessing the affinity tag. The observed early elution of the truncated 42-kDa band, in contrast with elution of the full-length, 45-kDa protein only following addition of the final elution step (Figure 3.3) supports the location of the truncation in the N-terminal section of eCBL.

Figure 4.5 shows the sequence of eCBL with the 6-His tag and the linker attached with possible cleavage sites for *E.coli* proteases I-VII (Gottesman, 1996). The underline sequence represents regions where a cleavage could occur to produce the protein that has a molecular weight of about 42 kDa. If the cut were to occur in the N-terminal region, the exiting polypeptide would have a molecular weight of 2.6-3.3 kDa, whereas if the cleavage occurred at the C-terminal region, a polypeptide weighing 2.7-3.4 kDa would emerge.

The chromatography under denaturing conditions suggests that the cleavage occurs at the N-terminal end of the protein, residues of the 9-15 region is on the dimer interface, although the crystal structure does not show any hydrogen bonds between the monomers (Clausen et al., 1996). Improper folding caused by the structural changes following the cleavage in this region could lead to difficulties with binding of the two monomers and subsequent loss of activity. While the figure only shows representation of proteases, the cleavage after eCBL-R15 residue, which is a cleavage site of Protease II, would result in a truncated protein of 41.65 kDa. (Pacaud, 1978). The C-terminal region of eCBL 365-371 residues, however are on the protein surface, allowing easy access for proteases but likely not causing major structural changes (Clausen et al., 1996).

His-tag *linker*
H H H H H H S S S G H I E G R M A D K K L D T Q L V N A G R S K K Y T L G A V N
S V I Q R A S S L V F D S V E A K K H A T R N R A N G E L E Y G R R G T L T H
F S L Q Q A M C E L E G G A G C V L F P C G A A A V A N S I L A F I E Q G D H
V L M T N T A Y E P S Q D F C S K I L S K L G V T T S W F D P L I G A D I V K
H L Q P N T K I V F L E S P G S I T M E V H D V P A I V A A V R S V V P D A I
I M I D N T W A A G V L F R A L D F G I D V S I Q A A T R Y L V G H S D A M I
G T A V C N A R C W E Q L R E N A Y L M G Q M V D A D T A Y I T S R G L R T L
G V R L R Q H H E S S L K V A E W L A E H P Q V A R V N H P A L P G S K G H E
F W R R D F T G S S G L F S F V L R R K L N N E E L A N Y L D N F S L F S M A
Y S W G G Y E S L I L A N Q P E H I A A I R P Q G E I D F S G T L I R L H I G
L E D V D D L I A D L D A G F A R I V

Figure 4.5. The eCBL sequence with the possible protease I-VII cleavage sites. The amino acids highlighted in green represent P1 sites of proteases, the amino acid after which the cleavage occurs (Song et al. 2012); underline residues represent the regions in which protein cleavage would result in the protein of molecular weight of about 42kDa.

While the overall protein backbones of eCBL and eCGS adopt the same folded structure, the pair of short segments at the N-terminus (residues 36-56 of eCBL) and C-terminus (residues 347-372 of eCBL), regions I and II, respectively, that shield the active-site adopt distinct conformations in these two enzymes (Clausen et al., 1996). With the goal of exploring whether these regions are determinants of reaction specificity, Manders and colleagues (2013) created chimeric constructs, by exchanging regions I and II of eCBL and eCGS and observed that both eCBL-based constructs were sufficiently stable to purify and retained eCBL activity but had no observable eCGS activity. Replacement of eCBL regions I or II with those of eCGS reduced L-Cth hydrolysis to by 43,000 and 240-fold, respectively (Manders et al., 2013). Interestingly, while this suggests a greater role for the N-terminal, region I in the eCBL active site, no truncation was observed for the eCBL-based chimeric constructs (Manders et al., 2013). This suggests that while substitution of region I does not destabilize the N-terminus, or the overall structure of eCBL, substitutions that interfere

with PLP binding may result in a shift in active-site dynamics, destabilizing the N-terminal domain as a result.

4.3. Residues flanking K210

The side chain of the catalytic base (K210) is also positioned by the interactions of the residues flanking it in primary sequence, T209 and Y211. A hydrogen bond between the phosphate group of PLP and T209 anchors the neighboring K210 with respect to the cofactor and the aromatic side chain of Y111 restricts movement of the K210 side chain away from the cofactor following substrate binding (Clausen et al., 1996). Two variants of T209 were constructed to probe the role of this conserved residue in the hydrolysis of L-Cth by eCBL. Similar to the aspartate, asparagine and threonine substitutions of A207, replacement of T209 with valine or asparagine resulted in a low yield of enzyme, compared to the wild-type eCBL, and eliminated L-Cth hydrolysis activity (Table 3.1).

PyMOL modeling of the valine and asparagine substitutions of T209 was performed in the context of the resting state of eCBL and the external aldimine complex of eCBL with the N-hydrazinocarbonylmethyl-2-Nitro-Benzamide inhibitor, in which the catalytic base is released from Schiff base linkage with the cofactor (Clausen et al., 1996; Ejim et al., 2007). The rotamers with the lowest van der Waals overlap are presented in Figure 4.6.

