**The Dissertation Committee for Matthew Wesley Bonem Certifies that this is the approved version of the following Dissertation:**

# Development of Human Enzyme Therapeutics and the Tools for their Engineering

**Committee:**

George Georgiou, Supervisor

Everett Stone, Co-Supervisor

Hal Alper

Jennifer Maynard

Stefano Tiziani

# Development of Human Enzyme Therapeutics and the Tools for their Engineering

**by**

**Matthew Wesley Bonem**

**Dissertation**

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

**Doctor of Philosophy**

**The University of Texas at Austin**

**December 2018**

# Dedication

To my wonderful parents – Mike and Bonnie.

# Acknowledgements

First I would like to acknowledge and thank my advisor, Dr. George Georgiou, for the great opportunity to work in, and be a part of his research group. The last 5 years have allowed me to grow as a scientist, in large part due to his mentorship and advice. Additionally, I would like acknowledge and thank my co-advisor, Dr. Everett Stone. Dr. Stone has been an invaluable soundboard while troubleshooting the plethora of problems that have arisen during these projects. It goes without saying, both Dr. Georgiou and Dr. Stone have been instrumental to this work, and it would not exist without them. I am extremely grateful for the support and experience they have provided through this work.

I would also like to thank all of my committee members, Dr. Hal Alper, Dr. Jennifer Maynard, and Dr. Stefano Tiziani, for their insight and feedback during my prelim.

I would also like to thank Dr. Brent Iverson for his support and insight during our subgroup meetings.

I would also like to thank all of the BIGG lab members. In particular, I would like to thank Ebru Cayir and Dr. Makiko Watanabe for their valuable contributions to chapters 2 and 3 respectively. Additionally, I'd like to acknowledge the undergraduate and high school research assistants that I had the pleasure of working with: Emma Rowlinson, Jenny Kim, Grace Keliher, Chelsea Cuellar, and Yawen Ren. Additionally, I'd like to thank my lab mentors, Dr. Olga Paley, Dr. Shira Cramer, and, particularly, Dr. Giulia Agnello. Lastly, I would also like to particularly thank Dr. Jonathan McDaniels, Norah Ashoura, Joseph DeSautelle, and Candice Lamb for their support, feedback, and fun times.

I am grateful for the opportunity to interact with and work with some of the people at Aeglea as well as their contributions to chapter 3. In particular, I would like to thank Dr. Scott Rowlinson, Joseph Tyler, and Dr. Giulia Agnello.

I would like to also acknowledge and thank my undergraduate research advisor, Dr. Laura Segatori, for her mentorship and encouragement to pursue a Ph.D.

Lastly, I would like to thank all of my friends and family for their support throughout graduate school. In particular, my parents, Mike and Bonnie, for their support throughout graduate school and before. My siblings, David, Jonathan, and Hope, as well as their spouses, Karlie and Kellie, for indulging countless conversations about science and graduate school. Masha Konopleva and Norah Ashoura for their support and friendship. My grandparents, Joseph and Diane, for their encouragement and serving as wonderful role models.

# Abstract

## Development of Human Enzyme Therapeutics and the Tools for their Engineering

Matthew Wesley Bonem, Ph.D.

The University of Texas at Austin, 2018

Supervisor: George Georgiou and Everett Stone

Enzyme therapeutics are a growing class of pharmaceuticals for the treatment of a plethora of diseases, including the treatment of cancer and metabolic disorders. In cancer, amino acid depletion therapy has seen success with the family of l-asparaginase enzymes utilized for the treatment of acute lymphoblastic leukemia (ALL). For metabolic disorders, a new treatment modality termed enzyme replacement therapy (ERT), has seen much success, in particular with lysosomal storage disorders. With the evolution of this field has come the need for improved tools and techniques for the engineering and characterization of potential drug candidates. This work highlights 1) the development of a tool for therapeutic enzyme development, 2) the validation of a therapeutic enzyme for treating homocystinuria, 3) the development of a tool for peptide synthesis and purification, 4) the identification of a scaffold enzyme to replace existing amino acid depletion treatments.

Homocystinuria is a metabolic disorder that disrupts the transsulfuration pathway, causing a series of symptoms that typically result in cardiac arrest. This work showcases a novel enzyme therapeutic based on human cystathionine-γ-lyase (hCGL) for the treatment of homocystinuria and demonstrates its efficacy in multiple mouse models, completely

preventing the neonatal lethality associated with homozygous deletion of the *cbs* gene. To further improve the homocystinuria enzyme therapeutic, we explored a number of genetic selection strategies for the directed evolution of enzyme variants with superior catalytic and biophysical properties. Specifically, we designed a genetic selection based on an auxotrophy for α-ketobutyrate and, correspondingly, isoleucine. Also as part of this dissertation, beyond the homocystinuria enzyme therapeutic, we designed and developed an novel, orthogonal technology for the expression and purification of peptides based on human asparaginase-like protein 1 (hASRGL1). Lastly, we explored various human scaffold proteins in an effort to develop a successor to the bacterial asparaginase currently used for the treatment of acute lymphoblastic leukemia.

# Table of Contents

# List of Tables

# List of Figures

# Chapter 1: Introduction

## 1.1 Protein Therapeutics

Recombinant insulin, arguably the most widely used protein therapeutic, was approved by the FDA in 1982, becoming the first clinically used protein produced by genetic engineering approaches[1,2]. Since then, the field of recombinant protein therapeutics has blossomed to target an array of diseases. Over the last few decades, therapeutic proteins have become a staple of modern medicine, with over 200 FDA-approved drugs[3]. There a number of key parameters that need to be satisfied by all therapeutic proteins. In particular, protein therapeutics must: 1) retain high activity and persistence in a biological environment, 2) have sufficient activity to be therapeutic relevant without excessive dosing requirements, 3) can be produced in sufficient yield in order to be economically viable, 4) limit non-specific interactions to avoid unintended side-effects, and finally 5) they should not elicit an immune response (because they are recognized as foreign to the human immune system). As the field of protein therapeutics has advanced, the drugs, their mechanisms of action, and the diseases they target have increased in complexity. This has been enabled, in large part, by advances in technology that circumvents the shortcoming of potential drug candidates. One such advancement, PEGylation, consists of covalently attaching polyethylene glycol polymer chains to a protein in order to increase its persistence in circulation and to decrease immunogenicity[4–6]. Another major advancement is the field of protein engineering, which, through the introduction of mutations to a protein scaffold, seeks to generate protein variants having better biological activity, stability in physiological fluids, lower immunogenicity and improved expression[7]. As protein engineering techniques have advanced, the number of FDA approved drugs has increased, addressing the unmet medical needs of various patient populations.

## 1.2 APPLICATIONS

As mentioned above, the field of protein therapeutics has expanded over the decades to address multiple different aspects of human health and disease. Below a few such applications relevant to this work, as well as their implications for protein therapeutics in general, are discussed.

### 1.2.1 Cancer & Amino Acid Depletion

With over 1.6 million new cancer cases and over 600,000 deaths in the U.S., cancer is the second leading cause of death in the U.S. and globally[8]. In spite of this, cancer treatment, until recently, has largely remained stagnant in its archaic modality. Chemotherapy and radiation treatment, although effective at killing tumor cells, also have the unintended side effects of killing normal cells, often causing inadvertent toxicities as severe as the cancer[9]. Although both of these therapeutic approaches have advanced greatly, the ability to target tumors specifically by chemotherapeutic agents or by radiotherapy has remained elusive, and thus the toxicity associated with these treatments has remained an issue. Almost a century ago, Warburg identified important metabolic differences in the central metabolism of cancer and normal cells[10–12]. Over the years numerous other metabolic abnormalities have since been identified in cancer cells, a number of which relate to amino acid biosynthesis and salvage pathway[13,14]. Originally, dietary restriction for specific amino acids was attempted as a way to exploit metabolic vulnerabilities of cancer cells, but, unfortunately, often failed to have a therapeutic effect. Alternatively, a number of protein therapeutics have been generated and are being evaluated for the treatment of cancer based on the basis of amino acid depletion. The most well-known of these, and only FDA approved treatment based on amino acid depletion, is

the family of L-asparaginases. The use of L-asparaginase as an anti-cancer agent dates back to the mid-1960s, where it was used for the treatment of acute lymphoblastic leukemia (ALL) to exploit its high demand for the non-essential amino acid, L-asparagine[15,16]. The use of L-asparaginase was limited to children and young adults due to the toxicity observed in adult patients[17]. Subsequently, a PEGylated variant of L-asparaginase marketed by the trade name Oncaspar, was developed, allowing for reduced dosing due to increase circulation persistence, but more importantly reduced its immunogenicity[18]. Nonetheless even with PEGylation Oncaspar induces a hypersensitivity reaction in approximately 30% of patients[19]. This limitation notwithstanding, the efficacy of L-asparaginase treatment cannot be understated, with 5-year survival rates increasing from 54% in the 1970s to 90% now that Oncaspar is included in the chemotherapeutic regiment used to treat childhood ALL[15]. Since the success of L-asparaginase, many other amino acid depletion treatments have been explored, including L-methionase, L-cyst(e)inase, L-arginase[20–22]. Inevitably, as more metabolic vulnerabilities are identified, additional amino acid depleting therapeutics, or even intermediate metabolite depleting therapeutics, will be developed.

Recent work has demonstrated that the efficacy of L-asparaginase, for a number of cancer cell lines[23], is due to its glutaminase activity, not its asparaginase activity. A number of hypothesis regarding the importance of glutamine have been proposed[24]: 1) glutamine, when glucose is scarce, can serve as a precursor for α-ketoglutarate, propagating the TCA cycle, 2) glutamine, as a precursor for glutathione, is essential for controlling reactive oxygen species (ROS), 3) glutamine serves as a ligand for a number of receptors, a number of which are expressed in certain cancer cell lines. Regardless of the exact mechanism, glutamine's importance to cancer proliferation highlights another metabolic vulnerability capable of being targeted via amino acid depletion therapy.

### 1.2.2 Inborn Errors of Metabolism & Enzyme Replacement Therapy

Another major field of protein therapeutic is enzyme replacement therapy (ERT). ERT typically targets inborn errors of metabolism (IEM) in which the patient either does not produce a specific enzyme or produces a defective version of the enzyme[25]. Ultimately the disorders disrupts metabolic regulation and flow, resulting in a disease state dependent on the specific mutation. ERT typically consists of administration of an active form of the missing/defective enzyme to patients, allowing for restoration of the metabolic pathway, and subsequently alleviation of the associated symptoms. Unfortunately, due to the compartmentalization of human biology on an organ, cellular, and organelle level, ERT is limited in which IEMs it can be applied to. Typically, the disease may affect the patient on a cellular level, which intravenously administered therapeutic enzymes often cannot affect directly. The most well-known example of ERT is in the treatment of lysomsal storage disorders, such as Cerezyme in the treatment of Gaucher disease[26,27]. Conveniently, for these therapeutics, intravenous administration is sufficient, as the therapeutic drug can be uptaken by cells. Regardless of the difficulties that arise from compartmentalization, ERT has shown much promise in the treatment of IEMs.

An IEM that may benefit from ERT is Homocystinuria. Homocystinuria is caused by a defect of the transsulfuration pathway (shown in Figure 1.1) that affects approximately 20,000 to 40,000 people worldwide, homocystinuria is often caused by mutations in the cystathionine-β-synthase (CBS) gene[28]. On a cellular basis, classical homocystinuria affects a critical point in the transsulfuration pathway, determining the fate of homocysteine and affecting the flux and subsequent regulation of this pathway. Homocysteine can be remethylated into methionine by the betaine-homocysteine s-methyltransferase (BHMT) enzyme[29]. Alternatively, homocysteine can be irreversibly committed to cysteine biosynthesis via the CBS enzyme[30,31]. Disruption of this

unidirectional flow to cysteine results in elevated levels of homocysteine and methionine in patients, as well as decreased levels of cysteine and cystathionine[32,33].

Ultimately, patients suffer from a variety of complications, including thromboembolic events, dislocation of ocular lenses, and spinal osteoporosis[28,34–37]. Patients first develop myopia, typically by the age of 1, which, if the disease was not detected via newborn screening, serves as the first indication of the disease. Another common indication of classical homocystinuria is developmental delay, which is usually more severe in B6-non-responsive patients[28]. As the patient ages, the disease progresses with ocular issues escalating to lens dislocation. Further mental decline also correlates with severity of disease; B6-reponsive patients have a mean IQ of 79, while for B6-nonresponsive patients, the mean IQ is 57[28]. B6-nonresponsive individuals identified by newborn screening with good treatment compliance see an improvement in IQ, with a mean IQ of 105 for this group[28]. By or around puberty, patients start to display severe connective tissue problems as evident by their marfanoid appearance and osteoporosis. Cardiovascular distress is typically the last, but most severe, symptom to arise, resulting in 23% mortality by the age of 30 for B6-non-responsive patients[28].

Current homocystinuria treatments primarily focus on decreasing the concentration of total homocysteine (tHcy), with a secondary focus on increasing total cysteine (tCys) levels. Initially, patients are treated with pyridoxine (vitamin B6) supplementation, to which less than half are responsive[33,38]. Unresponsive or partially responsive patients begin a low methionine diet, along with cysteine supplementation, in an effort to normalize the metabolic imbalance[39]. Betaine treatment can further help normalize the metabolic imbalance, by increasing the flux through BHMT, but recent research has shown limited efficacy[40]. As often observed, patient compliance with restrictive/medical diets can be poor, resulting in uncontrolled disease; there is a clear need for an effective non-dietary

treatment. Ideally, a next generation treatment for homocystinuria will limit patient burden, improve quality of life and remove dietary compliance issues. Recent work has showcased a novel enzyme replacement therapy (ERT) for the treatment of classical homocystinuria[41–43], although there is room for additional, improved treatment options.

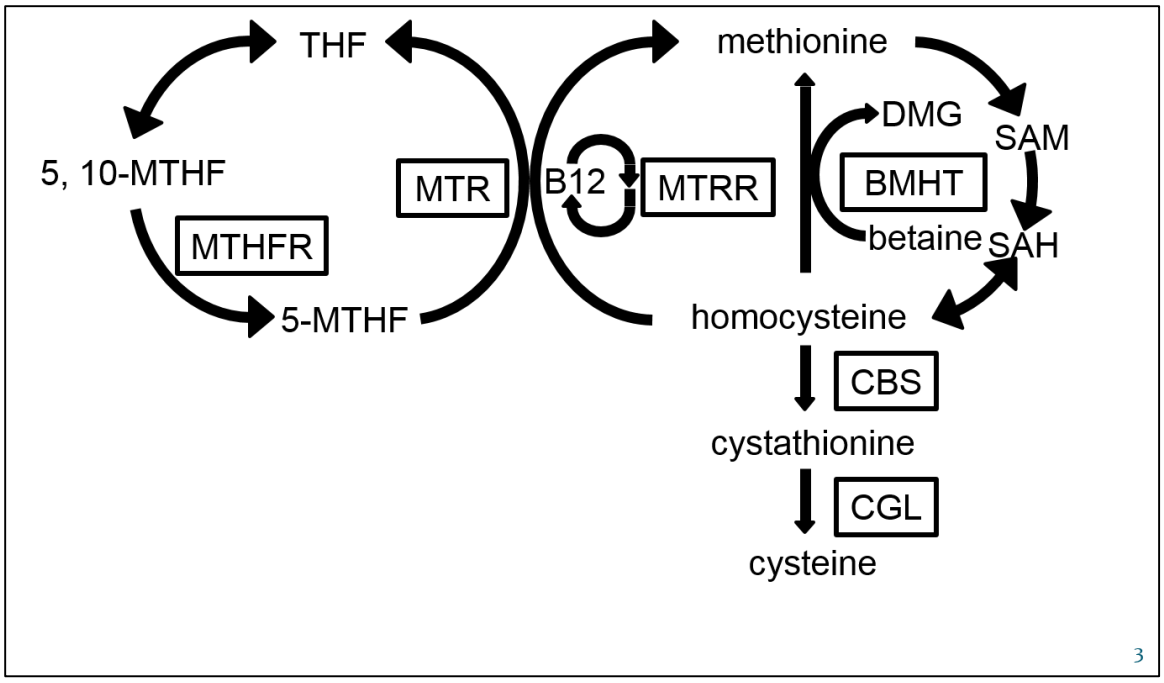


Figure 1.1 The Human Transsulfuration Pathway

The human transsulfuration pathway showing the metabolic pathway from methionine to cysteine. In homocystinuria, the cystathionine-β-synthase (CBS) enzyme is defective, resulting in elevated levels of homocysteine and methionine, concomitant with decreased levels of cysteine and cystathionine. **THF**: Tetrahydrofolate; **5,10-MTHF**: 5,10-Methylenetetrahydrofolate; **5-MTHF**: 5-Methyltetrahydrofolate; **B12**: Vitamin B12; **DMG**: Dimethylglycine; **SAM**: S-Adenosyl Methionine; **SAH**: S-Adenosyl Homocysteine; **MTHFR**: Methylenetetrahydrofolate reductase; **MTR**: 5-Methyltetrahydrofolate-Homocysteine Methyltransferase; **MTRR**: 5-Methyltetrahydrofolate-Homocysteine Methyltransferase Reductase; **BHMT**: Betaine-Homocysteine S-Methyltransferase; **CBS**: Cystathionine-β-synthase; **CGL**: Cystathionine-γ-lyase;

## 1.3 HUMAN ENZYME SCAFFOLDS

As discussed above, protein therapeutics have the potential to address many different unanswered problems in medicine. In an effort to avoid immunogenicity as seen with Oncaspar, many recent therapeutics have focused on using human enzymes rather than enzymes with bacterial origins. Below are a few enzymes that are relevant to this work.

### 1.3.1 Cystathionine Beta Synthase

Cystathionine-β-synthase (hCBS) catalyzes the reaction of homocysteine with serine, forming cystathionine and water[44]. hCBS contains multiple domains. The N-terminal domain, which is only present in higher order mammalian genes, is a heme binding domain[44]. The core catalytic domain of hCBS requires a co-factor, pyridoxal phosphate (PLP), in order to catalyze its native reaction. Lastly, its C-terminal domain has an S-adenosyl methionine sensor, serving as an allosteric regulator of its activity[45]. In order to be activated, hCBS requires the binding of S-adenosyl methionine to its C-terminal inhibitory domain. Due to its crucial role in the transsulfuration pathway and homocystinuria, hCBS would serve as a natural choice for ERT for homocystinuria. Unfortunately, hCBS, due to its multiple domains, has a number of issues that limit its potential as a therapeutic drug. One such limitation is its allosteric regulation by S-adenosyl methionine[44,46]. Recent work has developed truncated variants, removing both the C-terminal and N-terminal domains, resulting in activity without S-adenosyl methionine activation[47,48]. Additional recent work has re-engineered hCBS to improve its biophysical properties, improving its stability and activity[41].

### 1.3.2 Cystathionine Gamma-Lyase

An enzyme of the transsulfuration pathway, human cystathionine-γ-lyase (hCGL) catalyzes the α,γ-elimination of cystathionine to produce ammonia, water, cysteine, and α-ketobutyrate, as shown in Figure 1.2. Although capable of performing other α,β-eliminations and α,γ-eliminations, hCGL has significantly reduced activity and selectivity towards these substrates. The introduction of point mutations in and near the activity site has been shown to alter both the catalytic activity and selectivity of hCGL to instead favor degradation of cysteine and cystine[20]. An engineered variant having cyst(e)inase activity has shown promise both as a cancer treatment as well as a possible treatment for specific metabolic disorders. More importantly, this work highlights the amenability of hCGL to engineering in order to alter its activity and selectivity while retaining its stability and high expression levels. In light of this, hCGL could serve as a scaffold protein for the treatment of other thiol-related disorders.

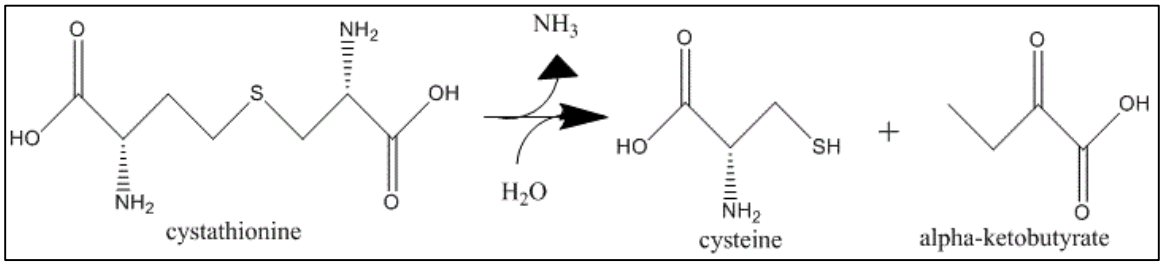


Figure 1.2 Cystathionine-γ-Lyase Reaction

The various reactions catalyzed by CGL

### 1.3.3 Asparaginase-Like Protein 1

Human asparaginase-like protein 1 (hASRGL1) is an N-terminal nucleophile hydrolase with β-aspartyl peptidase and asparaginase activity. Similar to the

aforementioned therapeutic proteins used for the treatment of ALL, hASRGL1 catalyzes the degradation of asparagine into aspartate. hASRGL1 initially forms a catalytically inactive dimer. The dimer then catalyzes an intramolecular cleavage of the peptide bond between Gly167 and Thr168, splitting one of the monomers into an α-subunit and β-subunit[49]. Only after this cleavage event for both dimers does hASRGL1 become catalytically active towards asparagine. As the intramolecular processing does not proceed to 100% cleavage, a heterogeneous mixture of active and inactive enzyme is inevitably formed, limiting hASRGL1's efficacy as a therapeutic protein. Recent work from our lab led to a circularly permutated version of hASRGL1, cp-hASRGL1[50]. In this variant, the C-terminal of the β-subunit is fused to the N-terminal of the α-subunit via a flexible linker. As such, the fusion protein, no longer contains the peptide bond between Gly167 and Thr168, circumventing the need for intramolecular processing. More importantly, this circularly permutated version retains its asparaginase activity.

### 1.3.4 Glutaminase Isoforms

Although the efficacy of Elspar and Oncaspar was originally attributed to their asparaginase activity, recent work has highlighted their glutaminase activity as a key factor in their anti-cancer efficacy for a number of cell lines[23,51]. Additionally, a number of studies have showcased the importance of glutamine to central metabolism in a number of cancer cell lines. In separate studies glutaminase inhibitors are being explored for the treatment of certain cancers[52–54]. In an effort to improve upon the success of Oncaspar, in particular reducing the immunogenicity, a human enzyme with glutaminase activity is highly desirable.

The human genome encodes two different glutaminases, each with multiple isoforms caused by mRNA splicing[55,56]. The first, kidney-type glutaminase (GLS1), has reasonable activity towards glutamine ($k_{cat}$ = 10 ± 1 $s^{-1}$; $K_M$ = 1.9 ± 0.4 mM; $k_{cat}/K_M$ = (5.3 ± 1.2) x $10^3$ $M^{-1}s^{-1}$) but requires inorganic phosphate at levels beyond physiological levels ($K_A$=76 ± 7 mM) to be active[55]. An elongated splice variant of GLS1, GAC, also has a strong dependence on inorganic phosphate, and more favorable kinetics ($k_{cat}$ = 22 ± 2 $s^{-1}$; $K_M$ = 1.4 ± 0.4 mM; $k_{cat}/K_M$ = (1.6 ± 0.5) x $10^4$ $M^{-1}s^{-1}$)[55]. The second type of glutaminase, liver-type glutaminase (GLS2), has a very slight dependence on inorganic phosphate ($K_A$ = 8 ± 1 mM) but has very poor activity towards glutamine ($k_{cat}$ = 2.5 ± 0.3 $s^{-1}$; $K_M$ = 4.0 ± 0.2 mM; $k_{cat}/K_M$ = (6.3 ± 0.8) x $10^2$ $M^{-1}s^{-1}$)[55].

# Chapter 2: Breaking the metabolic regulation of the branched chain amino acid biosynthesis pathway in order to engineer a protein therapeutic for homocystinuria

## 2.1 Abstract

Improving the biophysical properties of an enzyme is a time consuming task, which can be greatly aided by the usage of high throughput screens and selection. Genetic selections are arguably one of the strongest tools for this purpose. By creating an auxotrophy in a cell that in turn the enzyme of interest can rescue, the host organism's growth rate becomes coupled to enzyme activity, allowing for the identification of superior mutants. When successfully implemented, a single genetic selection experiment can survey the catalytic activity of $10^8$ mutant proteins over the course of a few days. This chapter discusses the development and optimization of a genetic selection designed to re-engineer cystathionine-γ-lyase (hCGL) into a homocysteine and homocystine degrading enzyme. Previous work identified the branched chain amino acid biosynthesis (BCAA) pathway as a target for such a genetic selection[57]. The previous work failed to address the plethora of metabolic feedback loops in this pathway as well as the toxicity associated with various metabolites in the pathway. As such, when applied to cases other than proof of concept experiments, the genetic selection did not function as desired. To this end, this work, beyond the identification of a superior hCGL mutant, aimed to remedy the genetic selection pressure. Although this work did not ultimately result in identifying superior hCGL mutants, it did create a number of hCGL libraries that can serve as the basis for future engineering efforts. Additionally, this work, through the introduction of additional, metabolic enzymes, as well as media additives to alter the metabolic regulation, successfully corrected the coupling between enzyme activity and growth rate. Lastly,

through the course of these experiments, the metabolic regulation of the branched chain amino acid biosynthesis pathway was further elucidated.

## 2.2 INTRODUCTION

Homocystinuria is an inborn error of metabolism disease (IEM) that affects the transsulfuration pathway, resulting in accumulation of homocysteine and methionine with a concomitant decrease in cysteine and cystathionine[28]. Treatment focuses on decreasing or preventing the accumulation of homocysteine, which can alleviate disease pathologies[39]. Enzyme replacement therapy (ERT) is a therapeutic approach for treating IEMs by administering a recombinant form of the missing or defective enzyme that causes the disease[27]. In the case of homocystinuria, ERT could involve the administration of a recombinant form of cystathionine-β-synthase (hCBS). Unfortunately, hCBS has a number of issues that limit its potential as a therapeutic drug. One such limitation is its allosteric regulation by S-adenosyl methionine[44,46]. In order to be activated, hCBS requires the binding of S-adenosyl methionine to its C-terminal inhibitory domain. Truncations of both the C-terminal and N-terminal domains that result in activity without S-adenosyl methionine activation[47,48]. Recent work has re-engineered hCBS to improve its biophysical properties, alleviating its allosteric regulation and improving its stability[41]. Alternatively, as demonstrated by recent work, hCGL is amenable to engineering to alter its specificity while retaining many properties favorable for a therapeutic drug[20]. Natively, hCGL catalyzes the α,γ-elimination of cystathionine into cysteine, α-ketobutyrate, and ammonia, but it also has some activity towards both homocysteine and homocysteine, as shown in Figure 2.1. Although hCGL would not be able to replenish cysteine and cystathionine levels, in the manner that hCBS does, depletion of homocysteine is the primary goal of current treatment options, and thus, an engineered hCGL represents an interesting and exciting opportunity for the treatment of homocystinuria.

