1-1-2013

Biophysical Approaches to Human Genetic Disease: ADSL Deficiency as a Model

Stephen P. Ray
University of Denver

BIOPHYSICAL APPROACHES TO HUMAN GENETIC DISEASE: ADSL DEFICIENCY AS A MODEL

A Dissertation

Presented to

the Faculty of Natural Sciences and Mathematics

University of Denver

---

in Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

by

Stephen P. Ray

August 2013

Advisor: Sean E. Shaheen

Author: Stephen P. Ray
Title: Biophysical Approaches to Human Genetic Disease: ADSL Deficiency as a Model
Advisor: Sean E. Shaheen
Degree Date: August 2013

Abstract

Biophysical tools are becoming more useful in approaching human genetic disease. The pathogenic mechanisms behind many disorders have been attributed to protein loss of function or gain in toxic function. For example, changes in the protein Adenylosuccinate Lyase (ADSL) lead to ADSL deficiency, a disorder that causes symptoms ranging from epilepsy to expression of autistic features. Biophysics offers different tools to study intrinsic properties of proteins. We have applied such tools to study the enzyme kinetics and protein stability of ADSL and two mutations (R426H and R303C) to formulate better hypotheses regarding the pathogenic mechanism. Enzyme kinetic findings indicate a) non-linear dependences of the activities on the substrate ratios due to competitive binding and b) distinct differences in the behaviors of the different mutations. Preliminary stability measurements indicate a) there is an intermediate state in the folding pathway and a fourth state is necessary for modeling, b) R426H, a mutation that causes severe phenotypes, is less stable, c) there is no evidence of a gain in toxic function for either mutation. This complete work supports our hypothesis that there are multiple mechanisms of pathogenesis of ADSL deficiency including instability of the functional ADSL tetramer, aggregation of ADSL, or diminution of enzyme activity. The signif-

icance of this is that development of an effective therapy may depend upon which mechanism is active in a particular patient.

Acknowledgements

David Patterson, Kingshuk Ghosh, Terry Wilkinson II, Nathan Duval, Guido Vacano, Patterson Lab, Roberta Colman, Scott Pegan, Michelle Deaton, Sion Ledbetter, Stephanie Spor, David Phillips, Brooke Hirsch, Paulina Filus

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 ADSL Deficiency

All living organisms either make purines themselves via the de novo purine biosynthesis pathway (Figure 1.1) or obtain them from their environment. Purines are vital components of all living cells, fulfilling a wide range of diverse biological functions. They are the building blocks of DNA and RNA, the basic elements of the cell programming machinery [40]. In addition they fulfill a variety of functions in the metabolism of the cell, of which the most important are regulation of cell metabolism and function, energy conservation and transport, intracellular signaling and regulation of gene transcription, formation of coenzymes and of active intermediates of phospholipid and carbohydrate metabolism [40, 85]. Purine metabolism is highly conserved, especially from yeast to mammalian species. This is likely due to the fact purine metabolism is the first known metabolic pathway to develop (Figure 1.1). It is believed the pathway was developed so early because prebiotic chemistry was insufficient to

keep up with the early demand for purines. Evolution of the pathway required the recruitment of multifunctional enzymes that are closely related in their folding structure [14]. The recruitment for the purine biosynthesis pathway was the first completed pathway that interestingly has not furthered evolved (see Figure 1.1), further depicting the importance of completing the pathway as soon as possible for life to progress.