Replacing T209 with valine increases the volume by $12 \text{ \AA}^3$ and eliminates its hydrogen-bonding capacity, thereby removing the strong ($2.4 \text{ \AA}$) interaction with the phosphate group of the cofactor (Darby and Creighton, 1993; Clausen et al., 1996). Modeling of the T209N substitution, which retains similar volume ($3 \text{ \AA}^3$ greater volume of asparagine) and

hydrogen-bonding capacity, suggests that the distance to the PLP phosphate is increased to 3.1 Å, decreasing its strength and/or potentially shifting the position of the adjacent K210 and likely repositions the cofactor (Darby and Creighton, 1993). The resulting lack of the measurable activity in both T209N and T209V variants supports the interpretation of a shift in the positioning between the catalytic base and the cofactor, identifying a key role for T209 in the maintenance of productive active-site architecture.

Figure 4.6. Comparison of PyMOL models of (A and B) T209V and (C and D) T209N in the context of the (A and C) the eCBL-*N*-hydrazinocarbonylmethyl-2-Nitro-Benzamide complex, in which the PLP cofactor and *N*-hydrazinocarbonylmethyl-2-Nitro-Benzamide inhibitor are shown in Schiff base linkage and (B and D) unliganded eCBL structure, where the PLP cofactor is bound to K210 (Clausen et al., 1996, PDB ID: 1CL1; Ejim et al., 2007, PDB ID: 2GQN). The disks represent gradient of increasing degrees, of van der Waals overlap between the protein and the substituted residues at position 209, with green indicating minimal overlap and red highlighting structural clashes.

The reduced stability of the T209 variants and the majority of the A207 substitutions, as indicated by the presence of the 42-kDa band on SDS-PAGE, is likely due to their effect on PLP binding within the active site and demonstrates that the cofactor is required not only for catalytic activity but also for enzyme stability. A similar effect has been observed for the fold-type I enzymes serine hydroxymethyltransferase and AAT and fold-type II representatives OASS and tryptophan synthase (Bartholmes, Balk and Kirschner, 1980; Artigues et al., 1994; Bettati et al., 2000; Deu and Kirsch, 2007) and is explained by the stabilizing effect of ligand binding on protein stability (Ross and Subramanian 1981; Niesen, Berglund and Vedadi, 2007).

The catalytic base (K210) of eCBL is also flanked by Y211, a position conserved among the enzymes of the γ -family of fold-type I as tyrosine or phenylalanine. The role of the hydroxyl group of eCBL-Y211, which forms hydrogen bonds to the side chain of histidine 374 and the backbone of glycine 309 (Clausen et al., 1996), was probed by substitution with phenylalanine. As the position of the Y211 backbone is influenced by interaction with the side chain of R262 (Figure 4.7), the R262K and R262L were constructed to explore the contribution of this residue to the L-Cth hydrolysis activity of eCBL.

Figure 4.7. The hydrogen bonding interactions of Y211 and R262. (A) R262 forms hydrogen bonds with G255, T256, Q265, H266, Y211 and L212. (B) Substitution of R262 with lysine eliminates the hydrogen bonds to Y211, L212, G255 and T256. (C) Substitution of R262 with leucine eliminates all but backbone hydrogen bonds. All distances indicated are in Å (Clausen et al., 1997, PDB ID: 1CL2).

While the conservative valine and asparagine substitutions of T209 with asparagine and valine eliminated activity and reduced stability, removal of the Y211 hydroxyl moiety, by replacement with phenylalanine, results in only a 15-fold decrease in k_{cat} . The difference in the effect of substitution of T209 and Y211, which both flank the catalytic base (K210), provides insight in the specific role of each residue. While T209 positions K210 with respect to the PLP, the Y211 side chain is directed away from the cofactor. The 15-fold reduction in k_{cat} observed for eCBL-Y211F suggests that the bulk of the aromatic side chain of this residue may play a greater role in positioning the catalytic base than the hydrogen bond formed by the side-chain hydroxyl moiety. Substitution of Y211 with a nonaromatic residue, such as leucine, may provide further insight on the role of the Y211 side chain in eCBL activity.

Arginine, with a van der Waals volume of $148 \text{ \AA}^3$, is second only to tryptophan ($163 \text{ \AA}^3$) in size. Other distinctive feature of the arginine side chain is the resonance stabilization of the guanidinium group that increases its pK_a to 12.5, such that it is positively charged, and its ability to form several hydrogen bonds (Darby and Creighton, 1993; Berg, Tymoczko and Stryer, 2008). This is illustrated by the hydrogen bonds between the side chain of eCBL-R262 and residues Y211, L212, G255 and T256 (Ejim et al., 2007). As a result, this arginine residue anchors several active-site residues with respect to each other and may play a role in defining the architecture of the active site. The effect of R262 substitution is, therefore, complex and difficult to interpret. This is compounded by the fact that the most conservative substitutions possible are lysine and leucine, which reduce the van der Waals volume by $13 \text{ \AA}^3$ and $24 \text{ \AA}^3$, respectively (Figure 4.7).