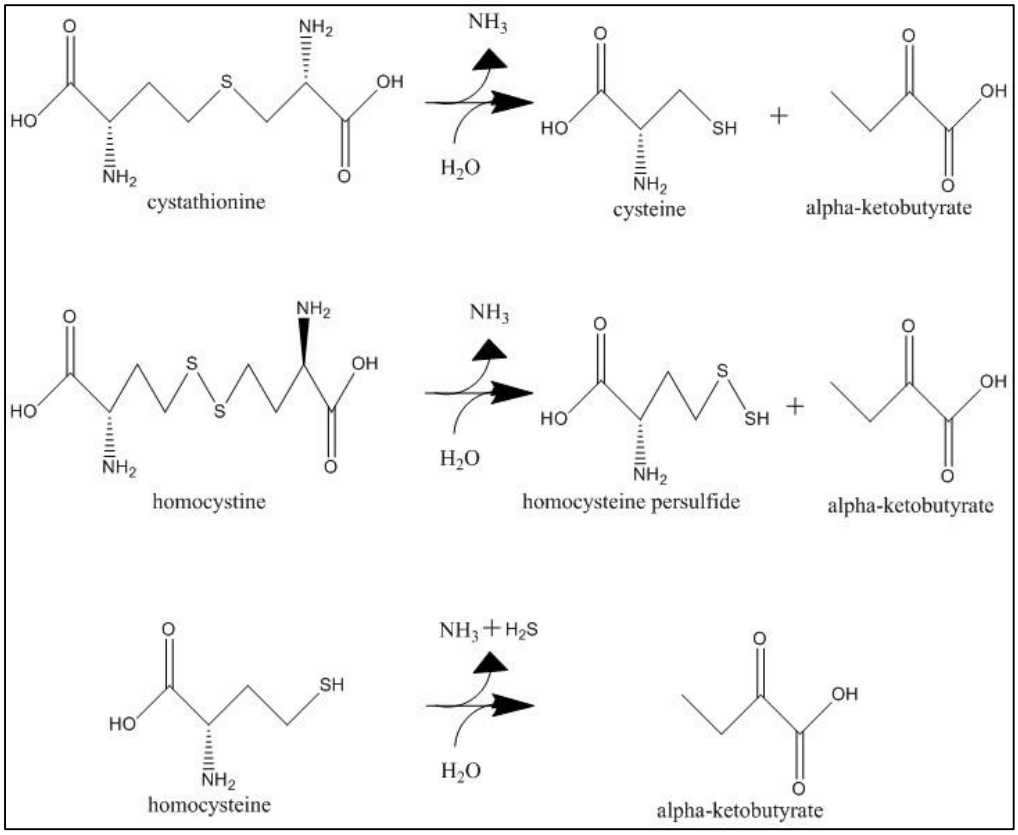


Figure 2.1 hCGL Reactions

The canonical reaction of hCGL, converting cystathionine into cysteine and α-ketobutyrate is shown at the top. The first, of many, alternative reactions converts homocystine into homocysteine and α-ketobutyrate. The second alternative reaction converts homocysteine into α-ketobutyrate.

In order to alter the activity and specificity of hCGL as required for therapeutic purposes, the protein inevitably requires engineering of its active site and secondary shell to favor homocysteine and homocystine. Although rational mutagenesis can expedite the process, such an undertaking inevitably requires sampling a large number of mutants for the desired profile. To that end, high through-put screening and selection methods are often employed to facilitate such a task. Genetic selections couple the desired enzymatic activity to a growth phenotype in a host organism, often *E. coli* or *S. cerevisiae*. This coupling is often accomplished by using host organisms engineered to be deficient in certain enzyme

genes, i.e. introducing an auxotrophy which the enzyme of interest can complement. In the simplest form, genetic selections can serve as a "live-dead selection", requiring a threshold activity for growth and thus removing any variants that are misfolded or have low activity. Ideally, the engineered organism's fitness is intricately coupled to the enzyme's activity. In such cases, a cell expressing a high activity variant has a fitness advantage over cells expressing worse variants. Thus, over multiple generations, cells expressing the superior variants can enrich due to their growth advantage, allowing for quick, high-throughput identification of improved variants.

As with any technique there are a number of hurdles involved in creating a successful genetic selections. The first is that the reaction of interest must produce a metabolite that either is essential, directly or indirectly, to growth. If the metabolite of interest is not natively utilized by the metabolism of the host organism, additional genes can be knocked out and knocked in to make the host organism dependent on said metabolite. Additionally, designing an auxotrophy for a ubiquitous metabolite, can be complicated because the metabolite is consumed or generated by a variety of metabolic pathways. Selections based on such metabolites require extensive genome editing which often leads to a non-viable organism. Beyond that, disruption of even a single metabolic pathway often has unintended, indirect effects due to the cell's metabolic regulation. The resulting phenotype can potentially decouple the enzyme activity from the organism's growth rate. Outside of the host organism and its metabolic pathway, it is possible that the enzyme of interest has moonlighting activity that produces unwanted metabolites or consumes needed metabolites, further convoluting the selection. Even if all of the aforementioned hurdles are addressed, there is no guarantee that growth rate and enzyme activity have a positive correlation over the range of the enzyme's activity. Due to these

hurdles, creating a successful genetic selection is a non-trivial task and, in spite of their power, successful genetic selections may not always be the best choice for engineering.

Previous studies led to the development of a strain of *E. coli*, *E. Coli* BL21 (DE3) *ΔmetAΔilvA*, here on referred to as Blimko, which is auxotrophic for α-ketobutyrate and, consequently, for isoleucine[57] as shown in Figure 2.2. In E. coli metabolism, α-ketobutyrate is only involved in one metabolic pathway, the branched chain amino acid (BCAA) biosynthesis pathway, and is a precursor to isoleucine. α-ketobutyrate is produced via the degradation of threonine by threonine deaminase (*ilvA*). Deletion of *ilvA* eliminates α-ketobutyrate biosynthesis. Additionally, to avoid biosynthesis of cystathionine or other thiol molecules that could serve as unwanted substrates for hCGL, *metA* was knocked out. Blimko cells, when grown in M9 minimal media, do not grow, but can be rescued by the addition of isoleucine and methionine. Additionally, by expressing a recombinant protein capable of producing α-ketobutyrate, such as methionine-γ-lyase (pMGL) from *P. putida*, Blimko cells can grow in the M9 minimal media supplemented only with methionine. In light of that, proof of concept experiments were conducted to assay how enzyme activity correlates with cell fitness. In brief, Blimko cells expressing pMGL and Blimko cells expressing hCGL were co-inoculated into the selective media. Over successive generations the Blimko cells expressing pMGL outgrew the Blimko cells expressing hCGL as evidenced by the increase in the number of colonies harboring the pMGL plasmid. This enrichment was attributed to pMGL's superior catalytic activity, and therefore its increased α-ketobutyrate production. In light of this, it was believed that the genetic selection could, identify variants with increased activity, thus serving as a useful tool in the development of a therapeutic drug for homocystinuria.

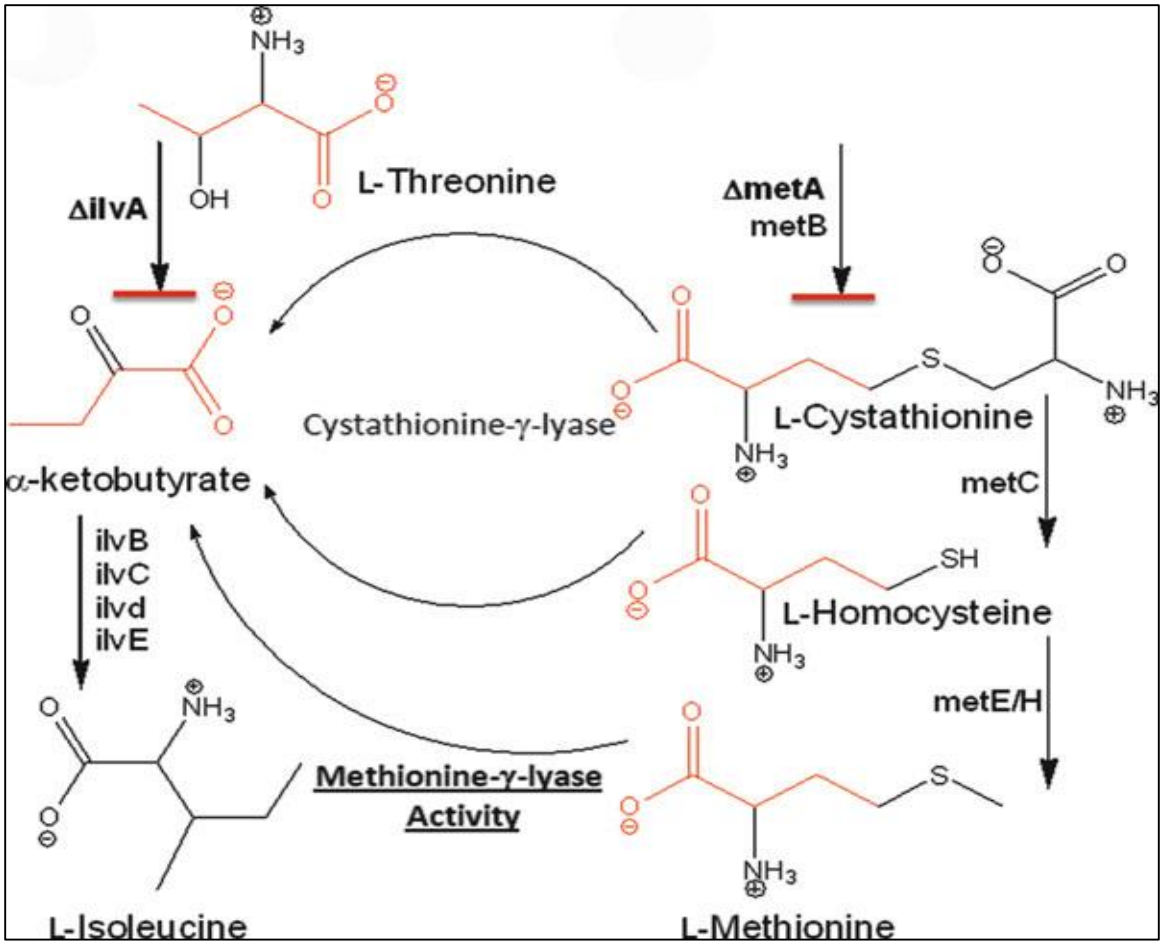


Figure 2.2 Genetic Selection Pathway in Blimko Strain

The branched chain amino acid (BCAA) biosynthesis pathway in the Blimko strain. The *ilvA* knock out prevents α-ketobutyrate production, causing an isoleucine auxotrophy. The *metA* deletion disrupts the transsulfuration pathway, preventing generation of the various metabolites which can serve as a source for α-ketobutyrate. Figure from: Paley et al., 2013, *Methods in Molecular Biology*

The work discussed here highlights the various attempts to develop a genetic selection for the engineering of a homocyst(e)inase enzyme. Multiple libraries were created, utilizing both rational and random design, to explore the vast mutational landscape for an enzyme with the desired activity profile. Although the proof of concept experiments for the genetic selection showed promise, they failed to address many underlying issues

with the selection. This work describes a number of strategies that were attempted in order to implement the proposed genetic selection approach. Through these efforts, many of the metabolic regulations of the BCAA biosynthesis pathway were disrupted and, as a consequence, new details regarding the pathway and its regulation were elucidated. Ultimately this work has built the framework for an α-ketobutyrate based genetic selection and, subsequently, an improved homocystinuria therapeutic drug.

## 2.3 MATERIALS AND METHODS

*Materials:* Oligonucleotides were purchased from Integrated DNA Technologies. Restriction enzymes, Taq polymerase, Phusion high fidelity DNA polymerase, T4 DNA ligase, corresponding buffers, and dNTPs were purchased from New England Biolabs. Difco 2xYT growth media, LB growth media, and TB growth media were all purchased from Becton Dickinson. Nickel-nitrotriacetic acid (Ni-NTA) resin and mini-prep kits were purchased from Qiagen. DL-Homocysteine, homocysteine, L-methionine, DL-cysteine, cystine, 3-methyl-2-benzothiazoninone hydrazine (MBTH), isopropyl β-D-1-thiogalactopyranoside (IPTG), kanamycin, ampicillin, chloramphenicol, and spectinomycin were all purchased from Sigma-Aldrich.

*Mutant Cloning:* Specific hCGL mutants were synthesized by overlap extension PCR. Briefly, an initial round of PCRs was conducted, utilizing gene-specific primers to encode the desired mutations coupled with plasmid-specific primers. Subsequently, a second PCR was initiated, using the products from the mutagenic first PCRs and the plasmid-specific primers. After the second PCR, the product was purified via gel extraction. The gel extracted DNA was digested using NcoI and EcoRI, after which it was ligated into the pET28a plasmid or pZE11 plasmid using the T7 ligase protocol. Ligated plasmid was then transformed into DH10β cells and plated on an antibiotic selective agar plate. Individual colonies were picked and their identify confirmed via Sanger sequencing. Confirmed plasmids were re-transformed into either BL21(DE3) cells or Blimko cells and again plated on antibiotic selective agar plates. Colonies were picked and used to inoculate overnight cultures, which were then frozen for subsequent usage.

*Library Construction:* Initial libraries were constructed using a pET28a plasmid containing the hCGL gene. Subsequent libraries were constructed using a modified version of the pZE11 plasmid containing the hCGL gene. Libraries were constructed in both

random and rational fashions. Libraries based on phylogenetic analysis were constructed via overlap extension PCR utilizing degenerate primers to encode the desired amino acids. Random libraries were also constructed via overlap extension PCR but utilized NNS codons to mutate specific residues. Additionally, a previous library of hCGL was constructed utilizing the pFunkel method[58]. Through the use of overlap extension PCRs this library was modified, serving to change the parental template.

*Genetic Selection:* In its final iteration, the homocyst(e)inase genetic selection consisted of a selective media and a nonselective media. The nonselective media consisted of LB media with 100 μg/mL of ampicillin and 50 μg/mL of spectinomycin. The recipe for the selective media is shown in Table 2.1. Previous iterations of the homocyst(e)inase genetic selection selective media utilized a similar recipe, lacking some of the additives as discussed in this manuscript. Libraries, after being transformed into the knock-out strain (BL21 (DE3) *ΔilvAΔmetA*) with the accompanying plasmids, were briefly grown up in the nonselective media at 37°C. Upon reaching an $A_{600}$ of 0.5 or higher, the cells were harvested via centrifugation (3500 rpm for 5 minutes) and washed with ice cold PBS. The centrifugation and wash steps were repeated three times to ensure the cells were thoroughly washed of the previous media. The final, washed cell pellet was resuspended and added to 10 mL of selective media at an $A_{600}$ of 0.1 to 0.01, depending on the library size and passage number. Upon reaching an $A_{600}$ of 0.5 or higher, the cells were harvested and washed, after which they were inoculated into nonselective media, to begin the next round of selection. A fraction of the harvest and washed cell pellet from each round was stored in 15% glycerol at -80°C for future selection and sequencing. Every third round was plated on ampicillin and spectinomycin selective plates. 10 colonies were randomly selected for Sanger sequencing in order to assess the evolution of the library.

| Component | Volume |
|---|---|
| 5x M9 Minimal Media Salts | 20 mL |
| 24% Sodium Pyruvate | 2 mL |
| 100 mM Homocysteine | 1 mL |
| 0.1 mg/mL 19 amino acid mix (minus isoleucine) | 40 mL |
| 1 M $MgSO_4$ | 0.2 mL |
| 0.1 M $CaCl_2$ | 0.1 mL |
| 10 mM Pyridoxine | 1 mL |
| 1 mg/mL Biotin | 0.1 mL |
| 1 mg/mL Thiamin | 0.1 mL |
| 100 mg/mL Ampicillin | 0.1 mL |
| 50 mg/mL Spectinomycin | 0.1 mL |
| 10 mg/mL Anhydrous Tetracycline | 0.1 mL |
| Water | 35.2 mL |
| Total | 100 mL |

Table 2.1 Selective Media Recipe

The recipe for the selective media utilized for the genetic selection in its final iteration. Previous iterations lacked various components, replacing them with water. The initial recipe utilized kanamycin instead of ampicillin and spectinomycin due to a difference in plasmid and therefore antibiotic resistance.

*Expression and Purification:* hCGL mutants identified via the genetic selection were expressed in order to assess the biophysical properties of the new mutations. In brief, the hCGL mutant, if not already in the pET28a plasmid, was excised via restriction enzyme digestion with NcoI and EcoRI from the pZE11 modified plasmid, and ligated into the pET28a plasmid via T4 ligation. The pET28a plasmid harboring the hCGL mutant was then transformed into BL21 (DE3) for expression. The cells were grown overnight at 37°C in 5 mL of 2xYT media with 50 μg/mL of kanamycin. The subsequent morning, the cultures were inoculated, at a 1:1000 dilution, into 0.5 L of fresh TB media containing 50 μg/mL of kanamycin. Upon reaching an $A_{600}$ of 0.6 to 0.8, the culture was moved to 25°C and 1 M isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM to induce T7 RNA polymerase expression, and subsequently the hCGL mutant expression. After overnight incubation, cells were harvested via centrifugation at 5,500

rpm for 20 minutes in a JA-10 rotor. The cell pellet, after removing the supernatant, was resuspended using 25 mL of pET lysis buffer (10 mM $NaH_2PO_4$, 30 mM imidazole, 300 mM NaCl, pH 8). The cell suspension was then lysed by two passes through a French pressure cell and then centrifuged at 14,000 rpm for 60 minutes in a JA-20 rotor. The supernatant from this centrifugation step was then filtered using a 0.45 μm membrane. The clarified supernatant was then applied to 2 mL of nickel-nitrotriacetic acid (Ni-NTA) resin that had been pre-equilibrated with 10 mL of pET lysis buffer. The resin was then washed with 60 mL (30 column volumes) of pET lysis buffer. Afterward, the resin was incubated with 5 mL of pET elution buffer (50 mM $NaH_2PO_4$, 250 mM imidazole, 300 mM NaCl, pH 8), equilibrated for 5 minutes, and then eluted, capturing the elution in a separate vessel. The elution step was then repeated to increase recovery. The elution was then buffer exchanged against PBS utilizing dialysis membranes over the course of 2 days, changing out the PBS every 12 hours. After completing the buffer exchange, glycerol was added to a final concentration of 10% v/v. The purified enzyme was then aliquoted, flash frozen with liquid nitrogen, and stored at -80°C. Purity of the proteins were analyzed via sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) on a 4% to 20% pre-cast Tris-Glycine gel run under reducing conditions (2.5% β-Mercaptoethanol (BME)). Protein concentration was determined via $A_{280}$ under denaturing conditions using a calculated extinction coefficient of 45,471 $M^{-1}cm^{-1}L$.

*Enzyme Activity Assay:* Kinetic analysis was achieved by measuring α-keto acid production from thiol substrates by reaction with 3-methyl-2-benzothiazolinone hydrazone hydrochloride (MBTH). The resulting product generates a chromophore, detectable at $\lambda_{320}$. Eppendorf tubes (1.5 mL) containing 220 μL of substrate in reaction buffer (100 mM $NaH_2PO_4$, 1 mM EDTA, pH 7.3) was pre-incubated at 37°C for 5 minutes. After which, 20.3 μL of enzyme solution was added and thoroughly mixed. After the pre-determined

reaction time, 26.7 μL of 50.0% w/v trichloroacetic acid was added to quench the reaction. After performing all the desired substrate concentrations, blanks, and controls in triplicate, 733 μL of MBTH solution ($2.7 \times 10^{-2}$ % w/v MBTH in a 1 M sodium acetate buffer, pH 5.0) was added and the mixture incubated at 50°C for 40 minutes. The mixture was then cooled for 15 minutes, transferred to cuvettes, and the $A_{320}$ was determined.

## 2.4 Results and Discussion

*hCGL Engineering & Library Construction*

Previous work identified a mutant of hCGL, hCGL-NLV, with improved activity towards homocysteine and homocystine (activity can be seen later in Figure 2.3). This variant has three mutations, E59N, R119L, and E339V, which are all around the active site and believed to be involved in the substrate binding/orientation. To engineer an improved therapeutic drug, hCGL-NLV was used as a template for further engineering. Table 2.2 displays the 8 libraries that were created, the targeted residues, the method for mutagenesis, and the rationale for the residues. In brief, the first library was based on the pFunkel method, allowing for a random sampling of the mutational landscape of hCGL-NLV[58]. Although a powerful method, since it theoretically contains all possible single mutations, the pFunkel method has the major limitation of not assessing the synergistic effects of multiple mutations. This can be problematic in cases where individual mutations have little or even adverse effects, but combine synergistically. To address this, additional libraries, allowing multiple amino acid substitutions in a single variant were designed. In order to assess the entire library, these libraries were restricted to targeted residues, requiring rational design. Libraries 2-6 all employ similar a rationale, as they all target highly unstructured regions of the enzyme near the active site. Recent work has shown these unstructured regions of an enzyme, due to their flexibility, can be vital to the activity of an enzyme[59]. Libraries 2-4 each target a different unstructured region, employing saturation mutagenesis to mutate the entire loop. Libraries 5 and 6 consider all three loops simultaneously, and, due to the large number of targeted residues could not utilize saturation mutagenesis. Instead, these libraries utilized partial saturation mutagenesis, with each residue's potential mutations dictated based on phylogenetic analysis. Library 5 utilized phylogenetic analysis of only CGL enzymes, while library 6 compared both CGL

and methionine-γ-lyase (MGL) enzymes. Library 7 and 8 also used targeted saturation mutagenesis based on rationale design. Instead of targeting unstructured regions of the enzyme, these libraries were focused on residues involved in the chemistry of the enzyme or the substrate binding. After creating these libraries, the genetic selection was utilized to identify mutants with improved activity for further characterization and subsequent mutagenesis.

| Library | Targeted Residues | Design | Method |
|---|---|---|---|
| 1 | All | Random | pFunkel |
| 2 | 51-55 | Rational | Complete Saturation |
| 3 | 329-333 | Rational | Complete Saturation |
| 4 | 346-350 | Rational | Complete Saturation |
| 5 | 50-58, 329-333, & 346-350 | Phylogenetic | Partial Saturation |
| 6 | 50-58, 329-333, & 346-350 | Phylogenetic | Partial Saturation |
| 7 | 59. 119, & 339 | Rational | Complete Saturation |
| 8 | 59, 116, 119, 241, & 242 | Rational | Complete Saturation |

Table 2.2 hCGL-NLV Libraries

A table of the libraries designed based on the hCGL-NLV scaffold. Listed are the targeted residues, the design method, and the mutagenesis method.

*Genetic Selection of Engineered Libraries*

Utilizing the aforementioned genetic selection and the previously reported minimal media, the engineered bacteria transformed with the constructed libraries were grown in the minimal media and selected for improved mutants[57]. Given the large diversity of mutants in each library (approximately $10^8$ variants), mutants with improved activity towards homocysteine, homocystine, methionine, or some combination of the three substrates were expected to be contained in each library. The sequences of a few of the clones that were enriched after growth for 2 weeks are presented in Table 2.3. The respective enzymes were purified and their kinetics with different thiol containing amino

acids were determined and compared to those of the parental template, hCGL-NLV (Figure 2.3). Unfortunately, none of the enzyme variants encoded by the enriched clones had improved kinetics towards any of the tested substrates. It is possible that the reason why these clones were enriched despite their lower catalytic activity is because they are either expressed at a higher level or because of higher stability within the cell (again resulting in a higher intracellular concentration of enzyme) and thus increased α-ketobutyrate production. We showed that for all of the enriched variants, the expression yield, based

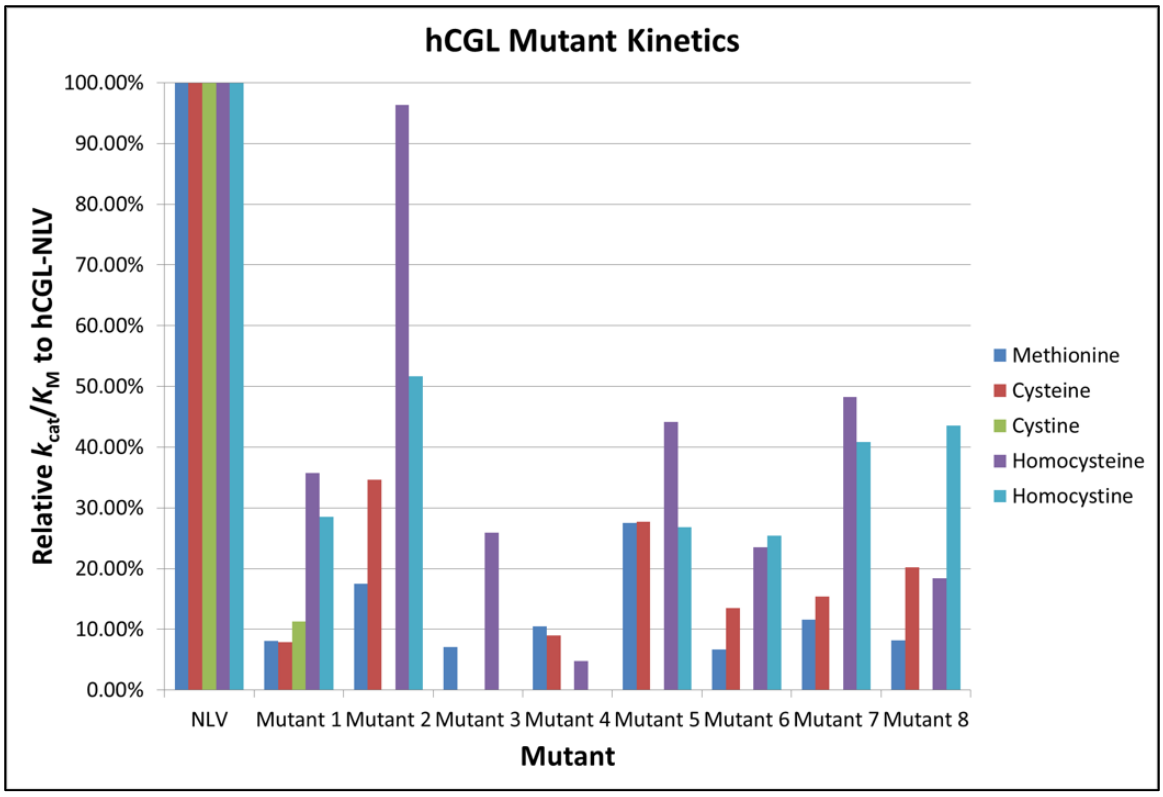


Figure 2.3 Kinetics of the hCGL Mutants

The kinetics of various mutants selected by the genetic selection. Mutations of the mutants can be seen in the accompanying table, Table 2.3. Each color represents a different substrate. The kinetics are normalized relatively to the kinetics of hCGL-NLV for that specific substrate.

| Mutant | Mutations |
|---|---|
| 1 | hCGL-NLV L350C |
| 2 | hCGL-NLV L294C |
| 3 | hCGL E59N R119D E339V |
| 4 | hCGL-NLV G116C N241D |
| 5 | hCGL-NLV G50D P52L Q54E H55E L329V N331S A348V |
| 6 | hCGL-NLV G50R A51S Q54I S65Q F58Y L329C N331S A348V |
| 7 | hCGL-NLV L337P |
| 8 | hCGL R119S E339P |

Table 2.3 Mutations of the Selected Mutants

The mutations of the mutants shown in Figure 2.3.

on Ni-NTA purification, was comparable to, or worse than, that of hCGL-NLV. Similarly these variants had comparable, or worse, thermal stability to hCGL-NLV. This analysis suggested that the genetic selection failed to enable enrichment of enzyme mutants with superior activity. To further investigate this hypothesis, cells expressing selected enriched variant enzymes, were co-cultured in the selective media to determine their relative fitness, i.e. whether a certain variant with lower catalytic activity somehow confers better growth relative to a more active variant. This study reveal that, as hypothesized, cells expressing enzyme variants with lower activity do enrich, usually with ~15 generations. Thus confounding biochemical processes distinct from α-ketobutyrate generation must have dominated the selection.