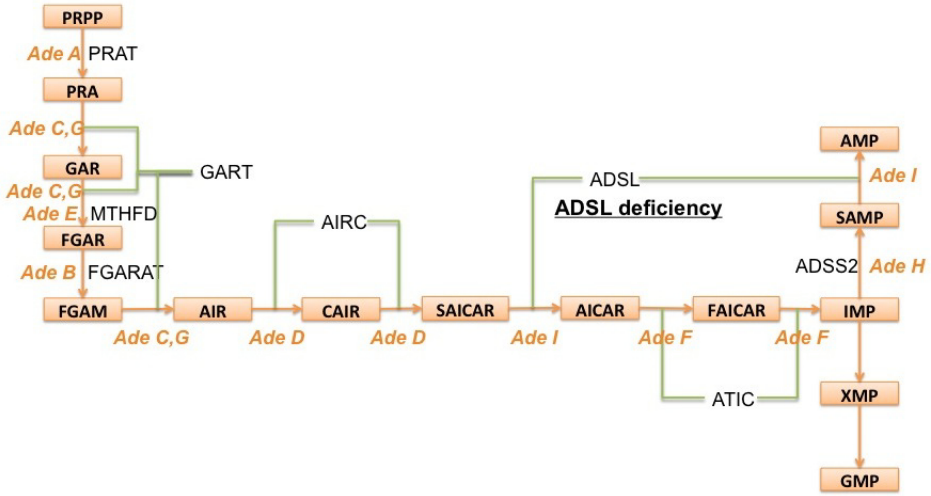


Figure 1.1: Purine pathway

Because of the critical nature of purine metabolism, disturbances within the pathway lead to human inborn errors of metabolism (A recent summary of these can be found in Knox et al [45]). For example, point mutations of the protein Adenylosuccinate Lyase (ADSL), an enzyme involved in two steps of the de novo purine biosynthesis pathway and the purine nucleotide cycle, lead to ADSL deficiency, an autosomal recessive disorder first reported in 1984 [36]. Since then many other patients in different ethnic groups have been detected with this disorder characterized by a spectrum of neurological and physio-

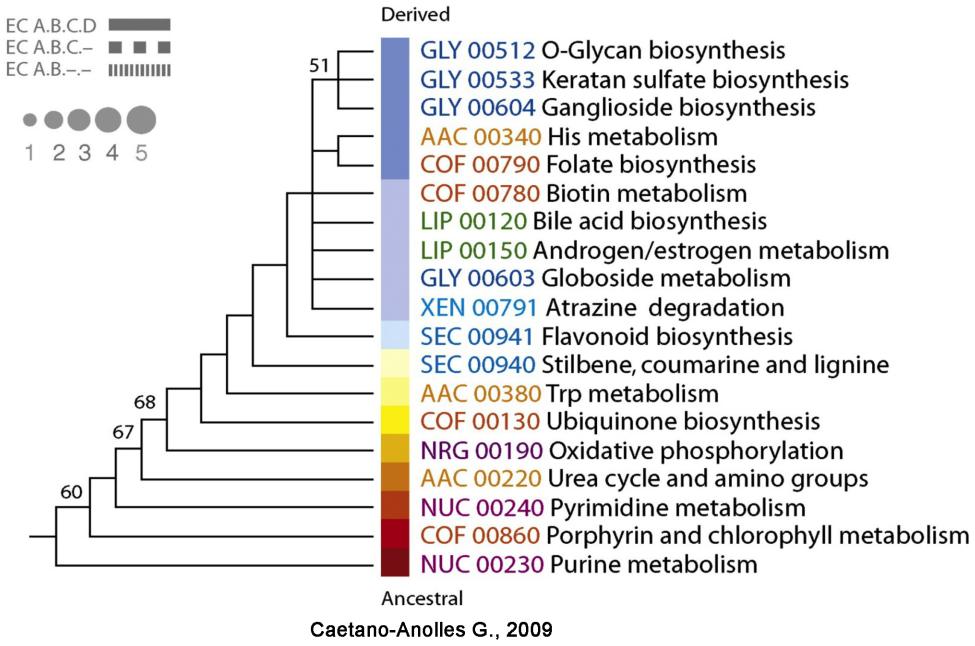


Figure 1.2: Metabolism evolution

logical symptoms such as psychomotor retardation, epilepsy, and expression of autistic features such as lack of eye contact, language and aggressive attitude [17]. Other phenotypes also include axial hypotonia, peripheral hypertonicity, ataxia, muscle wasting, growth retardation and strabismus [30,58,78]. One of these phenotypes, expression of autistic features, raises interest because it is estimated to affect more than 1 in 200 individuals in the United States. Since autistic features are common in ADSL deficiency [17], occurring in about 40 to 50% of cases, the study of ADSL deficiency, a single gene disorder, may reveal significant information regarding biomarkers for or possible pathogenic mechanisms leading to Autism Spectrum Disorder (ASD).