The R262L variant, which eliminates the charge and hydrogen-bonding capacity of arginine, exhibits a 42-kDa band, indicating truncation, and is not active. The 24-Å³ smaller in van der Waals volume of leucine, and accompanying shift in packing interactions, in combination with loss of all four hydrogen bonds formed by the R262 side chain likely both contribute to the reduced stability and loss of enzyme activity of this variant (Figure 4.7). Further refinement of the specific contribution of each of these hydrogen bonds would require construction of site-directed variants of the other three hydrogen-bonding partners (L212, G255 and T256) of R262.

In contrast with the loss of activity upon substitution of eCBL-R262 with leucine, the R262K variant retains eCBL activity, despite a 15-fold reduction in the turnover rate. This enzyme is unique as the steady state kinetic plot of activity versus substrate concentration is sigmoidal in shape rather than the rectangular hyperbola described by the Michaelis-Menten equation (Figure 3.6). The sigmoidal shape of the curve indicates positive cooperativity between active sites, suggesting that binding of substrate to one active site of the catalytic dimer results in a conformational change that increases substrate affinity in the other active site of the pair (Berg, Tymoczko and Stryer, 2008). Cooperativity has not previously been observed for wild-type eCBL or any of the 35 site-directed variants of this enzyme reported by our lab to date. The *in silico* modeling of eCBL-R262K identified 13 different possible rotamers, as a long side chain allows for a degree of flexibility, none of which can sustain the hydrogen bonds observed for R262. This suggests that the more drastic effect of substitution of R262 with leucine may be due to changes in packing interactions. However, it is also possible that neighboring residues shift in conformation to accommodate substitution of this arginine, thereby enabling some of the hydrogen bonds

formed by R262 to be maintained by a lysine in this position. The structure chosen for this *in silico* mutagenesis for the R262K variant was one with the AVG inhibitor in Schiff-base complex with PLP as there is no structure available with substrate bound at the active site and AVG is similar in structure to L-Cth (Clausen et al., 1997). The side chain of the lysine rotamer with the lowest degree of structural overlap adopts a different conformation from R262 and is within 4 Å of residues R257, T258 and the corresponding R262 (replaced with lysine in this R262 substitution variant) of the C-terminal domain of the adjacent monomer of the catalytic dimer (Clausen et al., 1997). The possibility of interactions with these residues or the loss of hydrogen bonds to Y211, L212, G255 and/or T256A may result in the observed change in cooperativity of the R262K variant, compared to wild-type eCBL.

4.4. eCBL residues W340 and L310

The indole nitrogen of the W340 side chain is within hydrogen-bonding distance of O3' (the hydroxyl group) of the cofactor and the α -carboxylate group of the aminoethoxyvinylglycine (AVG) inhibitor bound in Schiff base linkage with the PLP in the structure of the eCBL-AVG complex (Clausen et al., 1997). Analysis of the eCBL-W340F and R372A/K/L variants demonstrated that although the interaction of W340 with the α -carboxylate of the substrate or inhibitor contributes to substrate and AVG binding, it overshadowed by the salt bridge between the R372 guanidinium group and the α -carboxylate group of L-Cth or AVG (Lodha, Jaworski and Aitken, 2010). While a tryptophan in position 340 is conserved among bacterial CBL enzymes, in bacterial CGS and fungal CGL enzymes it is replaced with a conserved leucine. To explore the role of the

aromatic component of the W340 side chain in the hydrolysis of L-Cth by eCBL, this residue was replaced with alanine, leucine and histidine.

Replacement of the aromatic W340 side chain with the largest aliphatic amino acid, leucine ($124 \text{ \AA}^3$), resulted in a loss of six orders of magnitude in catalytic efficiency (Table 3.1). Similarly, substitution with aliphatic alanine ($67 \text{ \AA}^3$), which is 2.4-fold smaller than aromatic tryptophan ($163 \text{ \AA}^3$) eliminates eCBL activity (Table 3.1; Darby and Creighton, 1993). In contrast, the W340H substitution, although $45 \text{ \AA}^3$ smaller in van der Waals volume, preserves the aromatic character of the residue and the catalytic efficiency of this variant is reduced by only 63-fold. Additionally, while substitution of W340 with histidine results in the presence of a 42-kDa band, not reported for W340F, this band is more pronounced for the W340A and W340L variants (Table 3.1).

The tryptophan at position 340 is adjacent in sequence to S339, which is a conserved residue among the enzymes of the γ -subfamily of fold type I. The hydroxyl side chain of eCBL-S339 is thought to act as a determinant of reaction specificity by tethering the ϵ -amino group of the catalytic base, following its release from the Schiff-base linkage with the cofactor (Figure 4.8; Lodha and Aitken, 2011). While a serine in position 339 is conserved, its influence is thought to be modulated by interaction with neighboring residues (Manders et al., 2013; Lodha and Aitken, 2011). Without energy minimization, the value of *in silico* modeling for interpretation of the W340A and W340L variants is of limited value as substitution of the $96\text{-}\text{\AA}^3$ indole side chain of tryptophan with the more flexible aliphatic leucine, or the methyl group of alanine, likely causes significant reorganization within the active site. The position of W340 is influenced by packing interactions of L310 (Figure 1.7).