*Implications of Disrupting the Branched Chain Amino Acid Biosynthesis Pathway*

As discussed previously, there are a litany of reasons why a genetic selection may fail, and unfortunately the BCAA biosynthesis pathway contains many of these hurdles. Often one of the first examples given when discussing metabolic regulation, the enzymes of the BCAA biosynthesis pathway are known to be allosterically regulated. Additionally, flux through the pathway is further controlled by transcriptional regulation feedback

loops[60–63]. A schematic of relevant features of the BCAA pathway can be seen in Figure 2.4. In wild type *E. coli*, threonine deaminase (*ilvA*) occupies a critical role in the feedback regulation of the BCAA pathway. The *ilvA* enzyme is allosterically regulated by both valine and isoleucine, which activate and inhibit it respectively. The inhibition of *ilvA* by isoleucine reduces α-ketobutyrate production, causing the BCAA biosynthesis pathway to favor production of valine and leucine. Alternatively, activation of *ilvA* by valine enhances α-ketobutyrate production favoring production of isoleucine. Very simply, the two metabolites function to divert the BCAA pathway to produce whichever metabolite is less abundant. In the Blimko strain, because *ilvA* is deleted, this allosteric regulation by valine

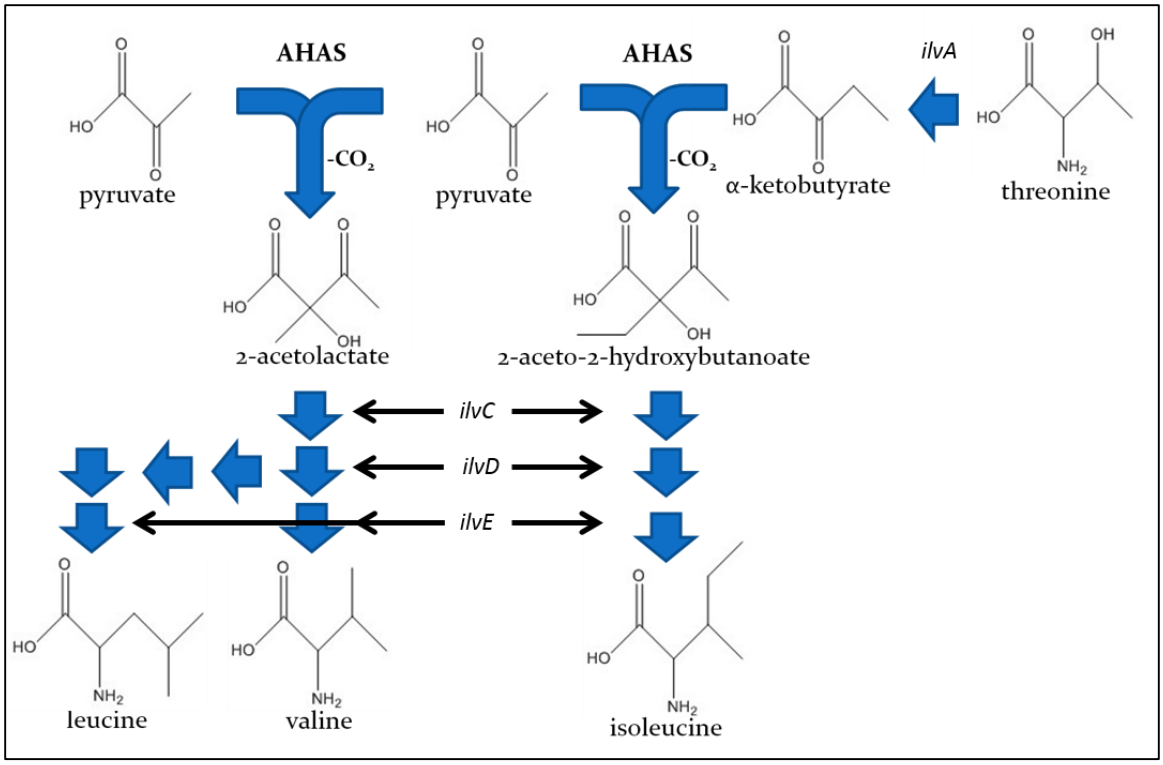


Figure 2.4 Branched Chain Amino Acid (BCAA) Biosynthesis Pathway

Above is an abridged version of the BCAA biosynthesis pathway as is relevant to this discussion. AHAS = acetohydroxy acid synthase

and isoleucine is abolished, disrupting control over BCAA levels. Unfortunately, this disruption does not completely explain the aberrant behavior observed in the genetic selection.

Flux through the BCAA biosynthesis pathway is also subject to additional allosteric and transcriptional regulation. Another critical enzyme in the pathway is acetohydroxy acid synthase (AHAS), which is shared by both the valine/leucine biosynthesis pathway and the isoleucine biosynthesis pathway[64]. AHAS catalyzes the carboligation of two α-keto acids, either consuming two pyruvate molecules to form 2-acetolactate, the precursor to valine and leucine, or consuming a pyruvate and an α-ketobutyrate molecule to form 2-aceto-2-hydroxybutanoate, the precursor to isoleucine. In *E. coli*, there are three AHAS isozymes: AHAS I (*ilvBN*), AHAS II (*ilvGM*), and AHAS III (*ilvIH*)[65]. The three AHAS isozymes have different kinetics and are subject to differential allosteric and transcriptional regulation[66]. A summary of what is known regarding the regulation of the AHAS isozymes as well as their substrate selectivity is presented Tables 2.4 and 2.5. AHAS II is unique relative to the other two isozymes in that it is not allosterically inhibited by any of the pathway end-products. Additionally, AHAS II has the highest selectivity for 2-aceto-2-hydroxybutanoate synthesis, meaning it increases metabolic flux towards the pathway for isoleucine biosynthesis. Early studies reported that the growth of *E. coli* K-12 strains is inhibited in minimal media supplemented with valine but not isoleucine[67] This is a consequence of the allosteric inhibition of AHAS I and III by valine and the fact that *E. coli* K-12 strains have a missense frameshift in the *ilvG* gene, disrupting expression of AHASII[65,68–70]. As a result, under these conditions the cells are unable to synthesize isoleucine, further confirmed by the lack of growth inhibition when *E. coli* K-12 strains are grown in rich media containing both isoleucine and valine. This frameshift mutation in *ilvG* is not present in *E. coli* B strains, however, AHAS II activity is not expressed in these

cells[71]. As with the metabolic disruption caused by the deletion of *ilvA*, this metabolic feedback loop could convolute the genetic selection, decoupling growth from the desired enzymatic activity. Unfortunate it still does not completely explain the results observed in the initial experiments, even when considered alongside the metabolic disruption caused by the *ilvA* knock out.

| Isozyme | Gene (map location) | Expression Repressed by: | Expression Induced by: | Feedback Inhibition[72,73] |
|---|---|---|---|---|
| AHAS I | *ilvBN* (82 min) | Valine Leucine | Cyclic AMP | Valine: $2 \times 10^{-4}$ M Isoleucine: $3 \times 10^{-3}$ M Leucine: $3 \times 10^{-2}$ M |
| AHAS II | *ilvGM* (84 min) | Valine Leucine Isoleucine | | |
| AHAS III | *ilvIH* (2 min) | Leucine | | Valine: $2 \times 10^{-5}$ M Isoleucine: $3 \times 10^{-3}$ M Leucine: $4 \times 10^{-2}$ M |

Table 2.4 Allosteric and Transcriptional Regulation of AHAS Isozymes

The allosteric and transcriptional regulation of AHAS isozymes. The reported concentration for feedback inhibition is for 50% inhibition under in vitro assay conditions. Table adapted from: Barak et al., 1987, *Journal of Bacteriology*. Data from various publications[74,75]

| Enzyme | pH | $K_M$ (mM) for pyruvate[a] | R[b] |
|---|---|---|---|
| AHAS I[c] | 6.5 | 1 | 1.3 |
| | 7.6 | | 1.0 ± 0.3 |
| | 8.5 | | 1.5 |
| AHAS II[d] | 7.6 | 10.6 | 185 |
| | 8.5 | | 175 |
| AHAS III[e] | 7.6 | 7 | 60 ± 10 |

Table 2.5 Selectivity of AHAS Isozymes

The selectivity of AHAS isozymes at various conditions. Table adapted from: Barak et al., 1987, *Journal of Bacteriology*.

[a] $K_M$ for pyruvate as the sole substrate under standard conditions

[b] Specificity ratio as defined by:

$$R = \frac{\frac{rate\ of\ 2-aceto-2-hydroxybutanoate\ formation}{rate\ of\ 2-acetolactate\ formation}}{\frac{2-ketobutyrate}{pyruvate}}$$

[c] Partially purified from *E. coli* K-12

[d] Purified from *E. coli* K-12

[e] Purified from *S. typhimurium*

Beyond the metabolic regulation observed in the pathway, it is possible that the genetic selection failed because the recombinant enzyme, hCGL, is either consuming essential metabolites or producing undesirable products. hCGL produces an α-keto acid as one of its products, as shown in Figure 2.1. However close examination of the reaction pathway reveals that hGCL produces a reactive enamine, 2-aminocrotonate or 2-aminoacrylate, which then readily react with water to form an α-keto acid. In the intracellular context of *E. coli*, the molarity of water is reduced relative to dilute solutions studied in vitro because the cytoplasm is a highly concentrated environment. Therefore,

the resolution of this reactive enamine may not occur as rapidly as expected. In fact a family of enzymes that catalyze this enamine/imine hydrolysis reaction has been discovered in other organisms[76]. Disruption of *yjgF*, one member of this family, in *S. enterica*, when coupled with induced high enamine levels, has been shown to arrest growth by inactivating multiple enzymes involved in amino acid biosynthesis[76]. In the context of the genetic selection, due to the high recombinant protein expression and relatively high substrate levels, it is possible that this reactive molecule accumulates at higher levels, providing an explanation for why more active catalytic variants failed to be enriched in the genetic selection.

In addition to 2-aminocrotonate, its downstream product, α-ketobutyrate, has been shown to cause complete growth inhibition when added at high levels to minimal media[77,78]. Further investigation revealed that α-ketobutyrate caused growth inhibition in M9 minimal media for a large number of carbon sources[77]. In the presence of exogenous α-ketobutyrate, cells accumulated guanosine pentaphosphate (ppGppp), an alarmone now known to be associated with amino acid starvation[79]. Ultimately, because of the partial alleviation of this growth inhibition by certain carbon sources and the presence of ppGppp, the authors concluded that α-ketobutyrate was disrupting the phosphoenolpyruvate (PEP) sugar phosphotransferase system. Thus accumulation of α-ketobutyrate in the cell may inhibit growth and contribute to the observed failure of the genetic selection to enrich for desired mutant enzymes with higher catalytic activity.

In summary[60], in *E. coli* BL21 grown on glucose, AHAS III is expressed, which favors the isoleucine pathway over the valine/leucine pathway as a consequence of its catalytic specificity. As isoleucine levels increase, *ilvA* is allosterically inhibited, resulting in a decrease in α-ketobutyrate levels. This inevitably causes AHAS III to favor valine and leucine production. The subsequent increase in valine levels then either 1) inhibits AHAS

III activity if all BCAAs are abundant, or 2) activates *ilvA* if isoleucine is depleted, effectively resetting the cycle. Exogenous, excess α-ketobutyrate causes toxicity via the disruption of this feedback loop, preventing AHAS III from switching from 2-aceto-2-hydroxybutyrate production to 2-acetolactate production. This causes a synthetic auxotrophy for valine and leucine, which can cause an increase in ppGppp, as previously observed, due to the amino acid deficiency. Additionally, the presence of valine in minimal media, without isoleucine, causes an auxotrophy for isoleucine due to the inhibition of AHAS III. Lastly, reactive enamine generation might become toxic when hGCL is overexpressed and high concentration of 2-aminocrotonate accumulate in the cell.

*Troubleshooting the BCAA Biosynthesis Pathway*

Previous work showed that the Blimko strain accurately enriched for pMGL over hCGL in the proof of concept experiments[57]. To further understand what might be limiting the dynamic range of the selection scheme we compared the fitness of mutants with different catalytic activities namely hCGL, hCGL-NLV and hCGL-IAV8. The kinetics of these mutants are reported in Table 2.6. hCGL-IAV8 has nearly 3-fold higher activity relative to hCGL-NLV for homocystine and has approximately fifty percent higher activity towards homocysteine. To determine the fitness conferred by each enzyme in the context of the genetic selection, competition experiments were conducted as described below. Blimko cells expressing the aforementioned variants were co-inoculated into the selective media. After ~15 generations, when inoculated at a 1:1 ratio, 10 out of 10 colonies contained the hCGL-IAV8 variant. After ~60 generations, when inoculated at a 1:1,000 ratio, 10 out of 10 colonies contained the hCGL-IAV8 variants. However Western blot analysis revealed that hCGL-NLV is expressed at nearly three times the level of hCGL-IAV8 in rich media. Thus, hCGL-NLV has a higher specific activity when considering the difference in relative expression, disproving the previously reported proof of concept

experiments. Additionally, this explained the results from the library screening. To further test this, hCGL-NLV was competed against hCGL-NDV, in a similar fashion. Unlike hCGL-IAV8, hCGL-NDV expresses at comparable levels to hCGL-NLV. When inoculated at a 1:1 ratio, 10 out of 10 colonies harbored the hCGL-NDV plasmid after ~15 generations. When inoculated at a 1:1,000 ratio, 10 out of 10 colonies harbored the hCGL-NDV plasmid after ~60 generations. In light of these experiments, determining the convoluting variables became the next priority.

| | | Methionine | DL- Homocysteine | Homocystine |
|---|---|---|---|---|
| hCGL-NLV | $k_{cat}$ ($s^{-1}$) | 9.2 | 7.2 | 4.5 |
| | $K_M$ (mM) | 12 | 0.91 | 0.59 |
| | $K_{cat}/K_M$ ($s^{-1}mM^{-1}$) | 0.75 | 8 | 7.6 |
| hCGL-IAV8 | $k_{cat}$ ($s^{-1}$) | 9.8 | 10 | 8.2 |
| | $K_M$ (mM) | 1.8 | 0.85 | 0.38 |
| | $K_{cat}/K_M$ ($s^{-1}mM^{-1}$) | 5.3 | 12 | 22 |
| hCGL-NDV | $k_{cat}$ ($s^{-1}$) | 1.1 | 6.8 | ND* |
| | $K_M$ (mM) | 22 | 3.3 | ND* |
| | $K_{cat}/K_M$ ($s^{-1}mM^{-1}$) | 0.05 | 2.1 | ND* |

Table 2.6 Kinetics of Proof of Concept hCGL Mutants

The kinetic properties of three hCGL mutants for methionine, DL-homocysteine, and homocystine. *ND = not detected (due to low activity)

From the previous section, there were three potential issues with the genetic selection to inspect: 1) α-ketobutyrate induced auxotrophy for valine and leucine, 2) α-ketobutyrate induced disruption of the PEP sugar phosphotransferase system, and 3) 2-

aminocrotonate mediated disruption of amino acid biosynthesis. To address the first problem we examined the growth of cells expressing hCGL-NLV in media with homocysteine in the presence of various amounts of α-ketobutyrate (Figure 2.6). Concentrations of less than 1 mM α-ketobutyrate did not cause any growth inhibition, but instead actually caused an increase in the $A_{600}$. However in the presence of 1 mM α-ketobutyrate, growth was significantly decreased at the first time point, although growth recovered by the second time point, consistent with the transient growth inhibition observed in the previously discussed experiments. At higher concentration of α-ketobutyrate, the growth inhibition is never overcome. Simultaneously, work was done to test the 2-aminocrotonate hypothesis as well as the valine/leucine auxotrophy hypothesis.

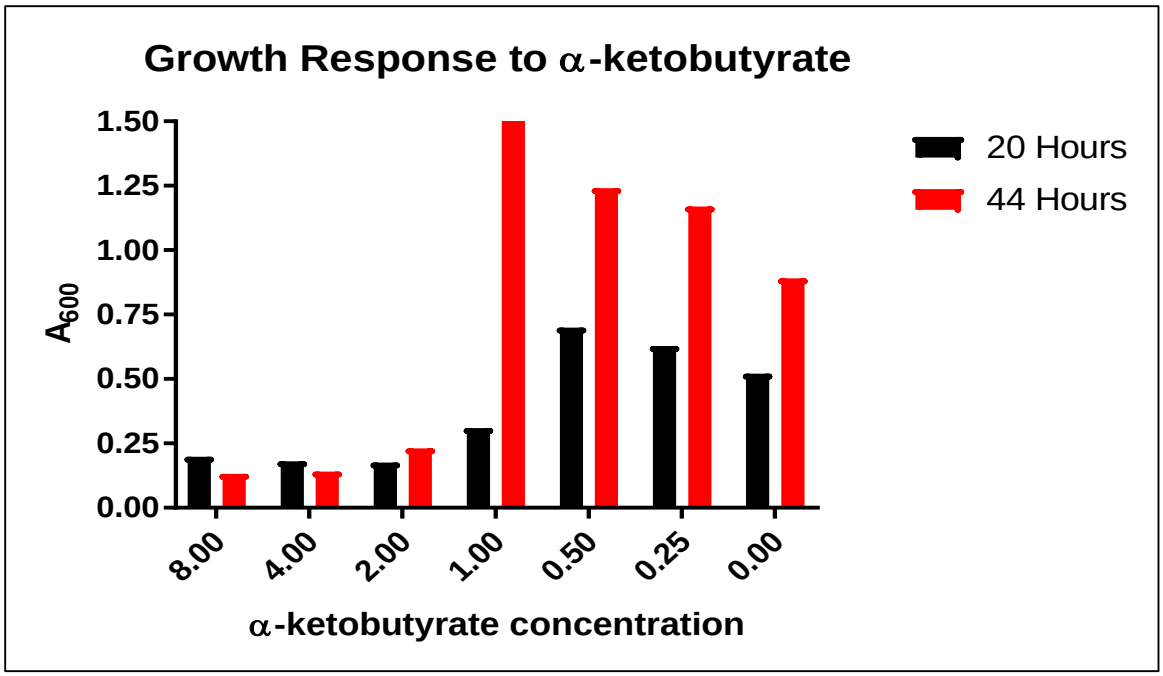


Figure 2.5 Growth Response to α-ketobutyrate

The growth of the Blimko strain, under the conditions of the original genetic selection, upon addition of exogenous α-ketobutyrate. Growth was assessed by measuring absorbance at 600 nm. Growth was assessed 20 and 44 hours after the initial inoculation.

In order to address the potential 2-aminocrotonate toxicity, a number of experiments were conducted. The competition experiments between Blimko cells expressing hCGL-NLV and Blimko cells expressing hCGL-NDV were repeated with a modified selective minimal media. All of the amino acids, except the BCAAs, were supplemented in an attempt to alleviate possible toxicity caused by 2-aminocrotonate. Unfortunately, the addition of these amino acids did not change the fitness of the Blimko cells, and thus, the cells expressing hCGL-NDV enriched as before. Next, recombinant expression of *yjgF* was attempted. *yjgF* is also known as reactive intermediate deaminase A (*ridA*), and catalyzes the hydrolysis of reactive enamines, such as 2-aminocrotonate[80]. By overexpressing *yjgF*, the hope was to provide an additional sink for 2-aminocrotonate. The competition experiments were repeated again, with the addition of *yjgF* expression. Again, the fitness of the cells was not affected as desired, and the Blimko cells expressing hCGL-NDV enriched as before. (An important aside, although not discussed in this work, simultaneous work was done adapting this genetic selection for a serine/threonine deaminase enzyme. Similar to the selection for hCGL, the aim of this work was to couple the generation of α-ketobutyrate to growth, serving to identify enzymes with superior activity. In this work it was found that expression of *yjgF* did alter the selection pressure, reducing the number of generation required for enrichment by nearly half. Unfortunately, the selection with *yjgF* expression enriched for the exact threonine deaminase mutant as was enriched for without *yjgF* expression. In spite of this, the results suggested that *yjgF* expression did alter the selection. One possible explanation for this accelerated selection is that *yjgF* accelerated the resolution of 2-aminocrotonate to α-ketobutyrate and thus allowing high activity mutants produce α-ketobutyrate more quickly. As such, the high activity mutants caused higher toxicity, causing the cells to die earlier in the selection, and increasing enrichment for the other mutants. Ultimately, the *yjgF* experiments did not

correct the genetic selection, but did provide evidence of the presence of 2-aminocrotonate.) In a final attempt to address any toxicity caused by 2-aminocrotonate, the hCGL mutants were cloned into a new plasmid in order to reduce expression and thus reduce 2-aminocrotonate generation. The pZE11 plasmid, which contains a tetracycline induced lambda phage promoter, was chosen because it allows for tight, titratable regulation of expression. The competition experiments were repeated again, utilizing the new pZE11 plasmids. Again, the fitness of the cells was not affected as desired, and the Blimko cells expressing hCGL-NDV enriched as before.

*Re-engineering the BCAA Biosynthesis Pathway*

Finally, recombinant expression of AHAS II was attempted in order to alleviate the growth arrest associated with valine. The two genes, *ilvG* and *ilvM*, were cloned into the pCDFDuet-1 plasmid. After transforming this new pCDFDuet-1-*ilvG*-*ilvM* plasmid into the Blimko strain, cells were grown in minimal media with 1 mM methionine (required for cell growth given the *metA* deletion), and with varying concentrations of valine/leucine as well as α-ketobutyrate. An empty plasmid control was included alongside the induced expression and repressed expression experiments. The results of the experiment were treated as binary, testing whether growth was observed after 48 hours, and can be seen in Table 2.7. As expected, the empty plasmid only grew when a low concentration of α-ketobutyrate was present without any valine or leucine. None of the strains grew when α-ketobutyrate was not present as expected as well. Induction of AHAS II caused complete growth inhibition under every media condition. Fortunately, leaky AHAS II expression showed significant improvements, growing under every condition except high α-ketobutyrate without valine or leucine. For the first time, the genetic selection strain was not inhibited by high α-ketobutyrate conditions. Additionally, this experiment suggested that the toxicity associated with α-ketobutyrate is due to the induced auxotrophy for valine

and leucine. To further improve these results, a way to avoid the toxicity associated with AHAS II overexpression was required.

| | Valine/Leucine Conc. (mM) | α-ketobutyrate Conc. (mM) | Empty Plasmid | Repressed AHAS II | Induced AHAS II |
|---|---|---|---|---|---|
| 1 | 0 | 0 | No | No | No |
| 2 | 0 | 1 | Yes | Yes | No |
| 3 | 1 | 1 | No | Yes | No |
| 4 | 10 | 1 | No | Yes | No |
| 5 | 0 | 10 | No | No | No |
| 6 | 1 | 10 | No | Yes | No |
| 7 | 10 | 10 | No | Yes | No |

Table 2.7 Growth of Blimko AHAS II Strain under Various Media Conditions

The Blimko strain harboring the pCDFDuet-1-*ilvG-ilvM* plasmid was tested for growth inhibition under a variety of media conditions. The results were treated as binary, simply surveying for growth after 48 hours in the respective media.

The cause of the AHAS II overexpression toxicity was not immediately apparent. As with the overexpression of any protein via the bacteriophage T7 RNA polymerase expression system, toxicity can be caused by high expression and subsequent misfolding[81]. In this specific case, AHAS II lacks the allosteric regulation observed in its isozymes, and, by driving expression via a non-native promoter, lacks the typical transcriptional regulation. As a result, the flux through this pathway is unregulated. This could cause a depletion of pyruvate, leading to the observed growth inhibition. Alternatively, it is possible that overexpression of AHAS II causes overproduction of the BCAAs, altering other transcriptional and allosteric regulation. In order to test the pyruvate hypothesis,

Blimko cells harboring the AHAS II plasmid were tested in minimal media containing pyruvate as the carbon source instead of glucose. Repeating select media conditions, such as media #3, it was found that the pyruvate-based media did indeed correct the toxicity associated with AHAS II induced expression. It should be noted, altering the carbon source to pyruvate, an inherently worse carbon source for *E. coli*, may have corrected the toxicity by slowing down the metabolism of the cells rather than supplementing an auxotrophy. Regardless of the exact mechanism, the pyruvate-based minimal media supplemented with valine and leucine, coupled with the pCDFDuet-1-*ilvG-ilvM* plasmid, allowed the Blimko strain to grow well under high α-ketobutyrate concentrations.

*Re-proving the Genetic Selection*

The two previously used mutants, hCGL-NLV and hCGL-NDV, were competed in the new genetic selection strain utilizing the new selective media (M9 minimal media with pyruvate, valine, and leucine). Blimko cells harboring the AHAS II plasmid and the hCGL-NLV plasmid, as well as Blimko cells harboring the AHAS II plasmid and the hCGL-NDV plasmid, were co-inoculated in the pyruvate-based selective media and allowed to grow. Enrichment was checked by sequencing every ~15 generations. Unfortunately competing the variants under this conditions still did not result in the desired enrichment, as hCGL-NDV, again, outgrew its superior counterpart, hCGL-NLV. As previously mentioned, simultaneous work was done adapting this selection for a serine/threonine deaminase, in which experiments suggested the toxicity of 2-aminocrotonate and 2-aminoacrylate was a contributing factor in the genetic selection. Therefore, the competition was repeated, supplementing with all of the amino acids except isoleucine, in hopes that these supplements could counteract any amino acid biosynthesis toxicity caused by 2-aminocrotonate. After ~15 generations, hCGL-NLV outgrew hCGL-NDV, suggesting the inverted correlation between activity and fitness had been finally corrected. Subsequent

experiments showed that hCGL-NLV enriched over hCGL-NDV when mixed at 1:100 and 1:1,000 ratios.

*Final Library Selection, Concluding Thoughts, and Future Directions*

The libraries previously discussed were recloned into the pZE11 plasmid for selection under the new system. For the initial tests, libraries 1 and 7, due to their smaller size, were tested to ensure reproducible results. After ~40 generations, the libraries converged to their respective variants, which were purified and characterized. Unfortunately, the new variants had worse catalytic activity compared to hCGL-NLV, the parental template. The selection was repeated to account for any possible experimental errors, and again, the worse variants were identified. Attempts at altering this selection pressure by changing the inducer concentration did change the identified mutant, but it was still inferior. Fortunately, utilizing the selection for the serine/threonine deaminase has shown promise though, as, under the same conditions, superior mutants have been successfully identified.