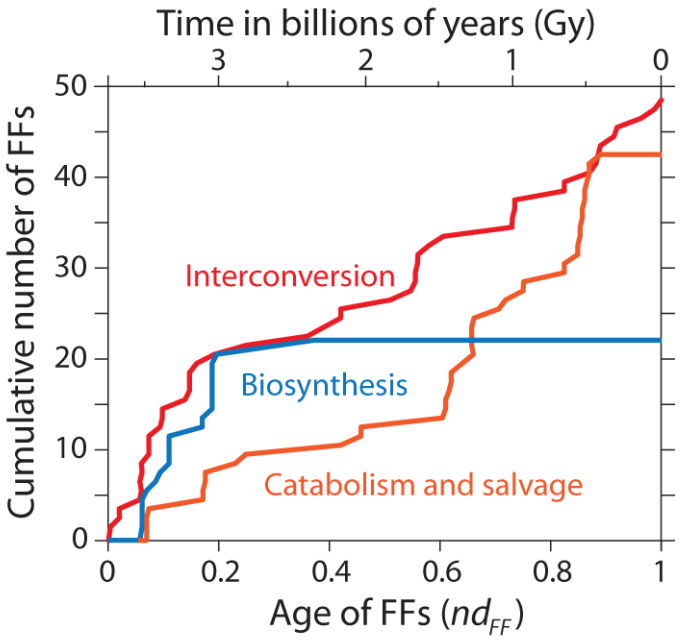


Figure 1.3: Early development of biochemical pathways

The leading hypothesis for the manifestation of ADSL deficiency is by elevated levels of ADSL substrates, 5-aminoimidazole-4-(N succinylaminoimidazolecarboxamide ribonucleotide) (SAICAR) and adenylosuccinate (SAMP), intracellularly. As SAICAR and SAMP accumulate, they are consequently dephosphorylated into succinylaminoimidazolecarboxamide riboside (SAICAr) and succinyladenosine (S-Ado) respectively and are transported out of the cell into urine, plasma and cerebrospinal fluid (CSF). The expected concentration of these dephosphorylated derivatives for a normal individual is $< 1\mu$M, or below detectable limits. Those with the deficiency accumulate a concentration of 150-500 μM [78]. The initial diagnosis is made by detection of SAICAr in urine using the Bratton-Marshall reaction for diazotizable amines followed by confirmatory analysis by other methods such as high performance liquid

chromatography (HPLC) with ultraviolet detection or liquid chromatography-tandem mass spectrometry (LC-MS/MS).

Three distinct phenotypic groups of ADSL deficiency have been distinguished: 1) neonatal fatal form, 2) severe childhood form (Type I), 3) moderate or mild form (Type II). A summary of these groups can be found in Table 1.1. The neonatal fatal form exhibits prenatal hyperkinesis, pulmonary hypoplasia and prenatal growth restriction, resulting in fatal neonatal encephalopathy [98]. The severe childhood form displays within the first months of life, encompassing the whole spectrum of disease symptoms and resulting in severe psychomotor retardation, developmental arrest, coma vigil, and often early death [98]. The moderate or mild form displays within the first years of life either with psychomotor retardation or hypotonia [98]. The first correlation found between biochemical diagnosis and phenotype expression has been the ratio of S-Ado to SAICAr. Patients with neonatal fatal form have a S-Ado:SAICAr ratio of <1. Patients with Type I have a S-Ado:SAICAr ratio of ~ 1. Patients with Type II have a S-Ado:SAICAr ratio of >1.

Currently, the mechanism by which ADSL deficiency directly causes the phenotypes of the disorder is not known. Consequently there is no effective treatment for ADSL deficiency either. Heterogeneity in disease phenotype has made it harder to attribute one single pathogenic mechanism behind the disease. Studies have not reported significant irregularities in most aspects of purine synthesis. The rate of synthesis is roughly normal as are growth rates of fibroblasts [86, 87]. Regular concentrations of purine nucleotides and ATP

Table 1.1: Phenotypic Groups, and Characteristics

Phenotypic Group	Phenotype Level	S-Ado:SAICAr	Symptoms
Neonatal	Fatal Neonatal	< 1	In utero death
Type I	Severe	~ 1	PMR, developmental arrest, often death
Type 2	Moderate	> 1	PMR, hypotonia, expression of Autistic features

are also reported [86, 87]. These findings seem to suggest that ADSL deficiency may not lead to the disorder by disruption of de novo purine synthesis. However, a significant weakness of many of these studies is that they do not adequately control for the presence of purines in the environment and purine salvage pathways, nor do they take into account the possibilities that de novo purine synthesis may be altered during particular stages of development or in particular tissues. Furthermore, the increased level of metabolites of ADSL substrates in urine, plasma and CSF points towards the critical role the enzyme defects may play in disease pathogenesis. Perhaps the most convincing evidence of enzyme defects involved in disease onset is the identification of few alterations in mutants of ADSL that are directly related to the severities of the symptom. These observations offer few possible scenarios about the disease mechanism. The resulting hypotheses regarding the pathogenesis include: 1) toxicity of high levels of SAICAR, AMPS, or their metabolites; 2) lack of ability to synthesize sufficient purines using the de novo purine biosynthetic pathway; 3) lack of a completely functional purine nucleotide cycle in muscle

and brain; 4) improper regulation of energy metabolism; or 5) inhibiting formation of a fully functional multi-enzyme purinosome complex [2, 57, 78, 98]. We hypothesize that there are multiple mechanisms of pathogenesis of ADSL deficiency including diminution of enzyme activity, instability of the functional ADSL tetramer, or aggregation of ADSL. The significance of this is that development of an effective therapy may depend upon which mechanism is active and the mechanism will depend on the patient's ADSL mutation.