A leucine is not a conserved residue with yCGL having methionine in place of leucine. The 15-fold decrease in catalytic efficiency observed upon substitution of eCBL-L310 with alanine may reflect the importance of W340 positioning for eCBL activity. Given that position 340 is conserved as leucine in both CGS and CGL, while substitution of eCBL-W340 with leucine reduces the catalytic efficiency of eCBL by six orders of magnitude and replacement with aromatic histidine or phenylalanine decreases catalytic efficiency by less than two orders of magnitude, future studies exploring the role of eCBL-W340 as a determinant of reaction specificity are warranted.

Figure 4.8. The relative positions of K210, S339 and W340 in the structure of the eCBL-AVG complex (Ejim et al. 2007, PDB ID: 2GQN). The hydrogen bond distances are shown in Å.

4.5. Conclusion

The goal of this study was to characterize the roles of 8 residues that influence the positioning and properties of the eCBL catalytic base (K210) and the PLP cofactor in the L-Cth hydrolysis activity of this enzyme. The residues that were targeted include those flanking K210 (T209 and Y211), the members of the hydrogen-bonding network that links the pyridinium nitrogen of the cofactor with the backbone of K210 (D185, T187 and A207) and residues that pack against (A207) or form hydrogen bonds with (D185, T209 and W340) the cofactor. Residues R262 and L310, which influence the positioning of Y211 and W340, respectively, were also selected for investigation.

A 42-kDa band, indicative of truncation, likely near the N-terminal of the eCBL polypeptide, was observed for 14 of the 20 site-directed variants investigated. Similar truncation has not been previously reported for site-directed variants of eCBL and, given that most affected variants correspond to substitutions of residues interacting with the cofactor (A207, T209 and W340), they may reflect a reduction in stability and/or shift in active-site dynamics that exposes a protease cleavage site that is protected in the wild-type enzyme.

The results described here demonstrate that the presence of an acidic residue (D185) interacting with the pyridinium nitrogen of the PLP, is less important for the β -elimination catalyzed by eCBL than for aminotransferase activity of the model PLP enzyme AAT. This suggests that the carbanion formed by abstraction of the α -proton of the L-Cth substrate of eCBL may not be stabilized via a quinonoid intermediate and that the reaction may proceed via an E2 mechanism, similar to β -elimination catalyzed by OASS. If this is the case, eCBL will be the first example of an E2 mechanism in fold-type I and the acidic residue at

position D185 of eCBL may be conserved primarily for its role in positioning of the PLP cofactor.

The intolerance of A207 and T209 to substitution highlights the role of these residues in the maintenance of active-site architecture, PLP binding and likely also in positioning of the catalytic base with respect to the cofactor. This study demonstrated that an aromatic residue at position 340 of eCBL is required and provides the foundation for further studies exploring the potential role of this position as a determinant of reaction specificity in CBL, CGS and CGL.

The PyMOL computational modeling suggests that the variants that did not show the activity were those with the greatest degree of van der Waals overlap, particularly with the cofactor. Adaptive changes during evolution have balanced protein stability and flexibility, required for activity, in a three-dimensional network of interactions such that any substitution requires accommodation in packing and polar interactions in order to maintain a stable, folded structure. While many of the residues discussed here, such as S339, T209 and D185, are not unique to eCBL (i.e. they may also be conserved in other PLP enzymes) they may still act as determinants of substrate or reaction specificity due to the influence of these interaction networks. Future work will build on the results of this study to explore the role of these residues, particularly W340, as determinants of specificity.