The difference in success for these two enzyme scaffolds suggests the failure to identify superior hCGL mutants could lie with the enzyme itself, rather than the strain and selection process. Based on this assumption, two major possibilities remain: 1) growth inhibition of high activity mutants is caused by substrate depletion or unintended side reactions, or 2) growth inhibition of high activity mutants is caused by *E. coli* regulation of homocysteine intracellular concentrations. In either case, further correcting the genetic selection may not be possible. In the former case, the minimal media is already supplemented with both cysteine and methionine, the two major alternative substrates for hCGL. In the latter case, homocysteine has been shown to cause toxicity in *E. coli*, ironically by disrupting *ilvA* activity, which would not affect this knock out strain[82]. Alternatively, thiol molecules, and correspondingly the redox state, are highly regulated in

most cells, including *E. coli*[83,84]. As such, the transport of homocysteine is limited to avoid disrupting this balance, which would result in low homocysteine concentrations and would limit enzyme activity. Alternatively, rather than the enzyme of interest being the problem, it is possible that α-ketobutyrate has not been thoroughly coupled to growth rate. In the proof of concepts, as well as the threonine deaminase project, the enriched mutants may have hit a sweet spot of activity, which was not achieved in the library screening. Figure 2.6 is a comprehensive figure showing the typical regulation of the BCAA biosynthesis pathway, the genetic selection and its various modifications, and the possible remaining issues preventing prior selection.

Regardless of the success in identifying superior hCGL mutants, this work has further elucidated the regulation and mechanisms of the BCAA biosynthesis pathway. Additionally, this work has developed libraries for future homocyst(e)inase screening as well as improved the genetic selection for identifying other α-ketobutyrate generating enzymes. Moving forward, the selection, as is, could be utilized as a live-dead screen. In doing so, inactive, misfolded, and low-activity mutants could be selected against, after which a 96-well plate screen could be utilized to identify superior mutants. In order to understand the selection further proteomics and metabolomics could be utilized to identify differences between the various media conditions as well as for different hCGL mutants. In doing so, a better understanding of the differences could be gathered, directing future engineering work.

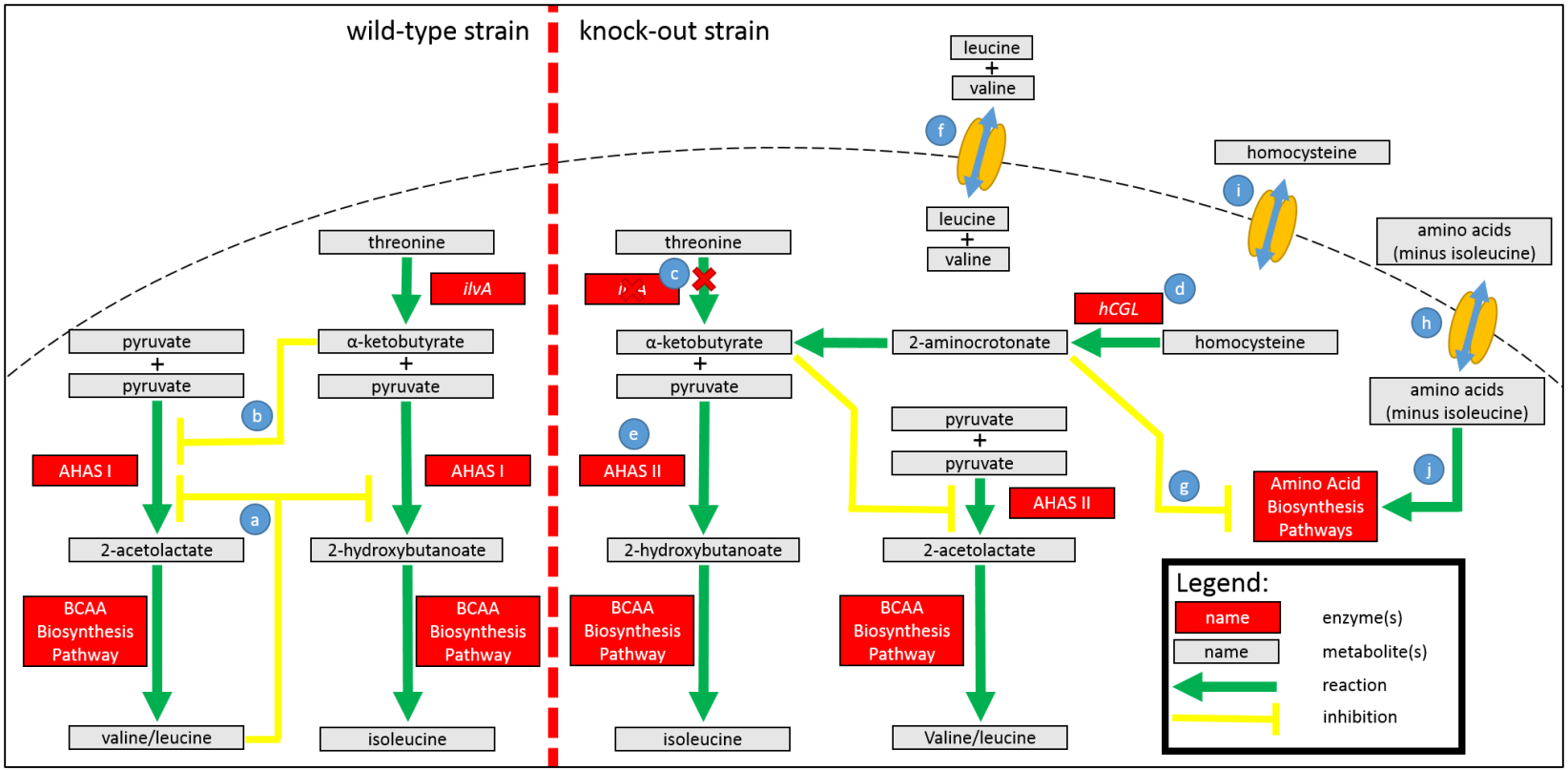


Figure 2.6 Overview of Final Genetic Selection

Overview of the branched chain amino acid biosynthesis pathway, the Blimko strain, and the various modifications made to the genetic selection process. On the left hand side, the wild-type BCAA biosynthesis pathway is shown, with the two major sources of regulation: a) valine/leucine inhibition of AHAS I and b) α-ketobutyrate inhibition of 2-acetolactate synthesis, limiting valine/leucine biosynthesis. On the right hand side, the knock-out strain is shown. c) The knock-out of *ilvA* prevents generation of α-ketobutyrate from threonine. d) Recombinant expression of hCGL allows for the generation of α-ketobutyrate from homocysteine. e,f) Expression of AHAS II, an alternative isozyme not natively expressed in most *E. coli* strains, lacks the allosteric regulation observed in AHAS I, allowing for valine and leucine to be supplemented in the media. g,h) Generation of 2-aminocrotonate by hCGL can limit growth by inhibiting amino acid biosynthesis pathways; supplementing the other amino acids, other than isoleucine, overcame this toxicity. i) It is possible that homocysteine transport limits the intracellular concentration thus limiting high activity mutants and altering the selection pressure. j) alternatively, hCGL could degrade essential metabolites, methionine and cysteine, limiting growth.

**ACKNOWLEDGEMENTS**

I would like to acknowledge Ebru Cayir for her contributions to testing the various modifications to the genetic selection and discussions for re-engineering the selection. I'd also like to acknowledge Emma Rowlinson, Jenny Kim, Grace Keliher, Chelsea Cuellar, and Yawen Ren for their assistance in purifying and characterizing the various mutants identified throughout this project.

# Chapter 3: Prevention of homocystinuria disease progression in a neonatal lethal murine model through systemic homocysteine depletion with a therapeutic enzyme

## 3.1 ABSTRACT

Homocystinuria, an inborn error of metabolism (IEM) disease of the transsulfuration pathway, is estimated to affect 20,000 to 40,000 people worldwide[28] and is primarily caused by defects in the cystathionine β-synthase (CBS) gene product. Impaired CBS activity results in the accumulation of homocysteine and methionine, and depletion of cystathionine and cysteine. This metabolic imbalance affects the ocular, skeletal, vascular, connective, and central nervous systems, ultimately reducing quality of life and often causing premature death[32]. Current treatments that are partially effective include protein restricted diets and vitamin B6 administration but due to their reliance on patient compliance do not often adequately control disease. Recent work has highlighted a potential CBS enzyme replacement therapy (ERT) for homocystinuria[41], demonstrating that controlling the levels of circulating homocysteine improves disease symptoms in murine models. The work presented here unveils an alternative enzyme therapeutic for homocystinuria, using a human cystathionine γ-lyase (hCGL) engineered to degrade both homocysteine and its disulfide form homocystine. The data presented in this work show that the proposed treatment effectively and safely depletes total homocysteine (tHcy) levels without adversely affecting total cysteine (tCys) or methionine levels and completely prevents the neonatal lethality and pathological abnormalities observed in a murine model of homocystinuria. Our data suggests that controlling tHCY levels is key to ameliorating disease and that the engineered human homocyst(e)inase provides several therapeutic advantages over the proposed CBS ERT strategy.

### 3.2 INTRODUCTION

Classical homocystinuria is primarily caused by mutations in the cystathionine β-synthase (CBS) gene[32] and is estimated to affect between 20,000 to 40,000 people worldwide[28]. Non-classical forms of homocystinuria are caused by defects in metabolic pathways related to the transsulfuration pathway (Figure 3.1.) such as folate metabolism.

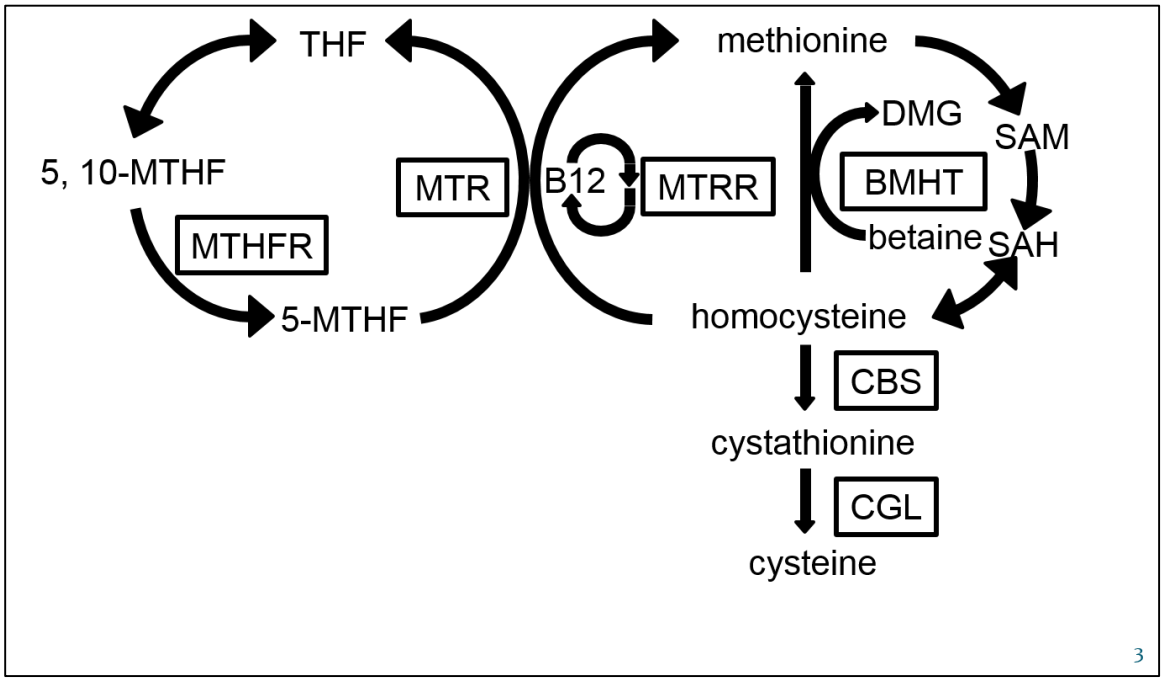


Figure 3.1 The Human Transsulfuration Pathway

The human transsulfuration pathway showing the metabolic pathway from methionine to cysteine. In homocystinuria, the cystathionine-β-synthase (CBS) enzyme is defective, resulting in elevated levels of homocysteine and methionine, concomitant with decreased levels of cysteine and cystathionine. **THF**: Tetrahydrofolate; **5,10-MTHF**: 5,10-Methylenetetrahydrofolate; **5-MTHF**: 5-Methyltetrahydrofolate; **B12**: Vitamin B12; **DMG**: Dimethylglycine; **SAM**: S-Adenosyl Methionine; **SAH**: S-Adenosyl Homocysteine; **MTHFR**: Methylenetetrahydrofolate reductase; **MTR**: 5-Methyltetrahydrofolate-Homocysteine Methyltransferase; **MTRR**: 5-Methyltetrahydrofolate-Homocysteine Methyltransferase Reductase; **BHMT**: Betaine-Homocysteine S-Methyltransferase; **CBS**: Cystathionine-β-synthase; **CGL**: Cystathionine-γ-lyase;

Although non-classical homocystinuria has some similarities in disease progression and symptoms, it ultimately require different treatment, and thus is largely excluded from this discussion. On a cellular basis, classical homocystinuria affects a critical point in the transsulfuration pathway, determining the fate of homocysteine and affecting the flux and subsequent regulation of this pathway. Homocysteine can be remethylated into methionine by the betaine-homocysteine s-methyltransferase (BHMT) enzyme[29]. Alternatively, homocysteine can be irreversibly committed to cysteine biosynthesis via the CBS enzyme[30,31]. Disruption of this unidirectional flow to cysteine results in elevated levels of homocysteine and methionine in patients, as well as decreased levels of cysteine and cystathionine[32,33].

Current homocystinuria treatments primarily focus on decreasing the concentration of total homocysteine (tHcy), with a secondary focus on increasing total cysteine (tCys) levels. Initially, patients are treated with pyridoxine (vitamin B6) supplementation, to which less than half are responsive[33,38]. Unresponsive or partially responsive patients begin a low methionine diet, along with cysteine supplementation, in an effort to normalize the metabolic imbalance[39]. Betaine treatment can further help normalize the metabolic imbalance, by increasing the flux through BHMT, but recent research has shown limited efficacy[40]. As often observed, patient compliance with restrictive/medical diets can be poor, resulting in uncontrolled disease; there is a clear need for an effective non-dietary treatment. Ideally, a next generation treatment for homocystinuria will limit patient burden, improve quality of life and remove dietary compliance issues.

As classical homocystinuria progresses, patients suffer from a variety of complications, including thromboembolic events, dislocation of ocular lenses, and spinal osteoporosis[28,34–37]. Patients first develop myopia, typically by the age of 1, which, if the disease was not detected via newborn screening, serves as the first indication of the disease.

Another common indication of classical homocystinuria is developmental delay, which is usually more severe in B6-non-responsive patients[28]. As the patient ages, the disease progresses with ocular issues escalating to lens dislocation. Further mental decline also correlates with severity of disease; B6-reponsive patients have a mean IQ of 79, while for B6-nonresponsive patients, the mean IQ is 57[28]. B6-nonresponsive individuals identified by newborn screening with good treatment compliance see an improvement in IQ, with a mean IQ of 105 for this group[28]. By or around puberty, patients start to display severe connective tissue problems as evident by their marfanoid appearance and osteoporosis. Cardiovascular distress is typically the last, but most severe, symptom to arise, resulting in 23% mortality by the age of 30 for B6-non-responsive patients[28].

Recent work presented a novel enzyme replacement therapy (ERT) for the treatment of classical homocystinuria[41–43]. Through intraperitoneal (IP) administration, CBS ERT for homocystinuria depleted tHcy levels while transiently increasing tCys and cystathionine levels in a transgenic murine model of homocystinuria (Tg-hCBS *cbs*$^{-/-}$)[41]. Importantly, despite the treatment being confined to the circulatory system, CBS ERT affected tissue health, suggesting that serum depletion of homocysteine, without tissue penetration, can effectively treat classical homocystinuria[41]. CBS ERT was also administered in a *cbs*$^{-/-}$ genetic murine model of homocystinuria to assess the therapeutic efficacy. The *cbs*$^{-/-}$ murine model differs drastically from the human disease, with almost all homozygous knock-out pups dying within 3-4 weeks of birth. Unlike the human disease, most organs are not affected adversely, with primarily the liver showing severe abnormalities. This drastic increase in severity is most likely due to the complete lack of CBS activity, whereas human patients typically have CBS mutations that result in an enzyme with residual activity. In transgenic murine models of homocystinuria, expression of hCBS mutants with reduced activity, alleviates the neonatal lethality observed in the

knock-out model, and instead, the mice present symptoms more closely resembling the human disease. In a preliminary work, administration of hCBS ERT in the knock-out mouse model partially rescued the neonatal lethality, with over 30% surviving past week 5[41,42]. Subsequent work further refined the treatment approach, resulting in over 90% survival concomitant with reduced liver pathology. Lastly, despite hCBS only being able to catalyze the condensation of serine with free, reduced homocysteine, the large depletion in tHcy suggests that the disulfide form, homocystine, and protein-bound forms were also affected by treatment. This suggests that the various thiol forms equilibrate relatively quickly, allowing hCBS to indirectly deplete homocystine and protein-bound homocysteine as they equilibrate upon the removal of free homocysteine.

In this work, rather than using a CBS ERT approach, a pharmacologically optimized cystathionine γ-lyase (hCGL) re-engineered to degrade homocysteine was used. The hCGL enzyme, which natively converts cystathionine into cysteine and other metabolites, is the last enzyme in the transsulfuration pathway. Due to its native enzyme chemistry, hCGL and its variants do not generate cystathionine, representing a major difference between this proposed therapeutic and the mechanism of action for the CBS ERT treatment. This work presents a variant of hCGL that successfully normalizes extremely elevated tHcy levels, completely prevents neonatal lethality in a murine model of homocystinuria, and repairs the liver abnormalities observed in these mice. Additionally, this work shows that serum-based tHcy depletion, without dietary or enzymatic tCys supplementation, is sufficient to rescue a murine model of homocystinuria, indicating that restoration of the transsulfuration pathway is not required to treat the disease. This work, to our knowledge, showcases the first time a metabolic disorder was corrected by a recombinant enzyme utilizing a novel catalytic reaction, a departure from the traditional

enzyme replacement therapy. More importantly, this work unveils a novel, promising therapeutic for homocystinuria patients.

### 3.3 MATERIALS AND METHODS

*Materials:* Unless otherwise noted below, chemicals were purchased from Sigma-Aldrich, cloning reagents from New England Biolabs, growth media from Becton Dickinson, chromatography reagents from Agilent Technologies, and mouse chow from Envigo, and mouse strains from the Jackson Laboratory.

*Expression and Purification:* hCGL mutants identified via the genetic selection were expressed in order to assess the biophysical properties of the new mutations. In brief, the hCGL mutant, if not already in the pET28a plasmid, was excised via restriction enzyme digestion with NcoI and EcoRI from the pZE11 modified plasmid, and ligated into the pET28a plasmid via T4 ligation. The pET28a plasmid harboring the hCGL mutant was then transformed into BL21 (DE3) for expression. The cells were grown overnight at 37°C in 5 mL of 2xYT media with 50 μg/mL of kanamycin. The subsequent morning, the cultures were inoculated, at a 1:1000 dilution, into 0.5 L of fresh TB media containing 50 μg/mL of kanamycin. Upon reaching an $A_{600}$ of 0.6 to 0.8, the culture was moved to 25°C and 1 M isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM to induce T7 RNA polymerase expression, and subsequently the hCGL mutant expression. After overnight incubation, cells were harvested via centrifugation at 5,500 rpm for 20 minutes in a JA-10 rotor. The cell pellet, after removing the supernatant, was resuspended using 25 mL of pET lysis buffer (10 mM $NaH_2PO_4$, 30 mM imidazole, 300 mM NaCl, pH 8). The cell suspension was then lysed by two passes through a French pressure cell and then centrifuged at 14,000 rpm for 60 minutes in a JA-20 rotor. The supernatant from this centrifugation step was then filtered using a 0.45 μm membrane. The clarified supernatant was then applied to 2 mL of nickel-nitrotriacetic acid (Ni-NTA) resin that had been pre-equilibrated with 10 mL of pET lysis buffer. The resin was then washed with 60 mL (30 column volumes) of pET lysis buffer. For use in mouse studies, the resin

was then washed with 100 column volumes of PBS with 1% triton X-114. Afterward, the resin was incubated with 5 mL of pET elution buffer (50 mM $NaH_2PO_4$, 250 mM imidazole, 300 mM NaCl, pH 8), equilibrated for 5 minutes, and then eluted, capturing the elution in a separate vessel. The elution step was then repeated to increase recovery. The elution was then buffer exchanged against PBS utilizing dialysis membranes over the course of 2 days, changing out the PBS every 12 hours. If the enzyme was to be used in a mouse study, it was then PEGylated using 100x molar excess, after which it was buffer exchanged again. After completing the buffer exchange, glycerol was added to a final concentration of 10% v/v. The purified enzyme was then aliquoted, flash frozen with liquid nitrogen, and stored at -80°C. Purity of the proteins were analyzed via sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) on a 4% to 20% pre-cast Tris-Glycine gel run under reducing conditions (2.5% β-Mercaptoethanol (BME)). Protein concentration was determined via $A_{280}$ under denaturing conditions using a calculated extinction coefficient of 45,471 $M^{-1}cm^{-1}L$.

*In vivo metabolite levels*. tHcy and tCys levels were analyzed via SBD-F derivatization on an Agilent 1200 series HPLC. Derivatization and chromatography were adapted from the referenced literature[85]. First 10 μL of sample (either plasma or controls dissolved in PBS), 5 μL of internal standard in PBS, and 5 μL of PBS were mixed. For the internal standard, cysteamine was used at a typical concentration of 100 μM. To this 20 μL mixture, 2 μL of 100 g/L tcis(2-carboxyethyl)phosphine (TCEP) dissolved in water was added. The resulting mixture was incubated at 25°C for 30 minutes to allow reduction of the disulfide bonds, after which 18 μL of 100 g/L trichloroacetic acid (TCA) dissolved in water with 1 mM EDTA was added and mixed. Next the mixture was centrifuged in a micro-centrifuge at 13,000 x g for 10 minutes. 22 μL of the resulting supernatant was transferred to a fresh Eppendorf tube. 4.4 μL of 1.55 M NaOH, 55 μL of borate buffer (125

mM borate, 4 mM EDTA, pH 9.5), and 22 μL of 1 g/L 7-fluorobenzofurazan-4-sulfonic acid (SBD-F) in borate buffer were added to the fresh Eppendorf. This final mixture was then incubated at 60°C for 60 minutes. The final derivatized mixture was then passed through a centrifugal unit to ensure no particulates clogged the HPLC, after which it was transferred to an HPLC vial for injection and analysis.

For the chromatographic separation of SBD-F derivatized thiols, an isocratic separation using a C18 column was utilized. The mobile was 97% 100 mM sodium acetate, pH 5.5 and 3% methanol. Between each run, the column was washed with a high organic flush. Cysteine eluted at ~2.8 minutes, homocysteine at ~3.7 minutes, glutathione at ~4.2 minutes, and cysteamine at ~5.5 minutes. These elution times vary depending on column, instrument setup, and slight run-to-run variation.

Methionine levels were analyzed via OPA derivatization on an Agilent 1200 series HPLC. Derivatization and chromatography were adapted from the referenced literature[86].

*Mice:* For the diet-induced hyperhomocysteinemia murine model, female C57BL/6J mice were obtained from the Jackson Laboratory. Mice began the decreased folate diet (Envigo, TD.98272) upon reaching 8 weeks of age and continued at least for 2 months. For the genetic homocystinuria murine model, *cbs*$^{+/-}$ mice were purchased from the Jackson Laboratory. Heterozygotes were bred together to obtain homozygous knockout (-/-) mice. The breeders for survival study were given water supplemented with 2% betaine throughout pregnancy and lactation. Genotyping was performed at the age of 8-10 days and homozygous knockout mice were enrolled into the study. All animal protocols were approved by the Institutional Animal Care and Use Committee (IACUC) at the University of Texas at Austin.

*ELISA:* IAV8-coated plates were prepared by incubating purified IAV8 enzyme in the well of ELISA plate and then wash and block with 2% milk PBST (PBS-0.05% tween)

at 4°C. Serial dilutions of mouse serum (3X series dilution) were then incubate in enzyme-coated plates and then analyzed to detect the titer of anti-IAV8 antibodies in the serum by using goat anti-mouse IgG antibody-peroxidase.

*Statistics*. All results are presented as mean ± SEM. For Kaplan-Meier plots of mouse survival, statistical significance was analyzed by the log-rank (Mantel-Cox) test. Statistical analysis of the various PD experiments was performed using two-way ANOVA followed by Bonferroni's multiple comparison test.

## 3.4 RESULTS AND DISCUSSION

*Administration of hCGL mutants in a diet-induced mouse model of hyperhomocysteinemia lowers tHcy without adversely affecting tCys*

Through the use of point mutations, previous work demonstrated that the selectivity of hCGL can be altered from the α,γ-elimination of cystathionine to instead favor the α,β-elimination of cysteine and cystine or the α,γ-elimination of methionine and homocysteine [20,87]. Two hCGL variants arising from these efforts were identified as potential homocystinuria drug candidates based on their activity and selectivity (Table 3.1). Ideally, a therapeutic based on hCGL would be capable of degrading homocysteine and homocystine, without degrading methionine, cysteine, or cystine. In doing so, the therapeutic would provide a sink for homocysteine and homocystine without adversely affecting cysteine and cystine, which are already depleted in homocystinuria patients. Additionally, although methionine can be elevated in patients, methionine levels begin to normalize upon reduction of homocysteine levels, and as such, a therapeutic without methionine activity avoids any toxicity from over depleting methionine levels. As seen in the table, IAV8, the first

| | L-Cysteine | Cystine | L-Methionine | DL-Homocysteine | Homocystine |
|---|---|---|---|---|---|
| | $k_{cat}/K_M$ ($mM^{-1}\ s^{-1}$) | $k_{cat}/K_M$ ($mM^{-1}\ s^{-1}$) | $k_{cat}/K_M$ ($mM^{-1}\ s^{-1}$) | $k_{cat}/K_M$ ($mM^{-1}\ s^{-1}$) | $k_{cat}/K_M$ ($mM^{-1}\ s^{-1}$) |
| WT-hCGL | 0.22 | 0.85 | ND* | 2.7 | 2.7 |
| IAV8 | 8.5 | 2.1 | 5.3 | 12 | 22 |
| IAV8-119D | 3.6 | 2.3 | 1.1 | 12 | 19 |

Table 3.1 Kinetics of hCGL and Engineered Mutants

The catalytic activity of hCGL and the engineered mutants used in this work.

identified mutant, has high activity towards both homocysteine and homocystine, but does not match the ideal substrate specificity profile since it also has relatively high catalytic activity towards methionine, cysteine, and cystine. Conversely, IAV8-119D has an improved substrate specificity profile at the expense of some activity towards homocystine. With these two variants, the efforts of this work examined the hypothesis that homocystinuria would be responsive to an enzyme-based treatment based on hCGL and that tHcy depletion, without an increase in tCys or cystathionine levels, is sufficient for a therapeutic effect.

First the pharmacodynamic (PD) properties of the above candidate enzymes, properties were assessed using a dietary-induced murine model of hyperhomocysteinemia[88]. By increasing the methionine concentration in the mouse feed, homocysteine levels were elevated over the course of two months, eventually surpassing 300 μM, emulating the concentrations found in human patients. This model only mimics the hyperhomocysteinemia symptom of homocystinuria, as the mice do not present with other symptoms typically seen in human patients. Additionally, because of the increased methionine consumption and a fully functional CBS enzyme, metabolic flux through the transsulfuration pathway is significantly altered from the disease state. As such this murine model is suitable for the assessment of tHcy PD properties but not for methionine PD properties nor for biological efficacy.