Although metabolites of the substrates SAICAR and SAMP are what are usually detected in patients, it has been shown that SAICAR is toxic when it accumulates in yeast cells [34, 64, 71, 72, 80]. In a recent study [55], the affects of defects in the *de novo* purine pathway on chronological life span (CLS) in yeast cells was tested. The results showed when any one protein in the pathway was absent, with the exception of ADSL, the CLS of the yeast was expanded. Cells missing ADSL could not survive even in the presence of adenine. This suggests ADSL is an essential enzyme at least for yeast. This study further demonstrates the importance of ADSL to be functional in the pathway and the toxic consequences of accumulating SAICAR.

Substrate accumulation can happen either by simple loss of enzymatic activity due to subtle changes in structure of ADSL or by drastic changes in the stability of the enzyme itself. Decrease in stability will result in fewer functional protein molecules or an increase in unfolded/misfolded proteins that may promote aggregation and/or degradation. If ADSL mutants have a higher propensity to aggregate, it is possible these aggregates could be

toxic. This is in accordance with the idea of mutations in proteins that can cause disease by Loss of Function or by Gain in Toxic Function [5, 33, 52]. The role of protein mis-aggregation/misfolding in human disorders is widely recognized, and a number of inborn errors of metabolism may fall into this class of disorder, for example phenylketonuria (PKU), a classic inborn error of metabolism [8, 46, 81, 94, 95]. Recent studies have also alluded to a novel idea of formation of the purinosome, a multi-enzyme complex, crucial in the de novo purine synthesis pathway [2, 89]. It is possible that the mutations in ADSL alter the purine metabolic pathway by inhibiting formation of a fully functional multi-enzyme purinosome complex leading to altered purine synthesis during selected phases of the cell cycle or during times of particularly high purine demand. There could also be a more profound effect of ADSL deficiency during fetal development when excess purines are unavailable and the embryo is likely to require the de novo pathway.

The unusual behavior of different mutants of human ADSL and their direct relevance to disease pathogenesis requires careful investigation. The proposed research is to study the enzyme activity, conformation, stability, and aggregation propensity of wild-type human ADSL as well as 4 mutants associated with ADSL deficiency (E80D, D87E, R303C, R426H) and A291V, a unique mutation that eliminates all detectable activity. These disease associated mutants were chosen because they represent missense mutations that cause different phenotypes on the ADSL Deficiency spectrum. Thus, we believe this study of the structure-function-stability relationship of the wild type (WT) and disease mutants of ADSL will advance our understanding of the mechanistic aspects of

ADSL deficiency and its relation to disease pathogenesis and therefore suggest new approaches towards therapy.

1.2 Disease Associated Mutations

Single site mutations in the amino acid sequence can be useful in evolution for protein structure or could be detrimental and cause disease. Whether a mutation is favorable or unfavorable depends on chemical factors such as composition, polarity, and molecular volume, which have been quantified by the Grantham Score (the higher the number the greater the difference). The work in this thesis will focus on two unfavorable missense mutations that are the standards for the extreme cases of ADSL deficiency: a substitution of a histidine for an arginine at amino acid position 426 (R426H) and a substitution of a cysteine for an arginine at amino acid position 303 (R303C). The R426H mutation is the most common occurring in 30% (see Table 1.3) of reported cases of ADSL deficiency and when homozygous is characterized by severe symptoms. R303C has been identified as a homozygous mutation in two patients with the mildest phenotypes and highest S-Ado/SAICAr ratios reported [78, 98]. Two other mutations of interest are a substitution of an aspartate for a glutamate at position 80 (E80D), and substitution of a glutamate for an aspartate at position 87 (D87E). These mutations are found in one heterozygous patient. These are of particular interest because the patient was initially diagnosed with Autism [76]. We have also investigated a substitution of a valine for an alanine at amino acid position 291 (A291V). This is a mutation that causes complete loss of activity and could provide informa-

Table 1.2: Grantham Score and Frequency in Disease and Evolution

ΔA.A.	Grantham Score	Frequency in Disease (%)	Frequency in Evolution (%)
R → H	29	2.5 - 12.5	*Has not been recorded
R → C	180	2.5 - 12.5	12.5 - 100
D → E	45	12.5 - 100	2.5 - 12.5
E → D	45	12.5 - 100	2.5 - 12.5
A → V	64	2.5 - 12.5	2.5 - 12.5
**G → V	121	12.5 - 100	12.5 - 100