REFERENCES

- Aitken, S.M. and Kirsch, J.F., 2003. Kinetics of the yeast cystathionine β -synthase forward and reverse reactions: continuous assays and the equilibrium constant for the reaction. *Biochemistry*, 42(2), pp.571-578.
- Aitken, S.M., Lodha, P.H. and Morneau, D.J., 2011. The enzymes of the transsulfuration pathways: active-site characterizations. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, 1814(11), pp.1511-1517.
- Alexander, F.W., Sandmeier, E., Mehta, P.K. and Christen, P., 1994. Evolutionary relationships among pyridoxal- 5'- phosphate- dependent enzymes. *The FEBS Journal*, 219(3), pp.953-960.
- Artigues, A., Iriarte, A. and Martinez-Carrion, M., 1994. Acid-induced reversible unfolding of mitochondrial aspartate aminotransferase. *Journal of Biological Chemistry*, 269(35), pp.21990-21999.
- Auger, S., Danchin, A. and Martin-Verstraete, I., 2002. Global expression profile of *Bacillus subtilis* grown in the presence of sulfate or methionine. *Journal of bacteriology*, 184(18), pp.5179-5186.
- Bach, R.D., Canepa, C. and Glukhovtsev, M.N., 1999. Influence of Electrostatic Effects on Activation Barriers in Enzymatic Reactions: Pyridoxal 5 '-Phosphate- Dependent Decarboxylation of α -Amino Acids. *Journal of the American Chemical Society*, 121(28), pp.6542-6555.
- Bartholmes, P., Balk, H. and Kirschner, K., 1980. Mechanism of reconstitution of the apo. beta. 2 subunit and the. alpha. 2apo. beta. 2 complex of tryptophan synthase with pyridoxal 5'-phosphate: kinetic studies. *Biochemistry*, 19(19), pp.4527-4533.
- Belfaiza, J., Parsot, C., Martel, A., De La Tour, C.B., Margarita, D., Cohen, G.N. and Saint-Girons, I., 1986. Evolution in biosynthetic pathways: two enzymes catalyzing consecutive steps in methionine biosynthesis originate from a common ancestor and possess a similar regulatory region. *Proceedings of the National Academy of Sciences*, 83(4), pp.867-871.
- Ben-Bassat, A., Bauer, K., Chang, S.Y., Myambo, K., Boosman, A. and Chang, S., 1987. Processing of the initiation methionine from proteins: properties of the *Escherichia coli* methionine aminopeptidase and its gene structure. *Journal of bacteriology*, 169(2), pp.751-757.
- Berg, J.M., Tymoczko, J.L. and Stryer, L., 2008. *Biochemistry 6th edition*. Macmillan. New York
- Berman, H.M., J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E., 2000, The Protein Data Bank. *Nucleic Acids Research*, 28, pp.235-242

Bigelow, D.J. and Squier, T.C., 2005. Redox modulation of cellular signaling and metabolism through reversible oxidation of methionine sensors in calcium regulatory proteins. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, 1703(2), pp.121-134.

Bornhorst, J.A. and Falke, J.J., 2000. Purification of proteins using polyhistidine affinity tags. *Methods in enzymology*, 326, pp.245-254.

Brattnshtbin, A. and Shemyakin, M.M., 1953. Theory of the process of amino-acid metabolism, catalysed by pyridoxal enzymes. *Biokhimiya*, 18, pp.393-411.

Brzovic, P., Holbrook, E.L., Greene, R.C. and Dunn, M.F., 1990. Reaction mechanism of *Escherichia coli* cystathionine. gamma.-synthase: direct evidence for a pyridoxamine derivative of vinylgloxylate as a key intermediate in pyridoxal phosphate dependent. gamma.-elimination and. gamma.-replacement reactions. *Biochemistry*, 29(2), pp.442-451.

Burkhard, P., Rao, G.J., Hohenester, E., Schnackerz, K.D., Cook, P.F. and Jansonius, J.N., 1998. Three-dimensional structure of O-acetylserine sulfhydrylase from *Salmonella typhimurium*1. *Journal of molecular biology*, 283(1), pp.121-133

Carey, F.A. and Sundberg, R.J., 2007. *Advanced organic chemistry: part A: structure and mechanisms*. Springer Science & Business Media.

Christen, P. and Mehta, P.K., 2001. From cofactor to enzymes. The molecular evolution of pyridoxal- 5'- phosphate- dependent enzymes. *The Chemical Record*, 1(6), pp.436-447.

Clausen, T., Huber, R., Laber, B., Pohlenz, H.D. and Messerschmidt, A., 1996. Crystal Structure of the Pyridoxal-5'- phosphate Dependent Cystathionine β -lyase from *Escherichia coli* at 1.83 Å. *Journal of molecular biology*, 262(2), pp.202-224.

Clausen, T., Huber, R., Messerschmidt, A., Pohlenz, H.D. and Laber, B., 1997. Slow-binding inhibition of *Escherichia coli* cystathionine β -lyase by L-aminoethoxyvinylglycine: a kinetic and X-ray study. *Biochemistry*, 36(41), pp.12633-12643.

Clausen, T., Huber, R., Prade, L., Wahl, M.C. and Messerschmidt, A., 1998. Crystal structure of *Escherichia coli* cystathionine γ - synthase at 1.5 Å resolution. *The EMBO journal*, 17(23), pp.6827-6838.

Cook, P.F. and Wedding, R.T., 1976. A reaction mechanism from steady state kinetic studies for O-acetylserine sulfhydrylase from *Salmonella typhimurium* LT-2. *Journal of Biological Chemistry*, 251(7), pp.2023-2029.

Cook, P.F., 2003. α , β -Elimination reaction of O-acetylserine sulfhydrylase. Is the pyridine ring required?. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, 1647(1-2), pp.66-69.

Cramer, A., Raillard, S.A., Bermudez, E. and Stemmer, W.P., 1998. DNA shuffling of a family of genes from diverse species accelerates directed evolution. *Nature*, 391(6664), p.288.

Darby, N.J. and Creighton, T.E., 1993. *Protein structure*. Oxford University Press, USA.

DeLano, W.L., 2002. Pymol: An open-source molecular graphics tool. *CCP4 Newsletter On Protein Crystallography*, 40, pp.82-92.