As research has shown that most pathological abnormalities are observed when tHcy levels exceed 100 μM, the threshold for severe hyperhomocysteinemia, we set this concentration as a target for PD studies[89]. Based on pilot dose ranging experiments, the mice were injected with a single 50 mg/kg intraperitoneal (IP) injection of either IAV8, IAV8-119D, or heat-inactivated enzyme as shown in Figures 3.2 and 3.3. The tHcy levels for day 0, as well as the control treatment, ranged from 250 to 350 μM, similar to

homocystinuria patients suffering from severe hyperhomocysteinemia. Administration of IAV8-119D depleted tHcy levels to 50 μM within the first day, while administration of the kinetically superior IAV8 depleted tHcy levels further to 25 μM. Both treatments maintained tHcy levels below the target concentration of 100 μM for multiple days. After 7 days, the tHcy levels rebounded to ~125 μM and ~100 μM for IAV8-119D and IAV8

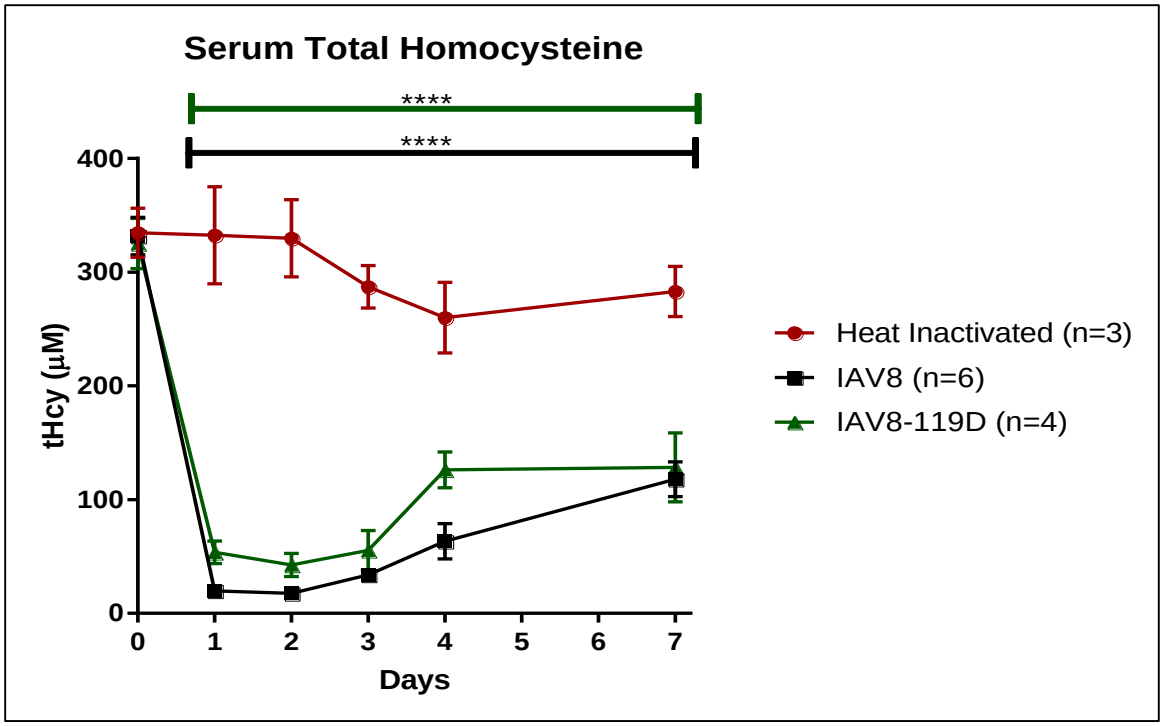


Figure 3.2 Serum Total Homocysteine upon Single Injection in Diet-Induced Hyperhomocysteinemia Mouse Model

PD profile of tHcy upon IP injection of 50 mg/kg IAV8 (n = 6), IAV8-119D (n = 4), and heat-inactivated control (n = 3). Plotted as mean ± SEM. Statistical significance was analyzed by two-way ANOVA followed by Bonferroni's multiple comparison test. *P<0.05; **P<0.01; ***P<0.001; ****P<0.0001

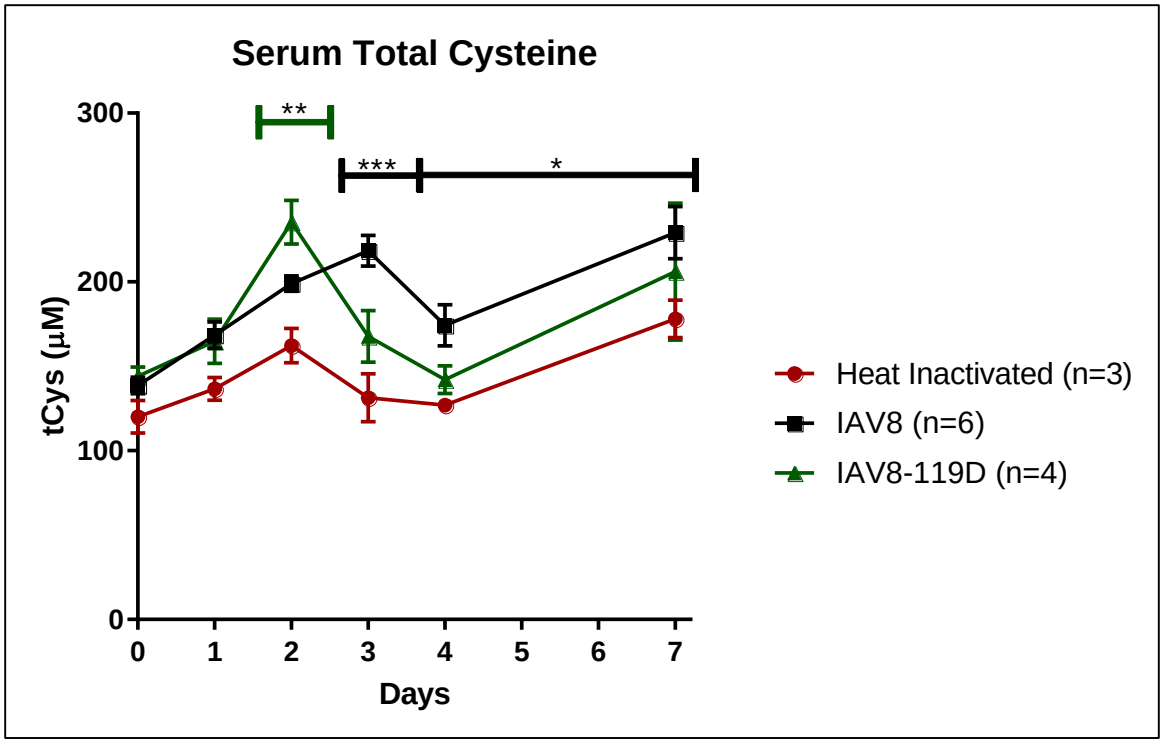


Figure 3.3 Serum Total Cysteine upon Single Injection in Diet-Induced Hyperhomocysteinemia Mouse Model

PD profile of tCys upon IP injection of 50 mg/kg IAV8 (n = 6), IAV8-119D (n = 4), and heat-inactivated control (n = 3). Plotted as mean ± SEM. Statistical significance was analyzed by two-way ANOVA followed by Bonferroni's multiple comparison test. *P<0.05; **P<0.01; ***P<0.001; ****P<0.0001

respectively. Analysis of the methionine levels for these studies showed that neither enzyme decreased the abnormal 3 mM methionine serum concentration to normal levels. In spite of the cysteine and cystine degrading activity of both enzymes, neither showed a statistically significant decrease in tCys levels, addressing one concern with using the hCGL scaffold as a therapeutic treatment for homocystinuria. This single dose study

showed the hCGL mutants effectively reduced tHcy levels without adversely affecting the mice, green lighting the next step in testing the treatment.

In order to assess efficacy over time and monitor for adverse side effects, IAV8 was tested in a multi-dose PD experiment. A second cohort of mice were given a 50 mg/kg IP injection of either IAV8 or heat-inactivated enzyme every 5 days for a month (Figure 3.4 and 3.5). Due to the required amount of blood for analysis, samples were only drawn at the time of every injection as well as 24 hours after every injection. Throughout the

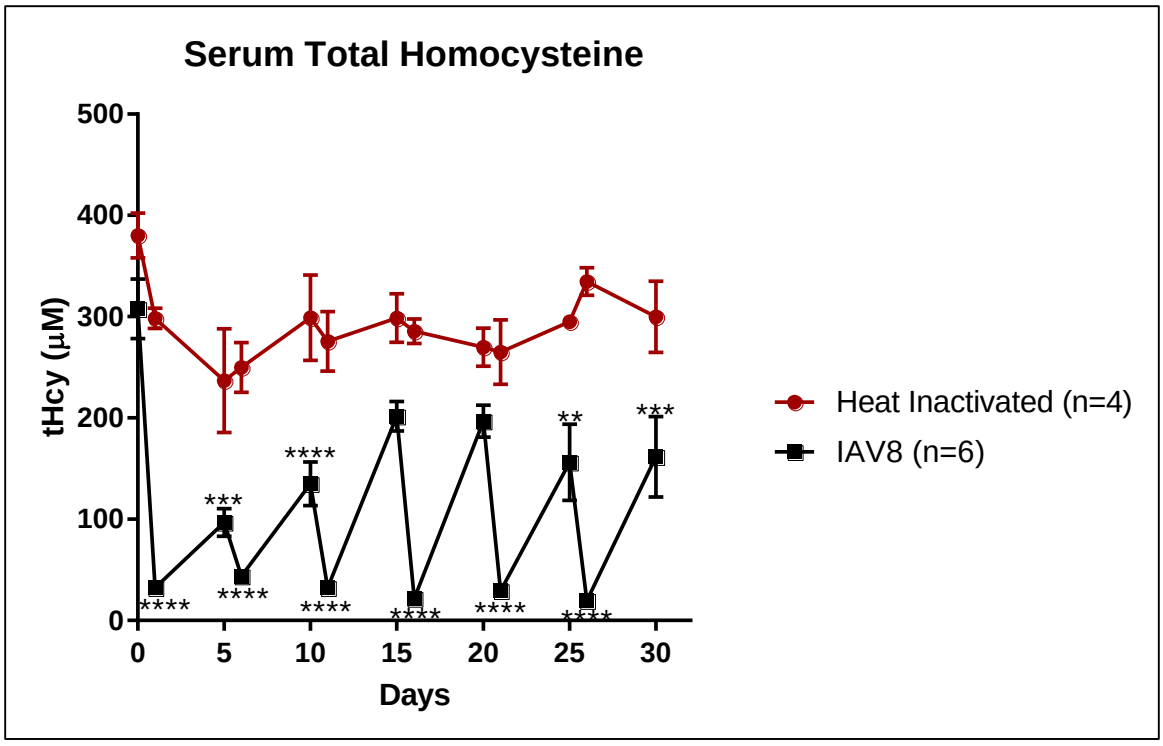


Figure 3.4 Serum Total Homocysteine upon Multiple Injections in Diet-Induced Hyperhomocysteinemia Mouse Model

PD profile of tHcy upon IP injection of 50 mg/kg IAV8 (n = 6) and heat-inactivated control (n = 4) every 5 days. Plotted as mean ± SEM. Statistical significance was analyzed by two-way ANOVA followed by Bonferroni's multiple comparison test. *P<0.05; **P<0.01; ***P<0.001; ****P<0.0001

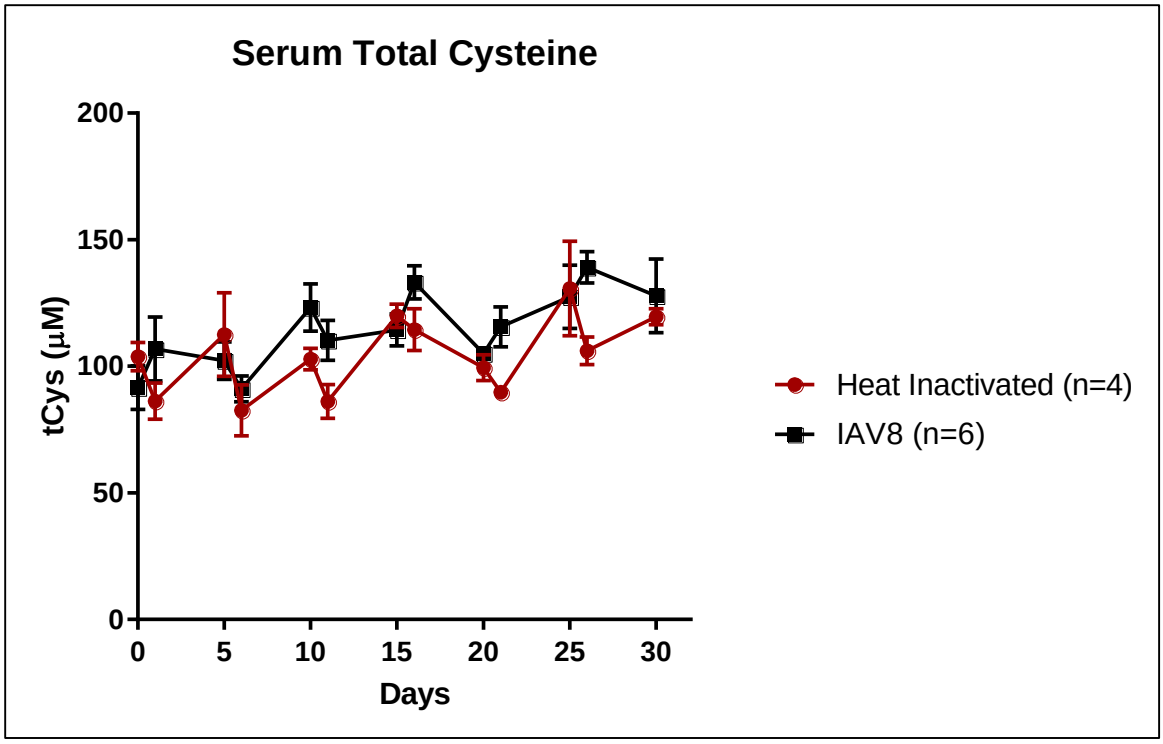


Figure 3.5 Serum Total Cysteine upon Multiple Injections in Diet-Induced Hyperhomocysteinemia Mouse Model

PD profile of tCys upon IP injection of 50 mg/kg IAV8 (n = 6) and heat-inactivated control (n = 4) every 5 days. Plotted as mean ± SEM. Statistical significance was analyzed by two-way ANOVA followed by Bonferroni's multiple comparison test. *P<0.05; **P<0.01; ***P<0.001; ****P<0.0001

study, neither the control nor treated mice showed any fluctuations in weight. Unsurprisingly, the tHcy PD profile followed the trend established by the single injection PD experiment. Although some variability was observed in the trough tHcy concentrations, the treatment repeatedly decreased the tHcy concentration below 50-25 μM. More substantial variability is seen in the peak tHcy concentrations, measured prior to each subsequent injection. The peak tHcy concentrations reach a maximum at days 15 and 20, which corresponds with the timeframe of peak antibody response to an antigen[90]. ELISAs performed on the day 0 and day 21/25 serum show an increase in antibody titer towards the enzyme but does not shown any neutralization (Figure 3.6). This suggests there was an

increase in antibody-mediated clearance, resulting in reduced efficacy during this time period, and likely explains the gradual increase in tHcy peak concentrations. The tCys PD

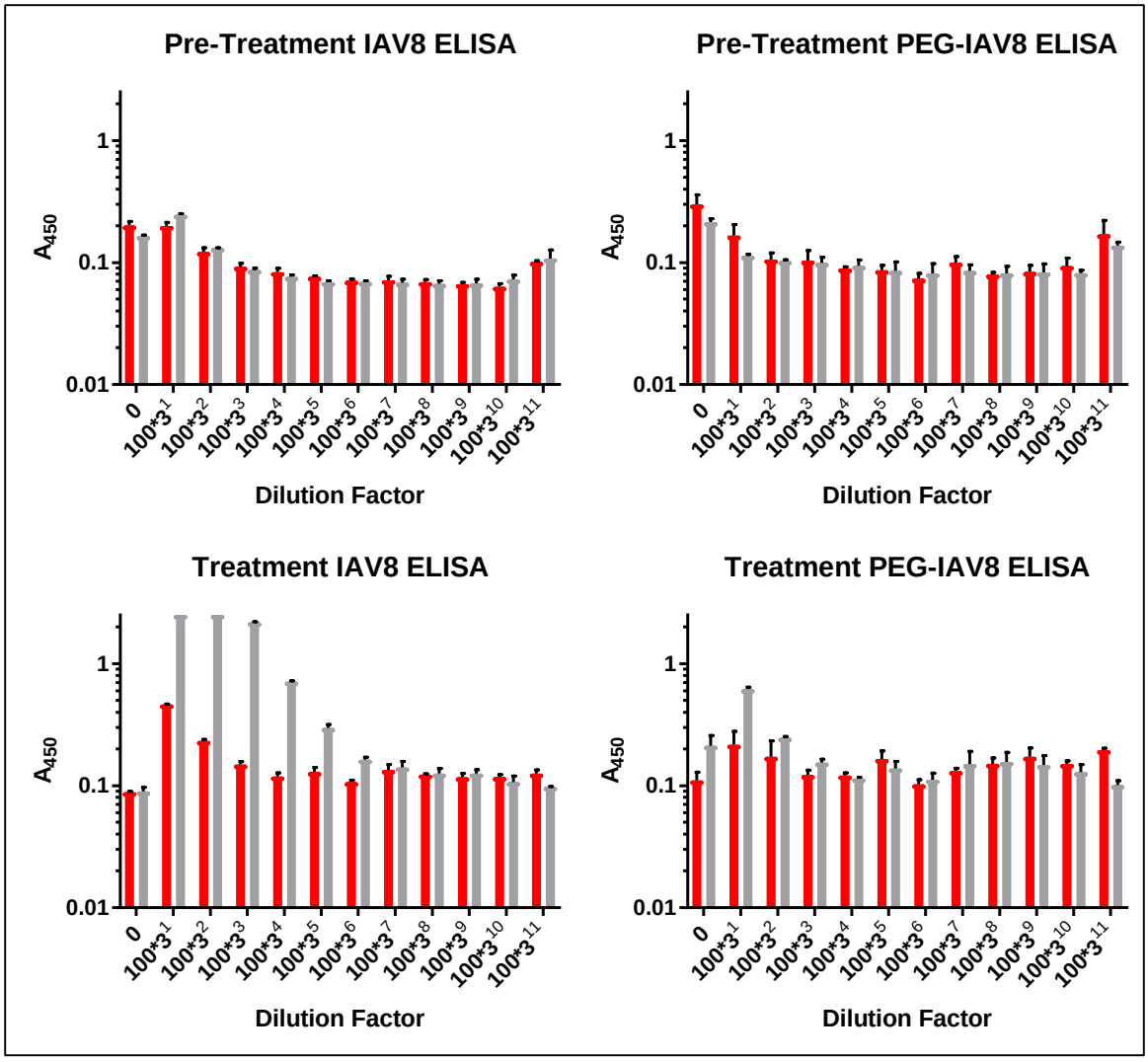


Figure 3.6 ELISA of Pre-Treatment and During Treatment of Antibody Response to PEGylated IAV8 and unPEGylated IAV8.

Red shows the ELISA of serum from untreated mice. Grey shows the ELISA of serum from treated mice. Pre-treatment ELISAs show minimal difference between the two groups across the various dilution factors. ELISAs of serum from treatment (day 21 and 25 for the treated and untreated mice respectively) show an antibody response to IAV8, although not to the PEGylated form. Plotted as mean ± SEM.

profile shows that, as mentioned in the single injection study, treatment does not lower tCys in spite of some intrinsic cysteine and cystine degradation activity.

*Administration of hCGL variants in a knock-out mouse model of homocystinuria completely prevents neonatal lethality and reverses liver abnormalities*

Based on the PD profiles for both candidates, we next tested the biological efficacy of the hCGL variants in the *cbs*$^{-/--}$ murine model of homocystinuria[91]. In this strain, the CBS gene is knocked out, resulting in a severe diseased state and causes a majority of the mice to die on or around day 25 post birth. The first experiment in this model was to confirm the PD data acquired in the diet-induced model. (Due to the young age, and therefore small size, of the mice, it was not possible to acquire multiple time points.) Pups were injected with either 20 mg/kg IP of IAV8, 25 mg/kg of IAV8-119D, or 20 mg/kg of

| Genotype | Treatment | Average tCys Concentration (μM) | Average tHCY Concentration (μM) | Average Met Concentration (μM) |
|---|---|---|---|---|
| *cbs*$^{+/+}$ | none | 230 ± 30 | 7 ± 1 | 54 ± 8 |
| *cbs*$^{+/-}$ | none | 260 ± 20 | 11 ± 2 | 54 ± 6 |
| *cbs*$^{-/-}$ | IAV8 20 mg/kg IP | 100 ± 10 | 90 ± 14 | 72 ± 12 |
| *cbs*$^{-/-}$ | Control 20 mg/kg IP | 86 ± 7 | 340 ± 50 | 830 ± 260 |
| *cbs*$^{-/-}$ | IAV8-119D 25 mg/kg IP | 95 ± 10 | 99 ± 13 | 88 ± 19 |

Table 3.2 Pharmacodynamic Response of Genetic Mouse Model of Homocystinuria to a Single Dose Treatment

The catalytic activity of hCGL and the engineered mutants used in this work. Presented as mean ± SEM.

heat-inactivated enzyme at day 22 and then sacrificed at day 23 for blood collection. This single dose PD study largely confirmed the trends established in the diet-induced single injection PD study as can be seen in Table 3.2; mainly that IAV8 lowers tHcy more than IAV8-119D, but, due to the increased dose of IAV8-119D, both lowered the tHcy concentration from ~375 μM to below 100 μM. Both treatments nearly normalized the methionine concentrations from the highly elevated levels. Additionally, neither injection lowered the tCys levels compared to the control knock-out mice, confirming that the treatment should not cause toxicity due to cysteine depletion nor methionine depletion.

Following the PD study, a multi-dose study was conducted to assess the biological efficacy of both candidate enzymes. The pups received their first dose at day 10, and then received an IP injection twice a week of either 20 mg/kg IP of IAV8, 25 mg/kg of IAV8-119D, or 20 mg/kg of heat-inactivated enzyme. The survival curves for all three treatments can be seen in Figure 3.7. As previously mentioned, and confirmed by this experiment, the majority of the untreated, *cbs*$^{-/-}$ pups died by day 25, whereas with the IAV8 injections, only 10% died by 25, and none died with the IAV8-119D injections. By day 50, the endpoint of treatment, none of the IAV8-119D treated mice had died, less than 40% of the IAV8 treated mice had died, and over 80% of the untreated mice had died. Some of the IAV8-119D treated mice were not sacrificed for histological examination, and all of them survived to day 100. Similar survival benefits have been observed with the aforementioned hCBS ERT in a similar murine model of homocystinuria, but required earlier intervention, beginning treatment at day 3 rather than day 10[41,42]. The increased survival of IAV8-119D treated mice over IAV8 treated mice is hypothesized to be due to the reduced methionine degrading activity of IAV8-119D. Although the single dose PD study of the *cbs*$^{-/-}$ pups did not show a significant difference in the methionine concentration, it is likely that over repeated dosing, IAV8 depleted methionine levels, more so than IAV8-119D. Although

depleting methionine can alleviate the accumulation of homocysteine, excessive methionine depletion can result in weight loss. Additionally, by depleting both homocysteine and methionine, it is possible that the intracellular equilibrium of s-adenosyl methionine and s-adenosyl homocysteine is disrupted, altering epigenetic regulation, further explaining the increased survival observed in the IAV8-119D treated mice. Regardless, both treatments improved survival without supplementation of cysteine, further confirming that cysteine generation is not required for successful treatment of this disease.

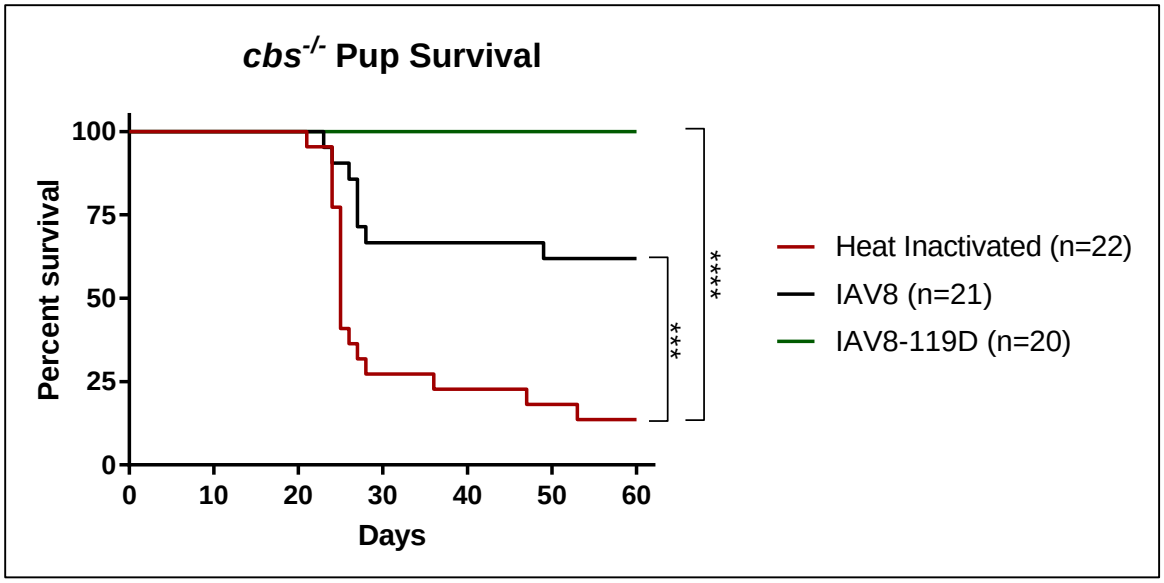


Figure 3.7 Kaplan-Meier Curves of Various hCGL Treatments in the Genetic Mouse Model of Homocystinuria

Mouse pups were enrolled at the age of 5 days, prior to genotyping. After confirming the genotype at day 10, knock-out pups continued the treatment. Treatment was twice weekly at 20 mg/kg for the heat inactivated, 20 mg/kg for the IAV8 treatment, and 25 mg/kg for the IAV8-119D treatment. Treatment was stopped at day 50, and survival monitored for 10 days without treatment. Statistical significance was analyzed by a log-rank (Mantel-Cox) test. *P<0.05; **P<0.01; ***P<0.001; ****P<0.0001

In addition to the binary metric of survival, livers of the afflicted mice were harvested at various time points in order to assess the pathological impact of the disease and potential amelioration brought about by treatment. Representative images of the liver sections can be seen in Figure 3.8. At day 23, the heterozygous mice showed no pathological abnormalities, whereas the untreated, homozygous knock-out mice showed broad confluent areas of hepatocyte necrosis as well as microvesicular steatosis in many of the viable hepatocytes. Mice treated with IAV8 displayed decreased necrosis, showing only occasional foci of hepatocyte necrosis and microvesicular steatosis. Also, livers assessed at day 60 from the knock-out mice treated with IAV8 showed only scattered, enlarged

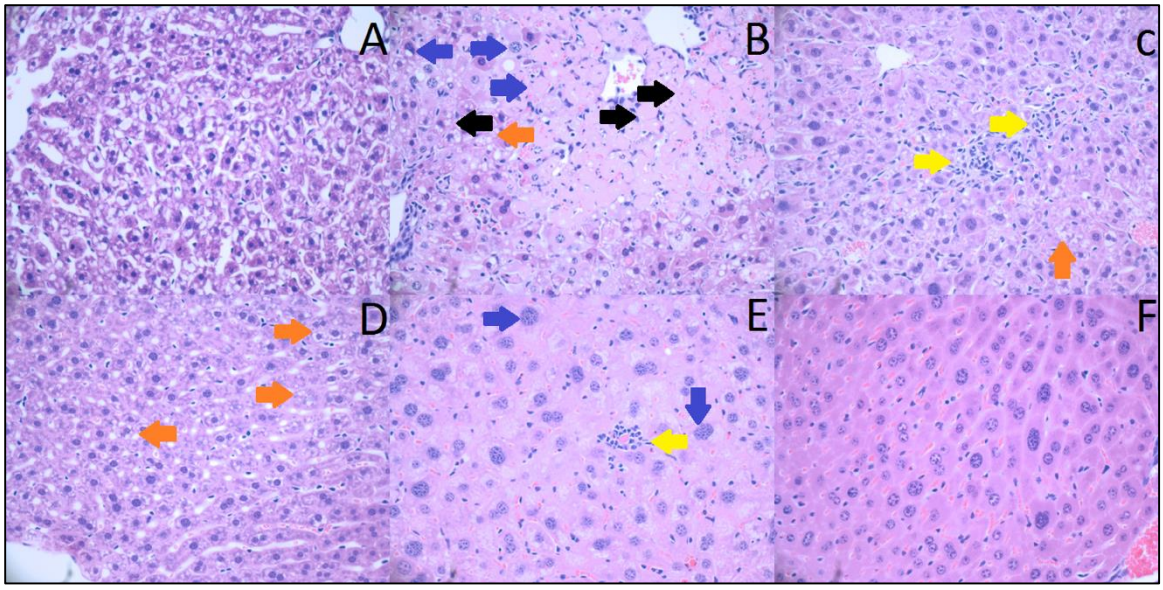


Figure 3.8 Treated and Untreated Liver Histology of the Genetic Mouse Model of Homocystinuria

40x liver sections of a) day 23 untreated *cbs*$^{-/+}$ pup, b) day 23 untreated *cbs*$^{-/-}$ pup, c) day 23 IAV8 treated *cbs*$^{-/-}$ pup, d) day 60 untreated *cbs*$^{+/+}$ pup, e) day 60 IAV8 treated *cbs*$^{-/-}$ pup, and f) day 60 IAV8-119D treated *cbs*$^{-/-}$ pup. Yellow arrows denote focal necrosis, blue arrows denote enlarge nucleus, black arrows denote large necrotic areas, and orange arrows denote microvesicular steatosis. As can be seen, treatment drastically reduces the disease progression in the liver.

nuclei and some focal areas of hepatocyte necrosis a significant improvement from the untreated mice and treated mice seen at day 23. Most importantly, *cbs*$^{-/-}$ mice treated with IAV8-119D show no pathological abnormalities at day 60, further confirming the superiority of this variant to both IAV8 and existing treatments for homocystinuria.