Della Ragione, F., Porcelli, M., Carteni-Farina, M., Zappia, V. and Pegg, A.E., 1985. Escherichia coli S-adenosylhomocysteine/5'-methylthioadenosine nucleosidase. Purification, substrate specificity and mechanism of action. *Biochemical Journal*, 232(2), pp.335-341.

Deu, E. and Kirsch, J.F., 2007. Cofactor-directed reversible denaturation pathways: the cofactor-stabilized Escherichia coli aspartate aminotransferase homodimer unfolds through a pathway that differs from that of the apoenzyme. *Biochemistry*, 46(19), pp.5819-5829.

Dong, R., Peng, Z., Zhang, Y. and Yang, J., 2017. mTM-align: an algorithm for fast and accurate multiple protein structure alignment. *Bioinformatics*, p.btx828.

Drennan, C.L., Huang, S., Drummond, J.T., Matthews, R.G. and Lidwig, M.L., 1994. How a protein binds B12: A 3.0 Å X-ray structure of B12-binding domains of methionine synthase. *Science*, 266(5191), pp.1669-1674.

Dunathan, H.C., 1966. Conformation and reaction specificity in pyridoxal phosphate enzymes. *Proceedings of the National Academy of Sciences*, 55(4), pp.712-716.

Ejim, L.J., Blanchard, J.E., Koteva, K.P., Sumerfield, R., Elowe, N.H., Chechetto, J.D., Brown, E.D., Junop, M.S. and Wright, G.D., 2007. Inhibitors of bacterial cystathionine β -lyase: leads for new antimicrobial agents and probes of enzyme structure and function. *Journal of medicinal chemistry*, 50(4), pp.755-764.

Eliot, A.C. and Kirsch, J.F., 2004. Pyridoxal phosphate enzymes: mechanistic, structural, and evolutionary considerations. *Annual review of biochemistry*, 73(1), pp.383-415.

Farsi, A., Lodha, P.H., Skanes, J.E., Los, H., Kalidindi, N. and Aitken, S.M., 2009. Interconversion of a pair of active-site residues in Escherichia coli cystathionine γ -synthase, E. coli cystathionine β -lyase, and Saccharomyces cerevisiae cystathionine γ -lyase and development of tools for the investigation of their mechanisms and reaction specificity. *Biochemistry and Cell Biology*, 87(2), pp.445-457.

Flavin, M. and Slaughter, C., 1967. Enzymatic synthesis of homocysteine or methionine directly from O-succinyl-homoserine. *Biochimica et Biophysica Acta (BBA)-Enzymology*, 132(2), pp.400-405.

Gilbert, H.F., 1995. Thiol/disulfide exchange equilibria and disulfide bond stability. *Methods in enzymology*, 251, pp.8-28.

Giovanelli, J. and Mudd, S.H., 1967. Synthesis of homocysteine and cysteine by enzyme extracts of spinach. *Biochemical and biophysical research communications*, 27(2), pp.150-156.

González, J.C., Peariso, K., Penner-Hahn, J.E. and Matthews, R.G., 1996. Cobalamin-independent methionine synthase from *Escherichia coli*: a zinc metalloenzyme. *Biochemistry*, 35(38), pp.12228-12234.

Gophna, U., Baptiste, E., Doolittle, W.F., Biran, D. and Ron, E.Z., 2005. Evolutionary plasticity of methionine biosynthesis. *Gene*, 355, pp.48-57.

Gottesman, S., 1996. Proteases and their targets in *Escherichia coli*. *Annual review of genetics*, 30(1), pp.465-506.

Gyorgy, P., 1934. Vitamin B2 and the Pellagra-like Dermatitis in Rats. *Nature*, 133(3361), pp.498-499.

Hayes, J.D. and McLellan, L.I., 1999. Glutathione and glutathione-dependent enzymes represent a co-ordinately regulated defence against oxidative stress. *Free radical research*, 31(4), pp.273-300.

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(1), pp.51-59.

Hopwood, E.M., Ahmed, D. and Aitken, S.M., 2014. A role for glutamate-333 of *Saccharomyces cerevisiae* cystathionine γ -lyase as a determinant of specificity. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, 1844(2), pp.465-472.

Hoshi, T. and Heinemann, S.H., 2001. Regulation of cell function by methionine oxidation and reduction. *The Journal of Physiology*, 531(1), pp.1-11.

Hyde, C.C., Ahmed, S.A., Padlan, E.A., Miles, E.W. and Davies, D.R., 1988. Three-dimensional structure of the tryptophan synthase α 2 β 2 multienzyme complex from *Salmonella typhimurium*. *Journal of Biological Chemistry*, 263(33), pp.17857-17871.

Jansonius, J.N., 1998. Structure, evolution and action of vitamin B6-dependent enzymes. *Current opinion in structural biology*, 8(6), pp.759-769.

Jaworski, A.F., Lodha, P.H., Manders, A.L. and Aitken, S.M., 2012. Exploration of the active site of *Escherichia coli* cystathionine γ - synthase. *Protein Science*, 21(11), pp.1662-1671.

- Jaworski, A.F. and Aitken, S.M., 2013. Exploration of the six tryptophan residues of *Escherichia coli* cystathionine β -lyase as probes of enzyme conformational change. *Archives of biochemistry and biophysics*, 538(2), pp.138-144.
- Jhee, K.H., McPhie, P. and Miles, E.W., 2000. Domain architecture of the heme-independent yeast cystathionine β -synthase provides insights into mechanisms of catalysis and regulation. *Biochemistry*, 39(34), pp.10548-10556.
- Kendrew, J.C., Bodo, G., Dintzis, H.M., Parrish, R.G., Wyckoff, H. and Phillips, D.C., 1958. A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature*, 181(4610), pp.662-666.
- Kery, V., Bukovska, G. and Kraus, J.P., 1994. Transsulfuration depends on heme in addition to pyridoxal 5'-phosphate. Cystathionine beta-synthase is a heme protein. *Journal of Biological Chemistry*, 269(41), pp.25283-25288.
- Kiene, R.P., Linn, L.J., González, J., Moran, M.A. and Bruton, J.A., 1999. Dimethylsulfoniopropionate and methanethiol are important precursors of methionine and protein-sulfur in marine bacterioplankton. *Applied and environmental microbiology*, 65(10), pp.4549-4558.
- Kirk, O., Borchert, T.V. and Fuglsang, C.C., 2002. Industrial enzyme applications. *Current opinion in biotechnology*, 13(4), pp.345-351.
- Koutmos, M., Datta, S., Patridge, K.A., Smith, J.L. and Matthews, R.G., 2009. Insights into the reactivation of cobalamin-dependent methionine synthase. *Proceedings of the National Academy of Sciences*, 106(44), pp.18527-18532.
- Limbach, H.H., Chan-Huot, M., Sharif, S., Tolstoy, P.M., Shenderovich, I.G., Denisov, G.S. and Toney, M.D., 2011. Critical hydrogen bonds and protonation states of pyridoxal 5'-phosphate revealed by NMR. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, 1814(11), pp.1426-1437.
- Lodha, P.H., Jaworski, A.F. and Aitken, S.M., 2010. Characterization of site-directed mutants of residues R58, R59, D116, W340 and R372 in the active site of *E. coli* cystathionine β -lyase. *Protein Science*, 19(3), pp.383-391.
- Lodha, P.H. and Aitken, S.M., 2011. Characterization of the side-chain hydroxyl moieties of residues Y56, Y111, Y238, Y338, and S339 as determinants of specificity in *E. coli* cystathionine β -lyase. *Biochemistry*, 50(45), pp.9876-9885.
- Manders, A.L., Jaworski, A.F., Ahmed, M. and Aitken, S.M., 2013. Exploration of structure–function relationships in *Escherichia coli* cystathionine γ -synthase and cystathionine β -lyase via chimeric constructs and site-specific substitutions. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, 1834(6), pp.1044-1053.
- Meister, A., 1988. Glutathione metabolism and its selective modification. *Journal of biological chemistry*, 263(33), pp.17205-17208.

Messerschmidt, A., Worbs, M., Steegborn, C., Wahl, M.C., Huber, R., Laber, B. and Clausen, T., 2003. Determinants of Enzymatic Specificity in the Cys-Met-Metabolism PLP-Dependent Enzyme Family: Crystal Structure of Cystathionine γ -Lyase from Yeast and Intrafamilial Structure Comparison. *Biological chemistry*, 384(3), pp.373-386.

Metzler, D.E., Ikawa, M. and Snell, E.E., 1954. A general mechanism for vitamin B6-catalyzed reactions. *Journal of the American Chemical Society*, 76(3), pp.648-652.

Miller, C.G., Strauch, K.L., Kukral, A.M., Miller, J.L., Wingfield, P.T., Mazzei, G.J., Werlen, R.C., Graber, P. and Movva, N.R., 1987. N-terminal methionine-specific peptidase in *Salmonella typhimurium*. *Proceedings of the National Academy of Sciences*, 84(9), pp.2718-2722.

Niesen, F.H., Berglund, H. and Vedadi, M., 2007. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nature protocols*, 2(9), p.2212.

Onuffer, J.J. and Kirsch, J.F., 1994. Characterization of the apparent negative cooperativity induced in *Escherichia coli* aspartate aminotransferase by the replacement of Asp222 with alanine. Evidence for an extremely slow conformational change. *Protein Engineering, Design and Selection*, 7(3), pp.413-424.

Pacaud, M., 1978. Protease II from *Escherichia coli*. *The FEBS Journal*, 82(2), pp.439-451.

Percudani, R. and Peracchi, A., 2003. A genomic overview of pyridoxal- phosphate-dependent enzymes. *EMBO reports*, 4(9), pp.850-854.

Ravanel, S., Gakière, B., Job, D. and Douce, R., 1998. The specific features of methionine biosynthesis and metabolism in plants. *Proceedings of the National Academy of Sciences*, 95(13), pp.7805-7812.