IAV8-119D accomplishes the classic homocystinuria treatment goal of lowering tHcy through an orthogonal mechanism while surpassing the efficacy seen in current and proposed treatments. IAV8-119D, despite its cysteine and cystine degrading activities, did not deplete tCys levels, alleviating a concern for a treatment based on the hCGL scaffold. Treatment completely prevented the neonatal lethality observed in the *cbs*$^{-/-}$ strain, a result which has not been reported by any other treatment to date. Lastly, this therapy prevented the liver pathological abnormalities typically seen in the murine model. In summary, IAV8-119D achieved the desired pharmacodynamic properties and biological efficacy without adverse side-effects, representing a promising and novel therapeutic option for homocystinuria patients.

*Future Directions*

This work demonstrated the efficacy of an enzyme therapeutic for homocystinuria based on the hCGL scaffold. Additionally, due to the difference in mechanism of action between hCGL treatment and the hCBS treatment, this work suggests that tHcy depletion alone can treat homocystinuria. Moving forward, future work should investigate auxiliary metabolites, such as s-adenosyl methionine, s-adenosyl homocysteine, and homocysteine thiolactone, to determine how their levels respond upon treatment. Additionally, future work could assess the response of metabolite levels intracellularly for the affected tissue. In both cases a comparison between IAV8 and IAV8-119D treatment could further explain the differences in survival observed in the knock-out mouse model. Furthermore, comparison to the hCBS treatment could determine if cysteine supplementation has any

effects not seen in the hCGL treatment. Beyond that, future work will need to include human patients or samples, in large part due to the differences between the murine models and the human disease.

**ACKNOWLEDGEMENTS**

I would like to acknowledge and thank Dr. Makiko Watanabe for her invaluable contributions to the projects and the work she did with the mice, Candice Lamb and Yuri Tanno for their contributions in purifying enzyme for the mouse studies, Dr. Wei-Cheng Lu for his work on this project, Christopher Daige and Danlee Enzler for preparing the tissue sections for histology, Precision Pathology for the histological analysis, and Dr. Scott Rowlinson and Aeglea Biotherapeutics for their contributions and assistance in this work.

# Chapter 4: High Yield Orthogonal Synthesis and Purification System for Peptides based on Human Asparaginase-Like Protein 1

## 4.1 Abstract

Therapeutic peptides are a growing field within modern medicine, with over 60 FDA-approved peptide drug, and many more in active trials[92]. Unfortunately, synthesis technology for peptides, especially at the laboratory-scale, has not kept pace with the growing needs of the field. Traditional synthesis techniques struggle with long peptides as well as specific difficult sequences, limiting the sequence space that research can survey[93]. More modern techniques have fused chemical techniques with biological tools, or even exclusively utilized biological tools for the synthesis of peptides[94,95]. Unfortunately, these techniques require a significantly longer time for synthesis than purely chemical methods. Additionally, many of the biological techniques have lower yields. This work developed an alternative synthesis and purification system for peptides based on human Asparaginase-like protein 1 (hASRGL1), a member of the N-terminal nucleophile superfamily. As such the active form of hASRGL1 is generated by an intramolecular cleavage step. Previous work created a circularly permutated variant, cp-hASRGL1, circumventing the need for this N-terminal hydrolase reaction[50]. By fusing a peptide of interest to the new N-terminal of the cp-hASRGL1, the peptide can be purified based on the affinity tag associated with cp-hASRGL1 and then the peptide can be isolated by a glycine-induced cleavage reaction. This system leaves a minimal scar, a C-terminal glycine, on the peptide of interest. Additionally, of the tested peptides, the theoretical yield ranges between ~5-10 mg/L, and, in these experiments, ~33% recovery of this theoretical yield was obtained. This system represents a novel, high yield, orthogonal synthesis and purification system for peptides and could improve the accessibility of peptides that are difficult to synthesize by conventional methods.

## 4.2 INTRODUCTION

Therapeutic peptide and proteins are of increasing medical importance. With over 60 FDA-approved peptide drugs, and many more in active trials, the need for improved synthesis methods both on a lab scale and an industrial scale is growing[92]. The predominant technology for peptide synthesis, solid phase peptide synthesis (SPPS), utilizes iterative coupling steps to grow a peptide chain[93]. In brief, the C-terminal residue is attached via its carboxyl acid group to an immobilized resin. The protecting group, which was originally attached to the N-terminal amine, is removed and the next amino acid is coupled to the exposed amine. This process is repeated until the desired sequence is acquired. Unfortunately, at each step, the protecting group on the chain can be prematurely removed or improperly removed, resulting in a sequence that is either too long or too short, respectively, in addition to containing an incorrect sequence. Additionally, due to reactive side chains of various amino acids, orthogonal protecting groups must be added to those amino acids to prevent branching peptides. As the peptide length increases, the complexity and cost of synthesis increase, while product purity decreases, limiting the technology to relatively short peptide chain lengths.

Alternatively, newer research has unveiled enzymes capable of ligating peptide chains, providing novel advances upon SPPS. In particular, chemo-enzymatic peptide synthesis (CEPS), couples the peptide creation of SPPS with enzymatic ligation via sortase or omniligase-1, allowing for the creation of longer peptide chains with higher purity[94,96]. Another advance in peptide synthesis completely avoids chemical synthesis, instead utilizing cell-based synthesis. The peptide of interest is expressed as a fusion with a protein splicing element and an affinity tag[95,97,98]. The fusion construct is purified via affinity chromatography, after which the peptide is isolated by inducing splicing and therefore released from the affinity column. This technology, unlike the previous two, has no

restrictions on the size of the peptide. Unfortunately, because this technology relies on cells to synthesize the peptide of interest rather than chemical processes, yield is often a limiting factor. Namely, the IMPACT (Intein Mediated Purification with an Affinity Chitin-binding Tag) system typical yields less than 5 mg/L for full-length proteins, and significantly less for peptides[99].

This works presents a novel, orthogonal biologic peptide synthesis and purification technology based on human asparaginase-like protein 1 (hASRGL1). hASRGL1 is an N-terminal hydrolase that upon activation has β-aspartyl peptidase and asparaginase activity[49]. hASRGL1, through an intramolecular mechanism, breaks the scissile bond between Gly167 and Thr168, creating two distinct subunits, the α-subunit which contains Asn2 to Gly167, and the β-subunit, which contains Thr168 to Pro308[50]. Without intervention, this self-cleavage reaction proceeds very slowly, only reaching a maximum of approximately 50% cleavage after 3 days at 37°C[100]. This reaction can be accelerated via the addition of glycine, achieving nearly completely cleavage within 48 hours at high concentrations of glycine[100]. In tandem to this catalysis discover, other work created a circularly permutated version of hASRGL1, referred to as cp-hASRGL1[50]. The circular permutation removed the requirement for intramolecular autoprocessing and allows expression and purification of fully active enzyme. This circular permutation links the C-terminal of the β-subunit to the N-terminal of the α-subunit via a 12 amino acid linker, thus removing removed the scissile bond between Gly167 and Thr168. A schematic of the cleavage reaction and the circularly permutated variant can be seen in Figure 4.1. Subsequent work revealed that by fusing an additional amino acid sequence to the N-terminal of cp-hASRGL1, and thus recreating the scissile peptide bond, cp-hASRGL1 retained its N-terminal hydrolase activity, cleaving the additional amino acid sequence. Following this discovery a self-cleaving, single-column purification system was devised

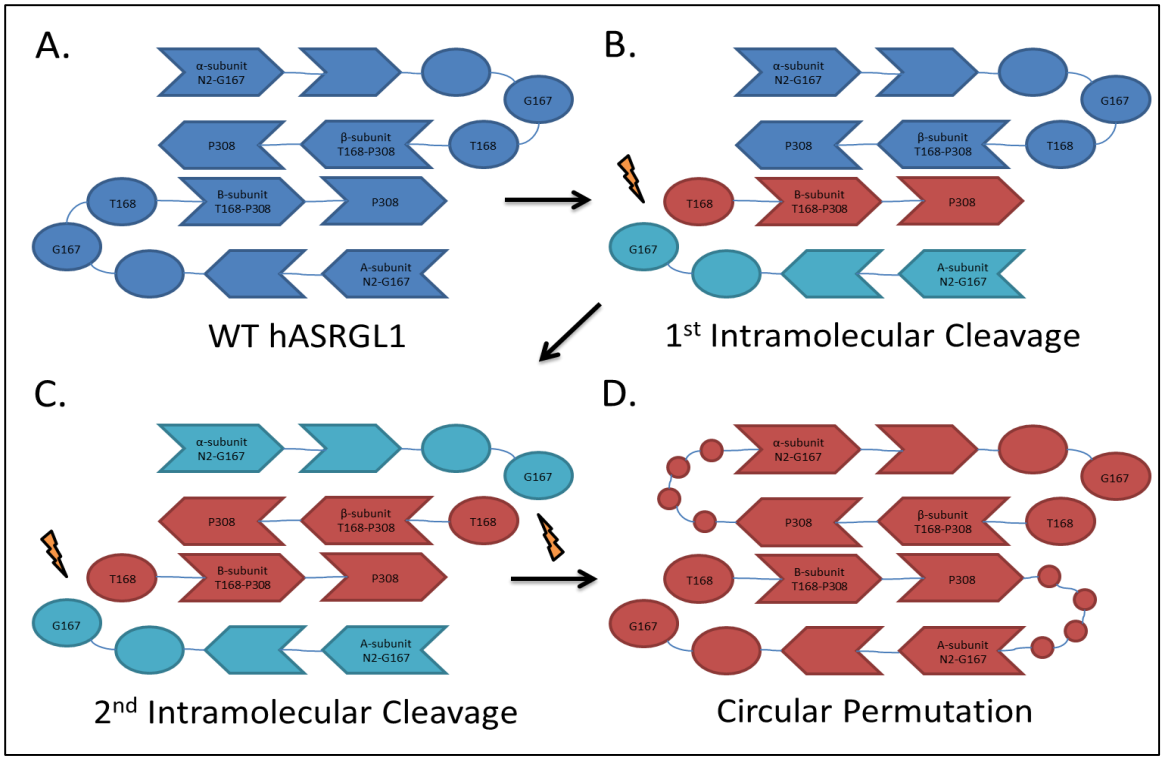


Figure 4.1 Cleavage Reaction of hASRGL1 and the Circularly Permutated Variant, cp-hASRGL1

A) uncleaved naïve form of hASRGL1, B) first cleavage event, C) second cleavage event, D) circularly permutated variant of hASRGL1, cp-hASRGL1. Adapted from: Li et al., 2012, ACS Chemical Biology

and tested. As fusions to the new N-terminal of cp-hASRGL1, peptides are expressed, bound to a Ni-NTA column, and isolated via glycine-induced cleavage. Due to the naturally high yield of hASRGL1 in *E. coli*, up to 10 mgs of peptide can be recovered per liter of induced *E. coli* culture. For a given a peptide sequence, the required plasmid can be generated and product subsequently created in less than 2 weeks. Due to the mechanism of the cleavage reaction, the peptide sequence has no apparent restrictions other than a C-terminal glycine scar. Overall, this technology provides a quick, efficient, high-yield synthesis technique for laboratory use.

## 4.2 MATERIALS AND METHODS

*Materials:* Oligonucleotides were purchased from Integrated DNA Technologies. Restriction enzymes, Taq polymerase, Phusion high fidelity DNA polymerase, T4 DNA ligase, Gibson master mix, corresponding buffers, and dNTPs were purchased from New England Biolabs. Difco 2xYT growth media, LB growth media, and TB growth media were all purchased from Becton Dickinson. Nickel-nitroacetic acid (Ni-NTA) resin and mini-prep kits were purchased from Qiagen. Isopropyl β-D-1-thiogalactopyranoside (IPTG), kanamycin, glycine, and all other unmentioned chemicals were all purchased from Sigma-Aldrich.

*Gibson Assembly & Cloning:* gBlocks for the various peptides and proteins to be expressed were ordered from Integrated DNA Technologies. gBlocks were then amplified using gene specific primers that also contained homology with the pET28a-cp-hASRGL1 plasmid. Simultaneously, the pET28a-cp-hASRGL1 plasmid was amplified via PCR in 2 separate pieces. Then, the three PCR products, two plasmid and one gene of interest, were purified via gel extraction. The gel extracted DNA was mixed with the Gibson assembly master mix and incubated at 50°C for 30 minutes. The mixture was desalted and then transformed into DH10β cells and plated on an antibiotic selective agar plate. Individual colonies were picked and identify confirmed via Sanger sequencing. Confirmed plasmids were re-transformed into either BL21(DE3) cells and again plated on antibiotic selective agar plates. Colonies were picked and used to inoculate overnight cultures, which were then frozen for subsequent usage.

*Enzyme Expression, Purification, & Cleavage*: The fusion constructs were cloned into the pET22a vector using a combination of PCR amplification, gel extraction, and Gibson assembly. *E. coli* BL21(DE3) cells harboring these plasmids were grown in terrific broth (TB) media at 37°C. Upon reaching mid-log phase ($OD_{600}$ of ~0.8) isopropyl-β-D-1-

thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM. The induced culture was further incubated for ~16 hours at 25°C, after which cells were collected by centrifugation and re-suspended in immobilized metal affinity chromatography (IMAC) buffer (10 mM $NaH_2PO_4$/30 mM imidazole/300 mM NaCl, pH 8) with a protease inhibitor tablet. The re-suspended cells were lysed via a french pressure cell and centrifuged. The supernatant was applied to a Ni-NTA column, washed with 20 column volumes of IMAC buffer. For on-column cleavage, ~5 column volumes of IMAC cleavage buffer (10 mM $NaPO_4$/30 mM imidazole/300 mM NaCl/50 mM glycine, pH 8) was added and incubated at the desired temperature for the desired time, after which it was eluted. For off-column cleavage, 5 column volumes of IMAC elution buffer (50 mM $NaPO_4$/250 mM imidazole/300 mM NaCl, pH 8) was added and eluted. The elution was then mixed with an equal volume of cleavage buffer (10 mM $NaPO_4$/30 mM imidazole/300 mM NaCl/500 mM glycine, pH 8) and incubated at the desired temperature for the desired time. Peptides were then purified from the cp-hASRGL1 via 10,000 MWCO centrifugal filter devices. Extinction coefficients for protein concentration determination were calculated based on amino acid sequence. Purity and cleavage efficiency were assessed by SDS-PAGE coupled with densitometry. Peptide sequences were confirmed via MALDI-MS. Peptide yield was further confirmed via a BCA assay.

## 4.3 Results and Discussion

*BFP-cp-hASRGL1 Fusion System Validation and Optimization*

This purification system was first tested by fusing a protein of interest, in this case blue fluorescent protein (BFP), to the N-terminal of the cp-hASRGL1 variant. The protein sequence around the scissile bond of hASRGL1, cp-hASRGL1, and the BFP-cp-hASRGL1 can be seen in Figure 4.2. In order to assess the natural expression and solubility of this

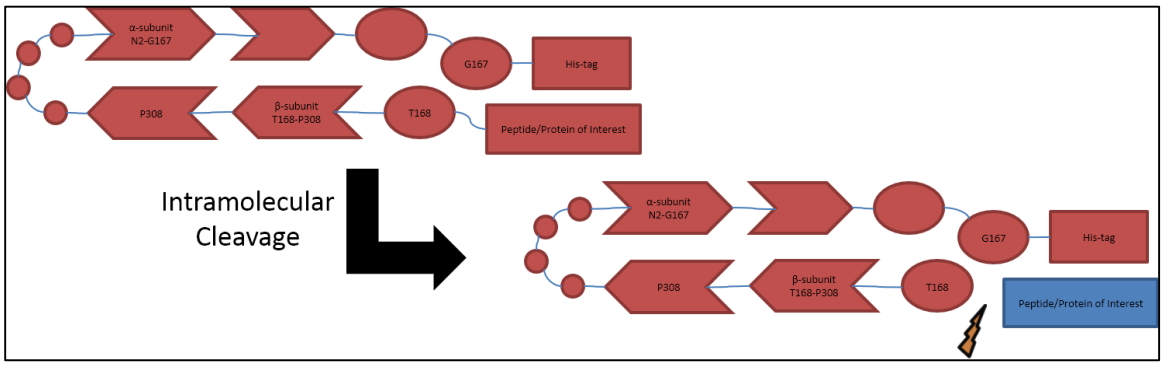


Figure 4.2 Schematic of cp-hASRGL1 Fusion and Intramolecular Cleavage

An N-terminal fusion of a peptide of interest to cp-hASRGL1. Cleavage, and therefore release of the peptide, is mediated by glycine addition.

system, high expressing and solubility-enhancing proteins, such as maltose binding protein (MBP) and green fluorescent protein (GFP), were not used for initial tests as they might artificially improve these parameters. A C-terminal His-tag was added to the cp-hASRGL1 protein to allow for affinity purification of the fusion construct. In brief, the fusion construct with induced via IPTG during mid-exponential phase. After induction, the cells were harvested and lysed, after which the lysate was bound to a Ni-NTA affinity column. In order to perform an on-column cleavage, glycine was added to the column, inducing cleavage, and thus liberating the protein of interest. Because glycine can act as a competitive ligand for Ni-NTA resin and therefore cause premature release of cp-

hASRGL1, an off-column cleavage method was also utilize for increased yield. To do so, the fusion construct was eluted via addition of imidazole, a competitive ligand for Ni-NTA. After elution, glycine was added to accelerate cleavage. Finally, the peptide of interest was isolated from the carrier/tag protein, cp-hASRGL by quickly buffer exchanging the glycine away, and passing the BFP and cp-hASRGL1 mixture over a fresh Ni-NTA column, rebinding cp-hASRGL1 and any uncleaved fusion proteins. The flow-through, a high purity BFP solution, was captured. After which, the purified BFP and the cleavage efficacy was characterized. Alternatively, although not applicable for the BFP purification, size exclusion centrifugal units were used for some of the other experiments, namely when expressing peptides rather than proteins. By using a molecular weight cutoff of ~10 kDa, the cp-hASRGL1 and any uncleaved fusion were trapped in the retentate, while the flowthrough was a highly pure solution of the peptide.

To assess the kinetics of this system, the BFP-cp-hASRGL1 fusion was expressed and the cleavage efficiency was measured, off-column at different temperatures over time (Figure 4.3). In each case, the fusion protein was expressed and purified using the aforementioned Ni-NTA column, and then eluted for off-column cleavage. The fusion was incubated with 250 mM glycine and incubated at various temperatures for various times, to determine the temperature-time profiles. The cleavage reaction was observed to be relatively slow at 4°C, cleaving less than 5% within 4 hours, and only 12% at 24 hours. Alternatively, at 25°C, the cleavage reaction showed decent cleavage within 4 hours, and more than 50% by 24 hours. Importantly, at 37°C, 75% of the BFP was cleaved within the first 4 hours, with less than 5% remaining uncleaved by 24 hours. Due to the slower cleavage rate at 4°C, this purification method may not be applicable for temperature sensitive peptides and proteins that cannot be incubated at 25°C or 37°C.

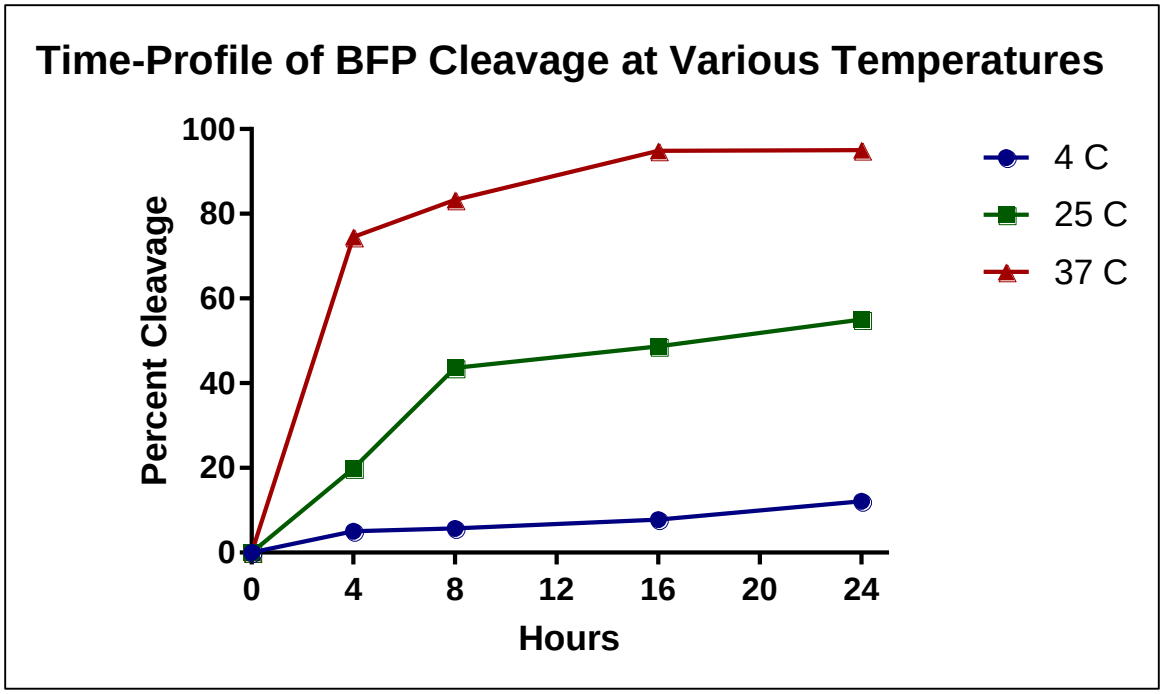


Figure 4.3 Temperature-Time Profile of BFP-cp-hASRGL1 Fusion Cleavage

The temperature-time profile of BFP-cp-hASRGL1 fusion cleavage efficiency at 4°C (blue circles), 25°C (green squares), and 37°C (red triangles).

In addition to testing the cleavage kinetics, the BFP fusion was used to assess the impact of the P1' residues on cleavage. Phylogenetic analysis of hASRGL1 shows the Gly167 and Thr168 residues are highly conserved, suggesting these specific amino acids are crucial to the N-terminal hydrolyase chemistry. The original BFP fusion construct was designed such that the final residue of the BFP-cp-hASRGL1 linker was a glycine, mimicking the native hASRGL1 sequence. This residue was mutated to the other 19 amino acid in order to determine how they affect the cleavage reaction. For all mutations, other than the original glycine, cleavage was not detected, confirming the glycine residue in the P1' position is essential. Additionally, other proteins were tested in this cleavage system, such as the pfu-sso7 fusion DNA polymerase[101]. It was found that certain proteins were incapable of being cleaved, which, based on experiments discussed below, was attributed

to steric clashes caused by the tertiary and quaternary structure of the fusion. In spite of the difficulty translating this system to other proteins and the strict requirement for a P1' glycine, the purification system still worked remarkable well with high yield, and warranted further investigation.