Ross, P.D. and Subramanian, S., 1981. Thermodynamics of protein association reactions: forces contributing to stability. *Biochemistry*, 20(11), pp.3096-3102.

Schein, C.H., 1990. Solubility as a function of protein structure and solvent components. *Nature Biotechnology*, 8(4), p.308.

Schneider, G., Käck, H. and Lindqvist, Y., 2000. The manifold of vitamin B 6 dependent enzymes. *Structure*, 8(1), pp.R1-R6.

Shapiro, S.K., 1958. Adenosylmethionine-homocysteine transmethylase. *Biochimica et biophysica acta*, 29(2), pp.405-409.

Sharif, S., Fogle, E., Toney, M.D., Denisov, G.S., Shenderovich, I.G., Buntkowsky, G., Tolstoy, P.M., Huot, M.C. and Limbach, H.H., 2007. NMR localization of protons in critical enzyme hydrogen bonds. *Journal of the American Chemical Society*, 129(31), pp.9558-9559.

Shaw, J.P., Petsko, G.A. and Ringe, D., 1997. Determination of the structure of alanine racemase from *Bacillus stearothermophilus* at 1.9-Å resolution. *Biochemistry*, 36(6), pp.1329-1342.

Singh, S., Madzelan, P. and Banerjee, R., 2007. Properties of an unusual heme cofactor in PLP-dependent cystathionine β -synthase. *Natural product reports*, 24(3), pp.631-639.

Song, J., Tan, H., Perry, A.J., Akutsu, T., Webb, G.I., Whisstock, J.C. and Pike, R.N., 2012. PROSPER: an integrated feature-based tool for predicting protease substrate cleavage sites. *PloS one*, 7(11), p.e50300.

Stemmer, W.P., 1994. Rapid evolution of a protein in vitro by DNA shuffling. *Nature*, 370(6488), p.389.

Tai, C.H. and Cook, P.F., 2001. Pyridoxal 5'-phosphate-dependent α , β -elimination reactions: mechanism of o-acetylserine sulphydrylase. *Accounts of chemical research*, 34(1), pp.49-59.

Taylor, R.T., 1982. B12-dependent methionine biosynthesis. *B*, 12, pp.307-355.

Taoka, S., Ohja, S., Shan, X., Kruger, W.D. and Banerjee, R., 1998. Evidence for heme-mediated redox regulation of human cystathionine β -synthase activity. *Journal of Biological Chemistry*, 273(39), pp.25179-25184.

Taoka, S., Lepore, B.W., Kabil, Ö., Ojha, S., Ringe, D. and Banerjee, R., 2002. Human cystathionine β -synthase is a heme sensor protein. Evidence that the redox sensor is heme and not the vicinal cysteines in the CXXC motif seen in the crystal structure of the truncated enzyme. *Biochemistry*, 41(33), pp.10454-10461.

Tobin, M.B., Gustafsson, C. and Huisman, G.W., 2000. Directed evolution: the 'rational' basis for 'irrational' design. *Current opinion in structural biology*, 10(4), pp.421-427.

Toney, M.D., 2001. Computational studies on nonenzymatic and enzymatic pyridoxal phosphate catalyzed decarboxylations of 2-aminoisobutyrate. *Biochemistry*, 40(5), pp.1378-1384.

Toney, M.D., 2011. Controlling reaction specificity in pyridoxal phosphate enzymes. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, 1814(11), pp.1407-1418.

Wang, S., Ma, J., Peng, J. and Xu, J., 2013. Protein structure alignment beyond spatial proximity. *Scientific reports*, 3, p.1448.

Whitfield, C.D., Steers, E.J. and Weissbach, H., 1970. Purification and properties of 5-methyltetrahydropteroyltriglutamate-homocysteine transmethylase. *Journal of Biological Chemistry*, 245(2), pp.390-401.

Woodyer, R., Chen, W. and Zhao, H., 2004. Outrunning nature: directed evolution of superior biocatalysts. *J. Chem. Educ.*, 81(1), p.126.

Yamagata, S., 1984. O-Acetylhomoserine sulfhydrylase of the fission yeast *Schizosaccharomyces pombe*: partial purification, characterization, and its probable role in homocysteine biosynthesis. *The Journal of Biochemistry*, 96(5), pp.1511-1523.

Yamagata, S., D'andrea, R.J., Fujisaki, S., Isaji, M. and Nakamura, K., 1993. Cloning and bacterial expression of the CYS3 gene encoding cystathionine gamma-lyase of *Saccharomyces cerevisiae* and the physicochemical and enzymatic properties of the protein. *Journal of bacteriology*, 175(15), pp.4800-4808.

Yano, T., Kuramitsu, S., Tanase, S., Morino, Y. and Kagamiyama, H., 1992. Role of Asp222 in the catalytic mechanism of *Escherichia coli* aspartate aminotransferase: the amino acid residue which enhances the function of the enzyme-bound coenzyme pyridoxal 5'-phosphate. *Biochemistry*, 31(25), pp.5878-5887.