*Extension of the System to Peptides and Further Validation*

To further assess the robustness of this purification scheme, multiple peptide fusions were expressed and characterized, the names of which, along with their sequences, are shown in Table 4.1. Each peptide was chosen due to their high price of purchase and

| Peptide | Sequence |
|---|---|
| Nesiritide-like | MSPKMVQGSGCFGRKMDRISSSSGLGCKVLRRHG |
| Teriparatide-like | MSVSEIQLMHNLGKHLNSMERVEWLRKKLQDVHNFG |
| Zadaxin-like | MSDAAVDTSSEITTKDLKEKKEVVEEAERG |

Table 4.1 Purified Peptides and their Sequences

A table of the purified peptides and their sequences in the context of the cp-hASRGL1 fusion.

their FDA approval as therapeutic peptides. Each fusion was expressed and purified, and the off-column cleavage efficiency was assessed at 37°C in the presence of 250 mM glycine. Preliminary cleavage tests showed that the Teriparatide-like and Zadaxin-like peptides achieved over 75% cleavage within 24 hours, whereas the Nesiritide-like peptide achieved less than 20% cleavage. Interestingly, the Nesiritide-like peptide contains cysteine residues, and thus a steric issue, as seen with the DNA polymerase, was suspected, potentially explaining the poor cleavage efficiency. The purification was repeated with the addition of 10 mM DTT, which resulted in improved cleavage efficiency to the level of

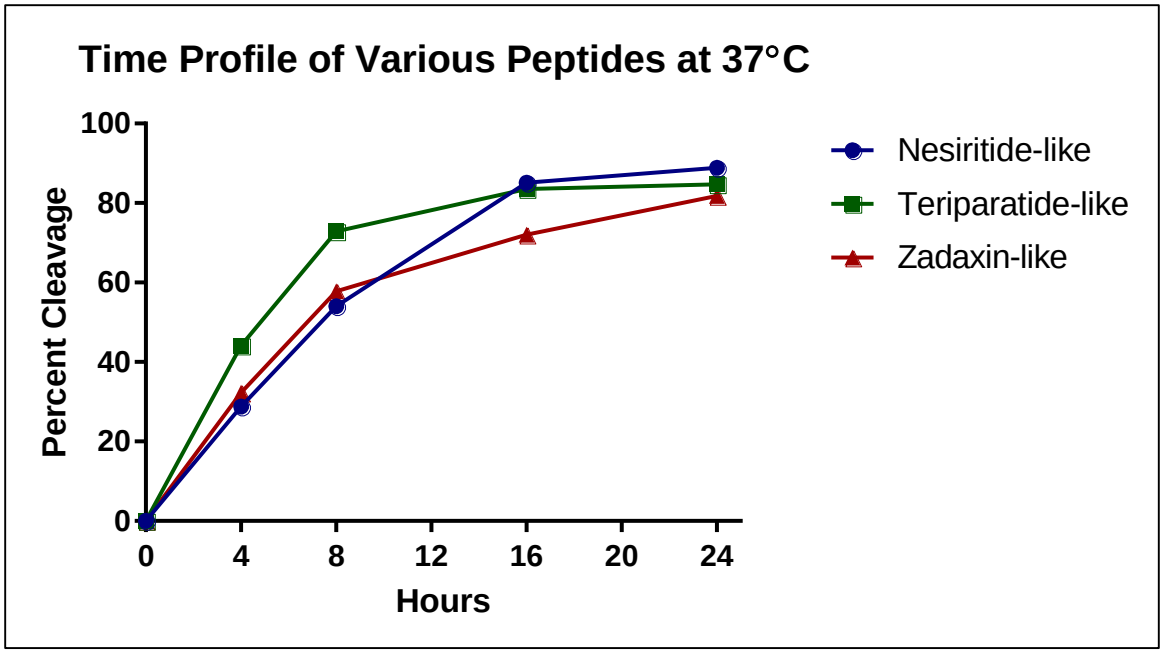


Figure 4.4 Time Profile of Peptide-cp-hASRGL1 Fusion Cleavage Efficiency

The time profile of peptide-cp-hASRGL1 fusion cleavage efficiency of the Nesiritide-like peptide (blue circles), the Teriparatide-like peptide (green squares), and the Zadaxin-like peptide (red triangles).

the other peptide constructs. The time profile of the cleavage can be seen in Figure 4.4. Although cleavage efficiency is important, the yield of the fusion construct also dictates the final purified quantity. For each peptide-fusion, a yield of ~75-100 mg/L of fusion protein was achieved, representing a theoretical yield of ~5-10 mg/L of peptide, depending on the specific peptide-fusion, its yield, and the relative size of the peptide compared to cp-hASRGL1. Using the off-column cleavage and purification scheme, 3.2 mg/L of the Teriparatide-like peptide was recovered at the final step. Given the ~90 mg/L yield of the Teriparaitde-like-cp-hASRGL1 fusion and the relative size of the peptide to fusion, this recovery represents ~32% of the theoretical yield.

In addition to assessing the yield and cleavage efficiency, the sequences of the purified peptides were confirmed by matrix assisted later desorption/ionization mass

spectrometry (MALDI-MS). The MALDI-MS spectra for both the Teriparatide-like peptide and the Zadaxin-like peptide can be seen in Figures 4.5 and 4.6 respectively. Given the sequences mentioned in Table 4.1, a molecular weight of 4306 Da was expected for the Teriparatide-like peptide. Without the N-terminal methionine, which is often cleaved off in *E. coli*, the expected size was 4175 Da. Similarly, for the Zadaxin-like peptide the expected sizes were 3297 Da and 3165 Da respectively. The variant without the N-terminal methionine can clearly be seen in both the MALDI-MS spectra within a few Daltons. Additionally, a second peak, often larger than the expected peak, can be seen in both spectra at ~+43. This second peak represents an acetonitrile adduct as the peptides were dissolved in acetonitrile for the MALDI-MS experiment. Thus the MALDI-MS analyses confirmed the sequences of the purified peptides and the absence of contaminants.

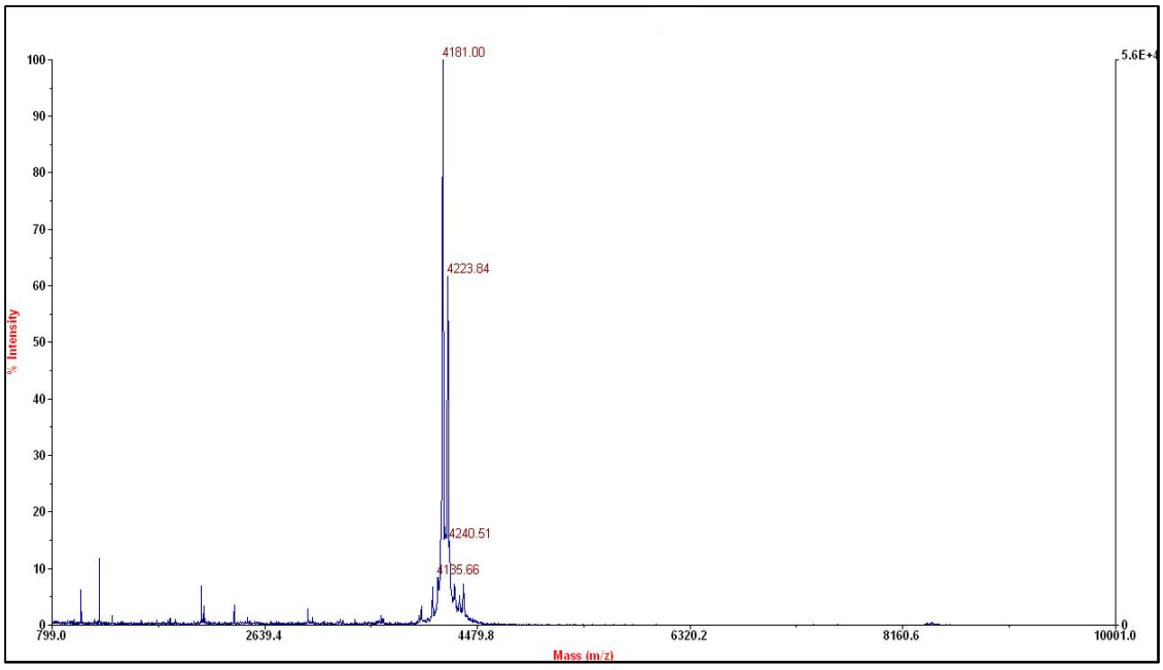


Figure 4.5 MALDI Spectrum of Purified Teriparatide-Like Peptide

MALDI spectrum of the purified Teriparatide-like peptide and its acetonitrile adduct.

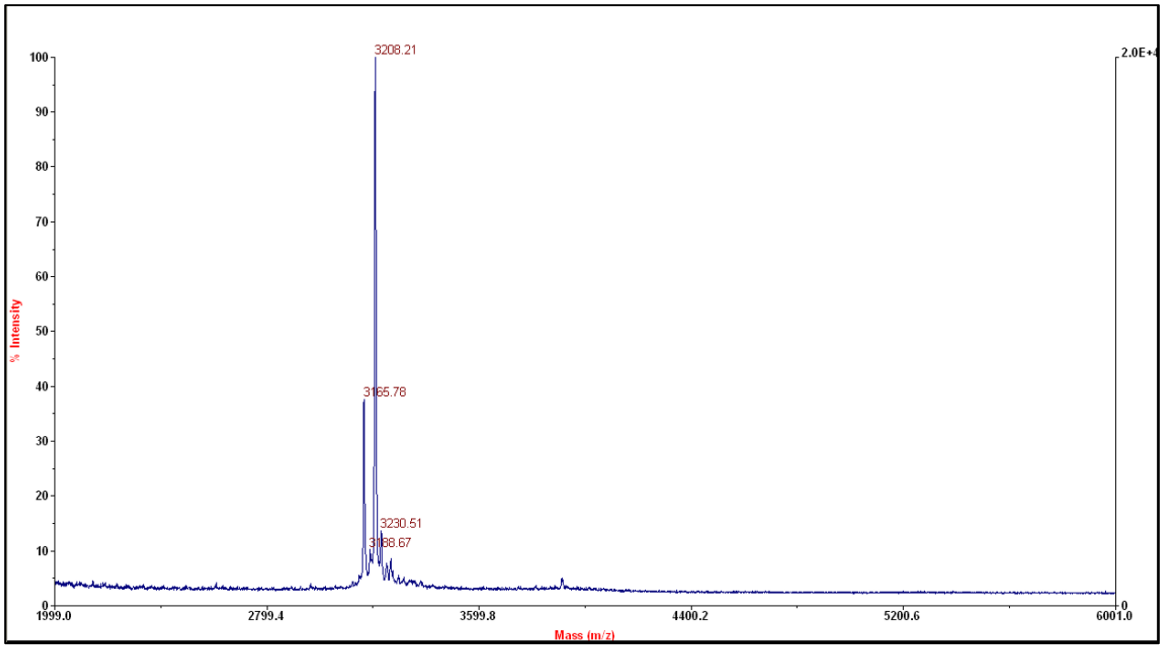


Figure 4.6 MALDI Spectrum of Purified Zadaxin-Like Peptide

MALDI spectrum of the purified Zadaxin-like peptide and its acetonitrile adduct.

*Conclusion and Future Directions*

This system represents a novel, high yield orthogonal method for the expression and purification of peptides. The system has been streamlined to allow for easy Gibson assembly. Based on the presented experiments, the system robustly handles a variety of peptide sequences without hindering cleavage efficiency or yield. Although the system requires a longer time for synthesis compared to chemical techniques, it is comparable to other biological methods, such as intein fusions, with a marked improvement in yield. The system does leave a C-terminal glycine scar, but does not impose any other sequence restrictions. This system improves the toolset for peptide synthesis, thereby enabling easily accessible peptides for future therapeutic peptide research.

Moving forward, future work for this project could focus on point mutations to hASRGL1 to improve the yield of the fusion proteins. Additional work could test different protein fusions with hASRGL1 to determine what sequences are capable of being cleaved and which sequences sterically hinder the cleavage reaction. The 6xHis affinity tag could be replaced to test on-column glycine mediated cleavage, as the 6xHis affinity tag dissociates at high glycine concentrations. Lastly, additional peptide sequences could be tested to confirm the versatility of the system.

**ACKNOWLEDGEMENTS**

I would like to acknowledge Dr. Maria Person and the UT Austin Proteomic Facility for their training and help with the MALDI runs.

# Chapter 5: Major Findings and Future Recommendations

## 5.1 MAJOR FINDINGS

### 5.1.1. Chapter 2: Breaking the metabolic regulation of the branched chain amino acid biosynthesis pathway in order to engineer a protein therapeutic for homocystinuria

This work had two major aims: 1) to engineer a variant of human cystathionine-γ-lyase (hCGL) with improved activity towards homocysteine and homocystine, and 2) to re-engineer the BL21(DE3)*ΔilvAΔmetA* strain and its associated genetic selection to aid in the identification of improved hCGL mutants[57]. Although the libraries that were assessed by this method did not result in improved variants, this work elucidated many aspects of the branched chain amino acid biosynthesis pathway that will improve future research, in addition to laying the groundwork for future engineering of an improved homocyst(e)inase. A multitude of libraries were constructed based on rational design that can be used for future selection and screens. By reintroducing *ilvGM*¸ the genes encoding acetohydroxy acid synthase isozyme II, along with the supplementation of valine and leucine, this work overcame the α-ketobutyrate toxicity observed in *E. coli*. Additionally, in doing so, this work provided strong evidence of the induced valine and leucine auxotrophy induced by excess α-ketobutyrate. Lastly, applying the selection outside the scope of this manuscript to a serine/threonine deaminase has shown much promise, successfully identifying variants with improved kinetics.

### 5.1.2. Chapter 3: Prevention of homocystinuria disease progression in a neonatal lethal murine model through systemic homocysteine depletion with a therapeutic enzyme

This work focused on testing and validating two similar homocystinuria treatments based on the hCGL scaffold. As mentioned previously, recent work has shown the efficacy

of treatment in murine models utilizing an hCBS mutant[41]. This work, first assessed the pharmacodynamic efficacy of two hCGL mutants in a diet-induced murine model of hyperhomocysteinemia. Both mutants depleted homocysteine levels without adversely affecting the mice and without depleting cysteine levels. In a month-long study with repeated treatment, homocysteine levels were effectively controlled through the study, again without any adverse effects observed. In a neonatal lethal genetic murine model of homocystinuria, both treatments improved survival, with complete survival using the variant with higher selectivity for homocysteine. Histology of the livers showed significant prevention of disease progression in the treated mice. In all, this work demonstrated a novel therapeutic option for homocystinuria, effectively depleting homocysteine levels, and preventing the neonatal lethality in a severe model of the disease.

### 5.1.3. Chapter 4: High Yield Orthogonal Synthesis and Purification System for Peptides based on Human Asparaginase-Like Protein 1

This work successful designed and validated a high yield, orthogonal synthesis and purification system for peptides based on human Asparaginase-like protein 1 (hASRGL1). The work here demonstrated the versatility of the system and its ability to handle different peptide lengths and sequences without adverse effects to yield, purity, or cleavage efficiency. Additionally, unlike some of the alternative synthesis techniques, this system has few restrictions on the peptide sequence, only requiring a C-terminal glycine. For the purified peptides, a high yield of ~10 mg/L was achieved, a huge improvement over existing biological synthesis techniques. Due to its high yield, ease of use, and minimal sequence restrictions, this technology will hopefully make peptides more readily accessible for future research and development.

# APPENDIX

## INTRODUCTION

Elspar and Oncaspar were the first amino acid depletion therapies utilized for the treatment of cancer[18,102]. Used for acute lymphoblastic leukemia (ALL), these enzymes deplete asparagine in the serum, exploiting these cells high demand for asparagine. The efficacy and success of these treatments cannot be understated, with 5-year survival rates drastically increasing from 54% in the 1970s to 90% now[15]. Unfortunately, due to the bacterial nature of these enzymes, a number of patients display immune reactions upon treatment, limiting administration in some cases[17,19]. PEGylated versions of the enzyme have been made, limiting the hypersensitivity for some patients, but it has still remained an issue in other cases. Additionally, recent research has demonstrated that, in for many cancer cell lines, the efficacy of L-asparaginases is due to their glutaminase degrading activity[51]. These discoveries have highlighted the need for a new generation of amino acid depletion treatment for ALL. To that end, this work has focused on developing a human-based enzyme capable of degrading glutaminase at therapeutically relevant levels with biophysical properties amenable to industrial production.

In order to develop a new treatment, identification of a scaffold for engineering is paramount. Within the human genome, there are 4 known glutaminases. The first, the kidney-type glutaminase, also known as GLS1, has relatively high activity towards glutamine ($k_{cat}$ = 10 ± 1 $s^{-1}$; $K_M$ = 1.9 ± 0.4 mM; $k_{cat}/K_M$ = (5.3 ± 1.2) x $10^3$ $M^{-1}s^{-1}$)[55]. Unfortunately, GLS1 has been shown to require phosphate concentrations, beyond physiological levels, to obtain its high activity. A second glutaminase, Glutaminase C (GAC), is a splicing variant of GLS1, showing similar activity ($k_{cat}$ = 22 ± 2 $s^{-1}$; $K_M$ = 1.4 ± 0.4 mM; $k_{cat}/K_M$ = (1.6 ± 0.5) x $10^4$ $M^{-1}s^{-1}$), but also requiring similar amounts of

phosphate to achieve this high activity. A third glutaminase, live-type glutaminase, also known as GLS2, does not have much dependence on phosphate but however displays poor catalytic activity ($k_{cat}$ = 2.5 ± 0.3 $s^{-1}$; $K_M$ = 4.0 ± 0.2 mM; $k_{cat}/K_M$ = (6.3 ± 0.8) x $10^2$ $M^{-1}s^{-1}$). A fourth type of glutaminase, a splicing variant of GLS2 has not been studied as exhaustively as the other three, but does have many properties similar to GLS2. Unfortunately, each of these scaffolds have significant problems, preventing their use as a therapeutic with further engineering. Additionally, none of these proteins have been expressed in *E. coli*, representing another hurdle to easy laboratory research. Fortunately, low catalytic activity, allosteric regulation by a small molecule, and poor expression in bacterial strains have all been addressed by protein engineering, and could be solved in this case as well.

Beyond the work done to understand the wild-type scaffolds, some additional work has been done exploring mutations that affect inhibitor binding, and correspondingly the phosphate dependence of the proteins[55,103]. A number of inhibitors have been developed in hopes of discovering a small molecule treatment for cancer. The most well studied inhibitor, bis-2-(5-phenylacetimido-1,2,4,thiadiazol-2-yl)ethyl sulfide, also known as BPTES, is known to alter the protein's quaternary structure, preventing the tetramer associated with phosphate binding[55]. These mutations that affect BPTES binding have not been fully explored, as there is limited knowledge on how they affect other biophysical properties, even, in some cases, the catalytic activity. Regardless, these mutations can serve as a basis for further engineering, hopefully allowing the phosphate dependence to be completely overcome. In spite of the scaffold-specific issues, the previous work done with both the wild-type scaffold and these mutants served as a solid basis for exploring the scaffolds as therapeutic proteins.

## MATERIALS AND METHODS

*Materials:* Oligonucleotides were purchased from Integrated DNA Technologies. Restriction enzymes, Taq polymerase, Phusion high fidelity DNA polymerase, T4 DNA ligase, corresponding buffers, and dNTPs were purchased from New England Biolabs. Difco 2xYT growth media, LB growth media, and TB growth media were all purchased from Becton Dickinson. Nickel-nitrotriacetic acid (Ni-NTA) resin and mini-prep kits were purchased from Qiagen. DL-Homocysteine, homocysteine, L-methionine, DL-cysteine, cystine, 3-methyl-2-benzothiazoninone hydrazine (MBTH), isopropyl β-D-1-thiogalactopyranoside (IPTG), kanamycin, and ampicillin were all purchased from Sigma-Aldrich.

*Mutant Cloning:* Specific Mutants were synthesized by overlap extension PCR. Briefly, an initial round of PCRs was conducted, utilizing gene-specific primers to encode the desired mutations coupled with plasmid-specific primers. Subsequently, a second PCR was initiated, using the products from the mutagenic first PCRs and the plasmid-specific primers. After the second PCR, the product was purified via gel extraction. The gel extracted DNA was digested using NcoI and EcoRI, after which it was ligated into the pET28a plasmid or pET14b plasmid using the T7 ligase protocol. Ligated plasmid was then transformed into DH10β cells and plated on an antibiotic selective agar plate. Individual colonies were picked and their identify confirmed via Sanger sequencing. Confirmed plasmids were re-transformed into either BL21(DE3) cells or C41 cells and again plated on antibiotic selective agar plates. Colonies were picked and used to inoculate overnight cultures, which were then frozen for subsequent usage.

*Expression and Purification:* Glutaminase constructs were expressed in order to assess the biophysical properties of the mutants and optimize their expression in *E. coli*. The pET28a plasmid harboring the glutaminase construct or the fusion construct in the

pET14b plasmid was then transformed into BL21 (DE3) for expression. The cells were grown overnight at 37°C in 5 mL of 2xYT media with 50 μg/mL of kanamycin for the pET28a plasmid or 100 μg/mL of ampicillin for the pET14b plasmid. The subsequent morning, the cultures were inoculated, at a 1:1000 dilution, into 0.5 L of fresh TB media containing the appropriate antibiotic. Upon reaching an $A_{600}$ of 0.6 to 0.8, the culture was moved to 25°C and 1 M isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM to induce T7 RNA polymerase expression, and subsequently the glutaminase construct expression. After overnight incubation, cells were harvested via centrifugation at 5,500 rpm for 20 minutes in a JA-10 rotor. The cell pellet, after removing the supernatant, was resuspended using 25 mL of Glutaminase lysis buffer (50 mM HEPES, 500 mM NaCl, 30 mM Imidazole, pH 7.5). The cell suspension was then lysed by two passes through a French pressure cell and then centrifuged at 14,000 rpm for 60 minutes in a JA-20 rotor. The supernatant from this centrifugation step was then filtered using a 0.45 μm membrane. The clarified supernatant was then applied to 2 mL of nickel-nitrotriacetic acid (Ni-NTA) resin that had been pre-equilibrated with 10 mL of pET lysis buffer. The resin was then washed with 60 mL (30 column volumes) of pET lysis buffer. Afterward, the resin was incubated with 5 mL of Glutaminase elution buffer (50 mM HEPES, 500 mM NaCl, 250 mM Imidazole, pH 7.5), equilibrated for 5 minutes, and then eluted, capturing the elution in a separate vessel. The elution step was then repeated to increase recovery. The elution was then buffer exchanged against a Glutaminase storage buffer (50 mM HEPES, 50 mM KCl, 2 mM DTT, pH 8.0) utilizing dialysis membranes over the course of 2 days, changing out the buffer every 12 hours. For fusion constructs, sumo protease was added at a 1:1000 molar ratio and incubated overnight at 4°C. The sumo protease, cleaved sumo tag, and uncleaved fusion was captured by passing the reaction over a fresh Ni-NTA column. The flowthrough was captured and buffer exchanged again.

After completing the final buffer exchange, glycerol was added to a final concentration of 10% v/v. The purified enzyme was then aliquoted, flash frozen with liquid nitrogen, and stored at -80°C. Purity of the proteins were analyzed via sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) on a 4% to 20% pre-cast Tris-Glycine gel run under reducing conditions (2.5% β-Mercaptoethanol (BME)). Protein concentration was determined via $A_{280}$ under denaturing conditions using a calculated extinction coefficient of 45,471 $M^{-1}cm^{-1}L$.

*Enzyme Activity Assay:* Kinetic analysis was achieved by measuring glutamate production from glutamine. For the kinetic reaction, 90 μL of substrate was pre-incubated at 37°C for 5 minutes. 10 μL of an enzyme solution was then added and thoroughly mixed. After the pre-determined reaction time, 10 μL of 50.0% w/v trichloroacetic acid was added to quench the reaction. The final mixture was then analyzed by derivatization with o-phthaldialdehyde (OPA). Derivatization and chromatography were adapted from the referenced literature[86].

## RESULTS AND DISCUSSION

### *Expression of Glutaminase Constructs*

Based on the isoforms and truncations previously reported, multiple glutaminase constructs were constructed, shown in Table A.1. As previously discussed, none of the isoforms have previously been purified from *E. coli*, a necessary step for any therapeutic protein. As such, before testing the kinetics and biophysical properties, an optimal expression system needed to be developed. For the first expression test, the constructs were expressed in *E. coli* BL21(DE3) in rich media, utilizing the pET28a plasmid. Utilizing this system, 4 of the 6 constructs were found to express soluble, detectable protein, although the yield was low as was the purity. The yields can be seen in Table A.1. Examination of the whole cell fraction revealed that the constructs expressed much better than the purification indicated, suggesting that a significant fraction of the protein was insoluble. To correct this, the constructs were re-expressed in the *E. coli* C41(DE3) strain in attempt to slow expression and increase the fraction of the expressed protein that is soluble. The results of this second expression test can also be seen in Table A.1. The C41 strain improved the soluble yield of two constructs, GACΔ70 and GACΔ133, as well as their purity, but did not improve GLS2 and GLS1Δ70 while decreasing the recovery for GLS1Δ133 and GAC. Thus the C41 strain, although improving the yield and purity of a few of the constructs, still did not completely address the insolubility observed with the scaffolds.

To further improve the expression, an N-terminal fusion with the sumo protein from *B. distachyon* was constructed for each glutaminase construct[104,105]. For these fusions, an N-terminal 6xHis tag was attached to the sumo protein, allowing for purification, after which the 6xHis tag and the sumo protein could be cleaved off by a sumo protease. The fusion constructs' yield was first assessed in *E. coli* BL21(DE3); the results of which can

be seen in Table A.1. The fusion increased the yield of the GLS1Δ133 construct compared to its native sequence. Surprisingly, the GLS1Δ70 fusion, which when expressed natively in the BL21 strain and the C41 strain was not detected, was expressed at relatively high levels. Unfortunately, the other 4 constructs did not show detectable expression. As a final test, all of the sumo fusion constructs were expressed in the C41 strain, to assess if slower expression would improve the yield. The results of this last expression test can also be seen in Table A.1. The sumo fusion constructs all expressed at detectable levels in the C41 strain, a number of which has good purity as well. Although not necessarily at the levels required for industrial production, this expression system allowed for sufficient production, enabling kinetic and biophysical characterization of the constructs. Since the GAC scaffold, specifically the GACΔ133 truncation, has the highest kinetics reported in literature, it was moved forward for characterization first.

| | BL21 (DE3) | | C41 (DE3) | |
|---|---|---|---|---|
| | N-term 6xHis | N-term Sumo | N-term 6xHis | N-term Sumo |
| GLS2 | ND | ND | ND | 31 mg/L* |
| GLS1Δ70 | ND | 22 mg/L* | ND | 31 mg/L |
| GLS1Δ133 | 20 mg/L* | 52 mg/L* | ND | 32 mg/L |
| GAC | 5 mg/L* | ND | ND | 12 mg/L* |
| GACΔ70 | 12 mg/L* | ND | 60 mg/L | 23 mg/L |
| GACΔ133 | 16 mg/L* | ND | 41 mg/L | 24 mg/L |

Table A.1 Glutaminase Constructs and Yield in Various *E. coli* Strains

The yield of the various glutaminase constructs when purified in the various *E. coli* strains. ND = not detectable; * = low purity

*Characterization of Glutaminase Constructs*

Upon successfully expressing the glutaminase constructs, the next step was characterization of their kinetics, beginning with the GACΔ133 construct. Due to the phosphate dependence of the scaffold, the kinetics were tested at utilizing the glutaminase storage buffer supplemented with 200 mM phosphate, to determine maximal activity, and also supplemented with 2 mM phosphate, to determine activity at physiological concentrations. Under the high phosphate conditions, the GACΔ133 construct was found to have a $k_{cat}$ of 550 $s^{-1}$, a $K_M$ of 3.4 mM, and a $k_{cat}/K_M$ of $1.6*10^5$ $s^{-1}mM^{-1}$. While the $K_M$ is within the expected range, based on literature, the $k_{cat}$ was significantly higher than expected, by an order of magnitude. This unusually high turnover rate was the first red flag during the characterization of the glutaminase scaffold.

Simultaneously to the characterization of the GACΔ133 scaffold, a number of mutants, based on literature, were cloned and expressed due to their decreased dependence

| | Mutations |
|---|---|
| Mutant 1 | GACΔ133 K311A |
| Mutant 2 | GACΔ133 K320A |
| Mutant 3 | GACΔ133 K328A |
| Mutant 4 | GACΔ133 K396A |
| Mutant 5 | GACΔ133 F318Y F322S |
| Mutant 6 | GACΔ133 Y394L |

Table A.2 GACΔ133 Mutants

The GACΔ133 mutants

on phosphate[55,103]. The list of these mutants can be seen in Table A.2. During the characterization of these mutants, it was discovered that the wild-type scaffold, GACΔ133, precipitously lost activity upon storage at 4°C, with less than 25% of its activity remaining after 3 days. Suspecting possible dissociation of the tetramer, and thus loss of activity, GACΔ133 was analyzed by SEC-FPLC immediately after purification and at various time points during storage. Additionally, the 6 mutants were analyzed by SEC-FPLC to determine if the mutations affected oligomerization and stability of the quaternary structure. The results comparing the 6 mutants with the wild-type, truncated scaffold can be seen in Figure A.1. A large variance can be seen in the elution volumes between 12.5 mL and 15 mL, corresponding to the size of the tetramer. All of the variants show some amount of monomer, as seen at ~16.5 mL. Interestingly, all of the variants show a peak at

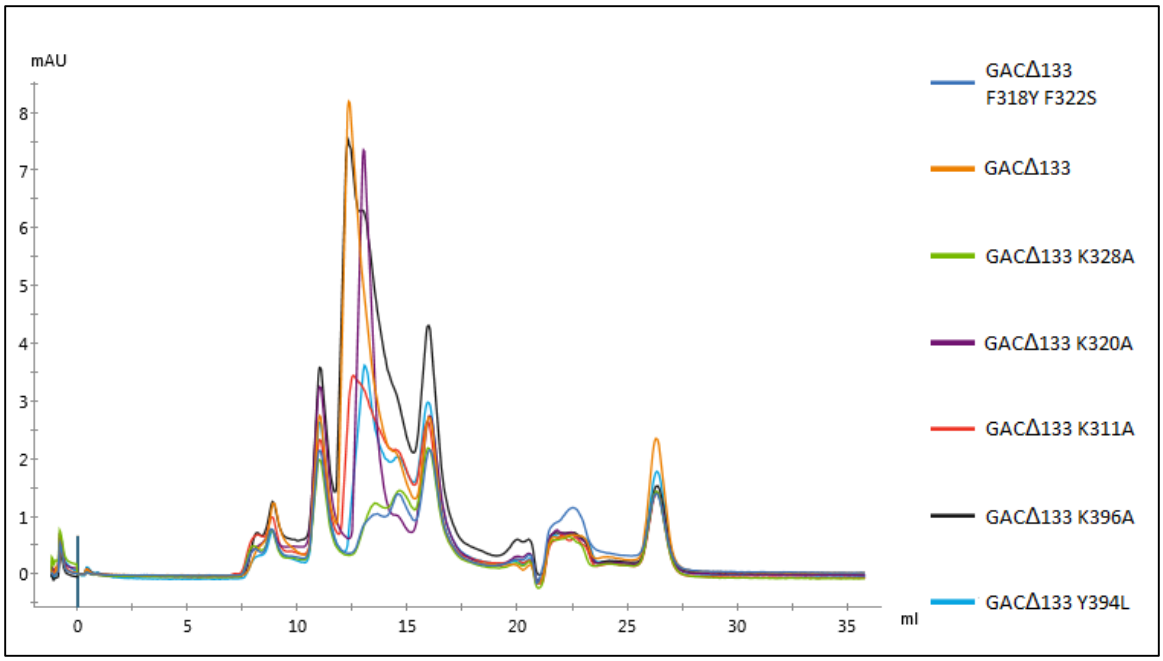


Figure A.1 SEC-FPLC Profile of GACΔ133 Mutants

SEC-FPLC profile of GACΔ133 mutants. The monomer is observed at ~16.5 mL, the tetramer at ~13 mL, and a decamer at ~11.5 mL.

~11.5 mL, which, based on the expected size for that elution time, corresponds to a decamer. The presence of this higher order oligomer raised further red flags, as having multiple oligomer forms would complicate purification and present heterogeneity issues.

In light of these results, we investigated the scaffold further, trying to understand the higher order oligomers as well as the unexpectedly high turnover rate. Recent literature had identified that GLS1 proteins are capable of forming long filamentous structures, spanning over a micron in length, in a phosphate dependent fashion[106]. When in this filamentous form, GLS1 activity significantly increased, as was previously observed under high phosphate activity, suggesting strong cooperativity between activity sites in this form. This discovery, also showed that some of the phosphate-independent mutations also assemble into this filament without phosphate. Additionally, it was found that GLS2, which does not display the phosphate-dependence observed in GLS1, does not assemble into the filaments, further confirming the link between filament formation, phosphate activation, and increased activity. Ultimately, these findings explained the higher catalytic activity as well as the higher order oligomers in the previous experiments.

Unfortunately, the requirement of this filamentous form for high activity represents a major hurdle for therapeutic development, relegating further research to focus on the GLS2 scaffold or engineering the GLS1 scaffold to avoid higher order assembly. Moving forward, future work could be focused on stabilizing an intermediate oligomer, to avoid the extremely long filaments, but preserve the cooperativity bestowed by such a form. Alternatively, work could focus on engineering the activity site of GLS2 to improve its activity without self-assembly as seen in GLS1. In either case, this work did develop a system for expressing the glutaminase constructs in *E. coli*, hopefully enabling future work on these proteins.

**ACKNOWLEDGEMENTS**

I would like to acknowledge and thank Dr. Giulia Agnello for her insight into and discussions of this project. I would also like to acknowledge and thank Dr. Christos Karamitros for his contribution of the sumo plasmid, the MBP-sumo plasmid, and the sumo protease plasmid.

# References


1. Clark, A. J. *et al.* Biosynthetic human insulin in the treatment of diabetes. A double-blind crossover trial in established diabetic patients. *Lancet* **2**, 354–357 (1982).
2. Leader, B., Baca, Q. J. & Golan, D. E. Protein therapeutics: a summary and pharmacological classification. *Nat Rev Drug Discov* **7**, 21–39 (2008).
3. Usmani, S. S. *et al.* THPdb: Database of FDA-approved peptide and protein therapeutics. *PLoS One* **12**, (2017).
4. Pasut, G. & Veronese, F. M. State of the art in PEGylation: The great versatility achieved after forty years of research. *Journal of Controlled Release* **161**, 461–472 (2012).
5. Molineux, G. Pegylation: engineering improved biopharmaceuticals for oncology. *Pharmacotherapy* **23**, 3S-8S (2003).
6. Zheng, J. *et al.* PEGylation is effective in reducing immunogenicity, immunotoxicity, and hepatotoxicity of α-momorcharin in vivo. *Immunopharmacol Immunotoxicol* **34**, 866–873 (2012).
7. Tobin, P. H., Richards, D. H., Callender, R. A. & Wilson, C. J. Protein Engineering: A New Frontier for Biological Therapeutics. *Curr Drug Metab* **15**, 743–756 (2014).
8. Siegel, R. L., Miller, K. D. & Jemal, A. Cancer statistics, 2017. *CA: A Cancer Journal for Clinicians* **67**, 7–30 (2017).

9. Arruebo, M. *et al.* Assessment of the Evolution of Cancer Treatment Therapies. *Cancers (Basel)* **3**, 3279–3330 (2011).
10. Warburg, O., Wind, F. & Negelein, E. THE METABOLISM OF TUMORS IN THE BODY. *J Gen Physiol* **8**, 519–530 (1927).
11. Potter, M., Newport, E. & Morten, K. J. The Warburg effect: 80 years on. *Biochem Soc Trans* **44**, 1499–1505 (2016).
12. Vander Heiden, M. G., Cantley, L. C. & Thompson, C. B. Understanding the Warburg Effect: The Metabolic Requirements of Cell Proliferation. *Science* **324**, 1029–1033 (2009).
13. Kano, Y. *et al.* Methionine Dependency of Cell Growth in Normal and Malignant Hematopoietic Cells. *Cancer Res* **42**, 3090–3092 (1982).
14. Scott, L., Lamb, J., Smith, S. & Wheatley, D. N. Single amino acid (arginine) deprivation: rapid and selective death of cultured transformed and malignant cells. *British Journal of Cancer* **83**, 800–810 (2000).
15. Egler, R. A., Ahuja, S. P. & Matloub, Y. L-asparaginase in the treatment of patients with acute lymphoblastic leukemia. *J Pharmacol Pharmacother* **7**, 62–71 (2016).
16. Ali, U. *et al.* L-asparaginase as a critical component to combat Acute Lymphoblastic Leukaemia (ALL): A novel approach to target ALL. *Eur. J. Pharmacol.* **771**, 199–210 (2016).
17. Oettgen, H. F. *et al.* Toxicity of E. coli L-asparaginase in man. *Cancer* **25**, 253–278 (1970).

18. Dinndorf, P. A., Gootenberg, J., Cohen, M. H., Keegan, P. & Pazdur, R. FDA drug approval summary: pegaspargase (oncaspar) for the first-line treatment of children with acute lymphoblastic leukemia (ALL). *Oncologist* **12**, 991–998 (2007).
19. Raetz, E. A. & Salzer, W. L. Tolerability and efficacy of L-asparaginase therapy in pediatric patients with acute lymphoblastic leukemia. *J. Pediatr. Hematol. Oncol.* **32**, 554–563 (2010).
20. Cramer, S. L. *et al.* Systemic depletion of L-cyst(e)ine with cyst(e)inase increases reactive oxygen species and suppresses tumor growth. *Nat. Med.* **23**, 120–127 (2017).
21. Hu, J. & Cheung, N.-K. V. Methionine depletion with recombinant methioninase: In vitro and in vivo efficacy against neuroblastoma and its synergism with chemotherapeutic drugs. *Int J Cancer* **124**, 1700–1706 (2009).
22. Glazer, E. S. *et al.* Bioengineered Human Arginase I with Enhanced Activity and Stability Controls Hepatocellular and Pancreatic Carcinoma Xenografts. *Transl Oncol* **4**, 138–146 (2011).
23. Chiu, M. *et al.* Glutamine depletion by crisantaspase hinders the growth of human hepatocellular carcinoma xenografts. *Br J Cancer* **111**, 1159–1167 (2014).
24. Shanware, N. P., Mullen, A. R., DeBerardinis, R. J. & Abraham, R. T. Glutamine: pleiotropic roles in tumor growth and stress resistance. *J Mol Med* **89**, 229–236 (2011).
25. Ezgu, F. Inborn Errors of Metabolism. *Adv Clin Chem* **73**, 195–250 (2016).

26. Futerman, A. H., Sussman, J. L., Horowitz, M., Silman, I. & Zimran, A. New directions in the treatment of Gaucher disease. *Trends Pharmacol. Sci.* **25**, 147–151 (2004).
27. Whittington, R. & Goa, K. L. Alglucerase. A review of its therapeutic use in Gaucher's disease. *Drugs* **44**, 72–93 (1992).
28. Sacharow, S. J., Picker, J. D. & Levy, H. L. Homocystinuria Caused by Cystathionine Beta-Synthase Deficiency. in *GeneReviews*® (eds. Adam, M. P. et al.) (University of Washington, Seattle, 1993).
29. Skiba, W. E., Taylor, M. P., Wells, M. S., Mangum, J. H. & Awad, W. M. Human hepatic methionine biosynthesis. Purification and characterization of betaine:homocysteine S-methyltransferase. *J. Biol. Chem.* **257**, 14944–14948 (1982).
30. Bukovska, G., Kery, V. & Kraus, J. P. Expression of human cystathionine beta-synthase in Escherichia coli: purification and characterization. *Protein Expr. Purif.* **5**, 442–448 (1994).
31. Kraus, J., Packman, S., Fowler, B. & Rosenberg, L. E. Purification and properties of cystathionine beta-synthase from human liver. Evidence for identical subunits. *J. Biol. Chem.* **253**, 6523–6528 (1978).
32. Mudd, S. H., Finkelstein, J. D., Irreverre, F. & Laster, L. HOMOCYSTINURIA: AN ENZYMATIC DEFECT. *Science* **143**, 1443–1445 (1964).
33. Mudd, S. H. *et al.* The natural history of homocystinuria due to cystathionine beta-synthase deficiency. *Am. J. Hum. Genet.* **37**, 1–31 (1985).

34. Yap, S., Rushe, H., Howard, P. M. & Naughten, E. R. The intellectual abilities of early-treated individuals with pyridoxine-nonresponsive homocystinuria due to cystathionine beta-synthase deficiency. *J. Inherit. Metab. Dis.* **24**, 437–447 (2001).
35. Yap, S. Classical homocystinuria: vascular risk and its prevention. *J. Inherit. Metab. Dis.* **26**, 259–265 (2003).
36. Mulvihill, A., Yap, S., O'Keefe, M., Howard, P. M. & Naughten, E. R. Ocular findings among patients with late-diagnosed or poorly controlled homocystinuria compared with a screened, well-controlled population. *J AAPOS* **5**, 311–315 (2001).
37. Shenoy, V., Mehendale, V., Prabhu, K., Shetty, R. & Rao, P. Correlation of Serum Homocysteine Levels with the Severity of Coronary Artery Disease. *Indian J Clin Biochem* **29**, 339–344 (2014).
38. Walter, J. H., Wraith, J. E., White, F. J., Bridge, C. & Till, J. Strategies for the treatment of cystathionine beta-synthase deficiency: the experience of the Willink Biochemical Genetics Unit over the past 30 years. *Eur. J. Pediatr.* **157 Suppl 2**, S71-76 (1998).
39. Komrower, G. M., Lambert, A. M., Cusworth, D. C. & Westall, R. G. Dietary treatment of homocystinuria. *Arch Dis Child* **41**, 666–671 (1966).
40. Sakamoto, A. & Sakura, N. Limited effectiveness of betaine therapy for cystathionine beta synthase deficiency. *Pediatr Int* **45**, 333–338 (2003).
41. Bublil, E. M. *et al.* Enzyme replacement with PEGylated cystathionine β-synthase ameliorates homocystinuria in murine model. *J. Clin. Invest.* **126**, 2372–2384 (2016).

42. Majtan, T. *et al.* Enzyme replacement prevents neonatal death, liver damage, and osteoporosis in murine homocystinuria. *FASEB J* fj.201700565R (2017). doi:10.1096/fj.201700565R

43. Majtan, T., Park, I., Bublil, E. M. & Kraus, J. P. Enzyme replacement therapy prevents loss of bone and fat mass in murine homocystinuria. *Hum. Mutat.* **39**, 210–218 (2018).

44. Meier, M., Janosik, M., Kery, V., Kraus, J. P. & Burkhard, P. Structure of human cystathionine β-synthase: a unique pyridoxal 5′-phosphate-dependent heme protein. *The EMBO Journal* **20**, 3910–3916 (2001).

45. Janosík, M., Kery, V., Gaustadnes, M., Maclean, K. N. & Kraus, J. P. Regulation of human cystathionine beta-synthase by S-adenosyl-L-methionine: evidence for two catalytically active conformations involving an autoinhibitory domain in the C-terminal region. *Biochemistry* **40**, 10625–10633 (2001).

46. Banerjee, R., Evande, R., Kabil, O., Ojha, S. & Taoka, S. Reaction mechanism and regulation of cystathionine beta-synthase. *Biochim. Biophys. Acta* **1647**, 30–35 (2003).

47. Miles, E. W. & Kraus, J. P. Cystathionine β-Synthase: Structure, Function, Regulation, and Location of Homocystinuria-causing Mutations. *J. Biol. Chem.* **279**, 29871–29874 (2004).

48. Frank, N., Kent, J. O., Meier, M. & Kraus, J. P. Purification and characterization of the wild type and truncated human cystathionine β-synthase enzymes expressed in E. coli. *Arch Biochem Biophys* **470**, 64–72 (2008).

49. Cantor, J. R., Stone, E. M., Chantranupong, L. & Georgiou, G. The human asparaginase-like protein 1 hASRGL1 is an Ntn hydrolase with beta-aspartyl peptidase activity. *Biochemistry* **48**, 11026–11031 (2009).

50. Li, W. *et al.* Uncoupling intramolecular processing and substrate hydrolysis in the N-terminal nucleophile hydrolase hASRGL1 by circular permutation. *ACS Chem. Biol.* **7**, 1840–1847 (2012).

51. Parmentier, J. H. *et al.* Glutaminase Activity Determines Cytotoxicity of L-Asparaginases on Most Leukemia Cell Lines. *Leuk Res* **39**, 757–762 (2015).

52. Fendt, S.-M. *et al.* Metformin Decreases Glucose Oxidation and Increases the Dependency of Prostate Cancer Cells on Reductive Glutamine Metabolism. *Cancer Res* **73**, 4429–4438 (2013).

53. Gross, M. I. *et al.* Antitumor Activity of the Glutaminase Inhibitor CB-839 in Triple-Negative Breast Cancer. *Mol Cancer Ther* **13**, 890–901 (2014).

54. Tardito, S. *et al.* L-Asparaginase and inhibitors of glutamine synthetase disclose glutamine addiction of β-catenin-mutated human hepatocellular carcinoma cells. *Curr Cancer Drug Targets* **11**, 929–943 (2011).

55. DeLaBarre, B. *et al.* Full-Length Human Glutaminase in Complex with an Allosteric Inhibitor. *Biochemistry* **50**, 10764–10770 (2011).

56. Szeliga, M. & Obara-Michlewska, M. Glutamine in neoplastic cells: focus on the expression and roles of glutaminases. *Neurochem. Int.* **55**, 71–75 (2009).

57. Paley, O. *et al.* GFP reporter screens for the engineering of amino acid degrading enzymes from libraries expressed in bacteria. *Methods Mol. Biol.* **978**, 31–44 (2013).

58. Firnberg, E. & Ostermeier, M. PFunkel: Efficient, Expansive, User-Defined Mutagenesis. *PLOS ONE* **7**, e52031 (2012).

59. Pabis, A., Risso, V. A., Sanchez-Ruiz, J. M. & Kamerlin, S. C. Cooperativity and flexibility in enzyme evolution. *Current Opinion in Structural Biology* **48**, 83–92 (2018).

60. Salmon, K. A., Yang, C.-R. & Hatfield, G. W. Biosynthesis and Regulation of the Branched-Chain Amino Acids†. *EcoSal Plus* **2**, (2006).

61. Gallagher, D. T., Chinchilla, D., Lau, H. & Eisenstein, E. Local and global control mechanisms in allosteric threonine deaminase. *Meth. Enzymol.* **380**, 85–106 (2004).

62. Changeux, J. P. The feedback control mechanisms of biosynthetic L-threonine deaminase by L-isoleucine. *Cold Spring Harb. Symp. Quant. Biol.* **26**, 313–318 (1961).

63. Merino, E. & Yanofsky, C. Transcription attenuation: a highly conserved regulatory strategy used by bacteria. *Trends Genet.* **21**, 260–264 (2005).

64. Gollop, N., Damri, B., Barak, Z. & Chipman, D. M. Kinetics and mechanism of acetohydroxy acid synthase isozyme III from Escherichia coli. *Biochemistry* **28**, 6310–6317 (1989).

65. Guardiola, J., De Felice, M., Lamberti, A. & Iaccarino, M. The acetolactate synthase isoenzymes of Escherichia coli K-12. *Molec. Gen. Genet.* **156**, 17–25 (1977).

66. Barak, Z., Chipman, D. M. & Gollop, N. Physiological implications of the specificity of acetohydroxy acid synthase isozymes of enteric bacteria. *J. Bacteriol.* **169**, 3750–3756 (1987).

67. Bonner, D. Further studies of mutant strains of Neurospora requiring isoleucine and valine. *J. Biol. Chem.* **166**, 545–554 (1946).

68. De Felice, M. *et al.* Growth inhibition of Escherichia coli K-12 by L-valine: a consequence of a regulatory pattern. *Mol. Gen. Genet.* **156**, 1–7 (1977).

69. Lu, M. F. & Umbarger, H. E. Effects of deletion and insertion mutations in the ilvM gene of Escherichia coli. *J Bacteriol* **169**, 600–604 (1987).

70. Parekh, B. S. & Hatfield, G. W. Growth rate-related regulation of the ilvGMEDA operon of Escherichia coli K-12 is a consequence of the polar frameshift mutation in the ilvG gene of this strain. *J. Bacteriol.* **179**, 2086–2088 (1997).

71. Coukell, M. B. & Polglase, W. J. Repression by glucose of acetohydroxy acid synthetase in Escherichia coli B. *Biochem J* **111**, 273–278 (1969).

72. De Felice, M., Squires, C. & Levinthal, M. A comparative study of the acetohydroxy acid synthase isoenzymes of Escherichia coli K-12. *Biochimica et Biophysica Acta (BBA) - General Subjects* **541**, 9–17 (1978).

73. Gollop, N., Chipman, D. M. & Barak, Z. Inhibition of acetohydroxy acid synthase by leucine. *Biochim. Biophys. Acta* **748**, 34–39 (1983).

74. de Felice, M., Lago, C. T., Squires, C. H. & Calvo, J. M. Acetohydroxy acid synthase isoenzymes of Escherichia coli K12 and Salmonella typhimurium. *Ann. Microbiol. (Paris)* **133**, 251–256 (1982).

75. Umbarger, H. E. Amino acid biosynthesis and its regulation. *Annu. Rev. Biochem.* **47**, 532–606 (1978).

76. Ernst, D. C. & Downs, D. M. 2-Aminoacrylate Stress Induces a Context-Dependent Glycine Requirement in ridA Strains of Salmonella enterica. *J Bacteriol* **198**, 536–543 (2016).

77. Daniel, J., Dondon, L. & Danchin, A. 2-ketobutyrate: A putative alarmone of Escherichia coli. *Mol Gen Genet* **190**, 452–458 (1983).

78. Danchin, A., Dondon, L. & Daniel, J. Metabolic alterations mediated by 2-ketobutyrate in Escherichia coli K12. *Molec. Gen. Genet.* **193**, 473–478 (1984).

79. Dalebroux, Z. D. & Swanson, M. S. ppGpp: magic beyond RNA polymerase. *Nature Reviews Microbiology* **10**, 203–212 (2012).

80. Lambrecht, J. A., Flynn, J. M. & Downs, D. M. Conserved YjgF protein family deaminates reactive enamine/imine intermediates of pyridoxal 5'-phosphate (PLP)-dependent enzyme reactions. *J. Biol. Chem.* **287**, 3454–3461 (2012).

81. Dumon-Seignovert, L., Cariot, G. & Vuillard, L. The toxicity of recombinant proteins in Escherichia coli: a comparison of overexpression in BL21(DE3), C41(DE3), and C43(DE3). *Protein Expr. Purif.* **37**, 203–206 (2004).

82. Tuite, N. L., Fraser, K. R. & O'Byrne, C. P. Homocysteine Toxicity in Escherichia coli Is Caused by a Perturbation of Branched-Chain Amino Acid Biosynthesis. *J. Bacteriol.* **187**, 4362–4371 (2005).

83. Javor, G. T., Stringer, C. D. & Ryu, J. Thiol-sensitive promoters of Escherichia coli. *J. Bacteriol.* **170**, 3291–3293 (1988).

84. Nakayashiki, T. *et al.* The tRNA Thiolation Pathway Modulates the Intracellular Redox State in Escherichia coli. *Journal of Bacteriology* **195**, 2039–2049 (2013).

85. Daskalakis, I. *et al.* Determination of plasma total homocysteine and cysteine using HPLC with fluorescence detection and an ammonium 7-fluoro-2, 1, 3-benzoxadiazole-4-sulphonate (SBD-F) derivatization protocol optimized for antioxidant concentration, derivatization reagent concentration, temperature and matrix pH. *Biomed. Chromatogr.* **10**, 205–212 (1996).
86. Gardner, W. S. & Miller, W. H. Reverse-phase liquid chromatographic analysis of amino acids after reaction with o-phthalaldehyde. *Anal. Biochem.* **101**, 61–65 (1980).
87. Stone, E. *et al.* De novo engineering of a human cystathionine-γ-lyase for systemic (L)-Methionine depletion cancer therapy. *ACS Chem. Biol.* **7**, 1822–1829 (2012).
88. Velez-Carrasco, W., Merkel, M., Twiss, C. O. & Smith, J. D. Dietary methionine effects on plasma homocysteine and HDL metabolism in mice. *J Nutr Biochem* **19**, 362–370 (2008).
89. Gupta, S. *et al.* Mouse models of cystathionine beta-synthase deficiency reveal significant threshold effects of hyperhomocysteinemia. *FASEB J.* **23**, 883–893 (2009).
90. Halperin, B. A. *et al.* Kinetics of the antibody response to tetanus-diphtheria-acellular pertussis vaccine in women of childbearing age and postpartum women. *Clin. Infect. Dis.* **53**, 885–892 (2011).
91. Watanabe, M. *et al.* Mice deficient in cystathionine beta-synthase: animal models for mild and severe homocyst(e)inemia. *Proc Natl Acad Sci U S A* **92**, 1585–1589 (1995).

92. Lau, J. L. & Dunn, M. K. Therapeutic peptides: Historical perspectives, current development trends, and future directions. *Bioorg. Med. Chem.* **26**, 2700–2707 (2018).

93. Jaradat, D. M. M. Thirteen decades of peptide synthesis: key developments in solid phase peptide synthesis and amide bond formation utilized in peptide ligation. *Amino Acids* **50**, 39–68 (2018).

94. Schmidt, M. *et al.* Omniligase-1: A Powerful Tool for Peptide Head-to-Tail Cyclization. *Advanced Synthesis & Catalysis* **359**, 2050–2055 (2017).

95. Xu, M. Q., Paulus, H. & Chong, S. Fusions to self-splicing inteins for protein purification. *Meth. Enzymol.* **326**, 376–418 (2000).

96. Mao, H., Hart, S. A., Schink, A. & Pollok, B. A. Sortase-mediated protein ligation: a new method for protein engineering. *J. Am. Chem. Soc.* **126**, 2670–2671 (2004).

97. Chong, S. *et al.* Utilizing the C-terminal cleavage activity of a protein splicing element to purify recombinant proteins in a single chromatographic step. *Nucleic Acids Res* **26**, 5109–5115 (1998).

98. Chong, S. *et al.* Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element. *Gene* **192**, 271–281 (1997).

99. Banki, M. R. & Wood, D. W. Inteins and affinity resin substitutes for protein purification and scale up. *Microb Cell Fact* **4**, 32 (2005).

100. Su, Y. *et al.* Free glycine accelerates the autoproteolytic activation of human asparaginase. *Chem. Biol.* **20**, 533–540 (2013).

101. Wang, Y. *et al.* A novel strategy to engineer DNA polymerases for enhanced processivity and improved performance in vitro. *Nucleic Acids Res* **32**, 1197–1207 (2004).

102. Avramis, V. I. & Tiwari, P. N. Asparaginase (native ASNase or pegylated ASNase) in the treatment of acute lymphoblastic leukemia. *Int J Nanomedicine* **1**, 241–254 (2006).

103. McDonald, C. J., Acheff, E., Kennedy, R., Taylor, L. & Curthoys, N. P. Effect of lysine to alanine mutations on the phosphate activation and BPTES inhibition of glutaminase. *Neurochem Int* **88**, 10–14 (2015).

104. Frey, S. & Görlich, D. A new set of highly efficient, tag-cleaving proteases for purifying recombinant proteins. *J Chromatogr A* **1337**, 95–105 (2014).

105. Frey, S. & Görlich, D. Purification of protein complexes of defined subunit stoichiometry using a set of orthogonal, tag-cleaving proteases. *J Chromatogr A* **1337**, 106–115 (2014).

106. Ferreira, A. P. S. *et al.* Active Glutaminase C Self-assembles into a Supratetrameric Oligomer That Can Be Disrupted by an Allosteric Inhibitor. *J Biol Chem* **288**, 28009–28020 (2013).