*H → R substitution is in the top 2.5% of substitutions in evolution
**G → V is the mutation that causes sickle cell anemia

Grantham, 1974 [31]; Wu et al, 2007 [93]

Table 1.3: Number of Cases Reported of Mutations in ADSL Deficiency

Mutation	Homozygote	Heterozygote	Total	%) Reported
E80D*	-	1	1	1.6
D87E*	-	1	1	1.6
A291V	-	-	-	-
R303C	2	-	-	3.3
R426H	12	6	18	30

*Found in same patient

tion regarding a highly conserved loop in the enzyme structure. The general location of these mutations on the monomer is depicted in Figure 1.4.

1.2.1 R426H (Arginine to Histidine)

The amino acid R426 is located at the surface of the substrate channel and is important for mediating the interaction of Q409 and D422. When histidine replaces the arginine, it disrupts the arginine-mediated interaction and causes instability [98]. The mutation of arginine to histidine has a Grantham score of

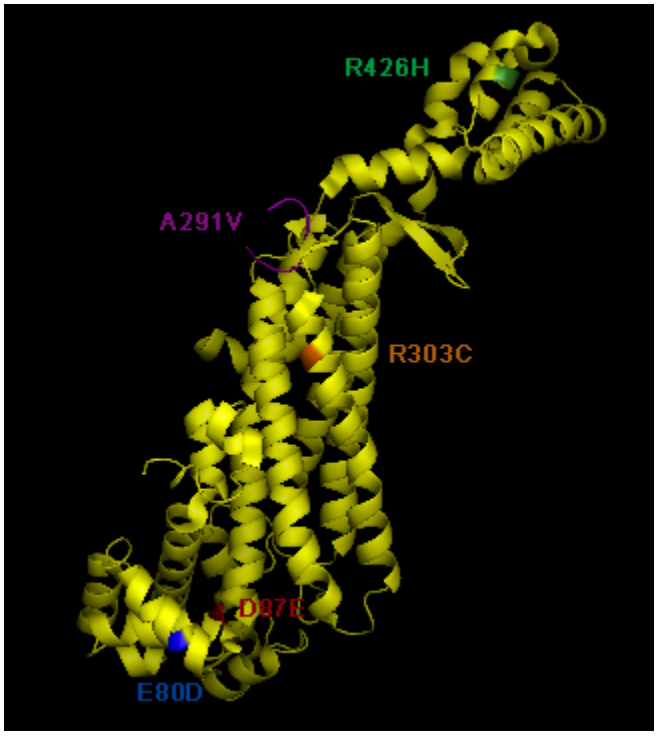
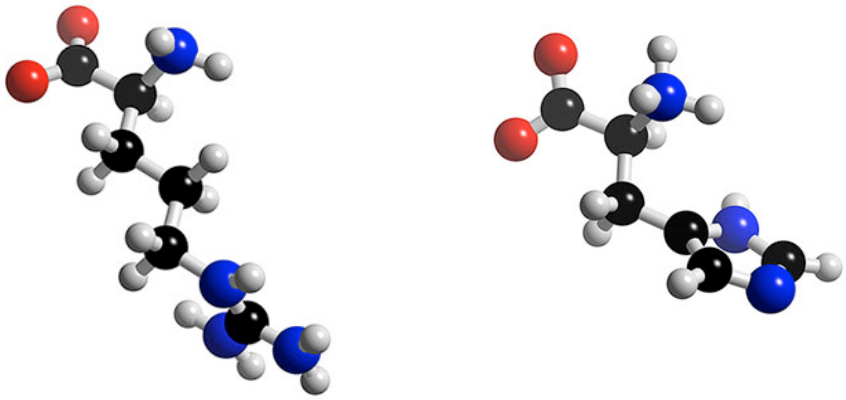


Figure 1.4: Human ADSL monomer with labeled clinically associated mutations in this study

29, meaning the two amino acids are very similar on a chemical level. However, this particular mutation is frequently associated with human genetic disease (2.5-12.5%) and has never been recorded in evolution (Figure 1.2) [31, 93]. Interestingly, the histidine to arginine mutation is one of the most frequent amino acid substitutions in evolution [93]. This may be due to the difference in pKa values. Arginine has a pKa value of 12.5 compared to histidine pKa value of 6.5. This means that histidine is more susceptible to losing its positive charge when small shifts in physiological pH levels occur.



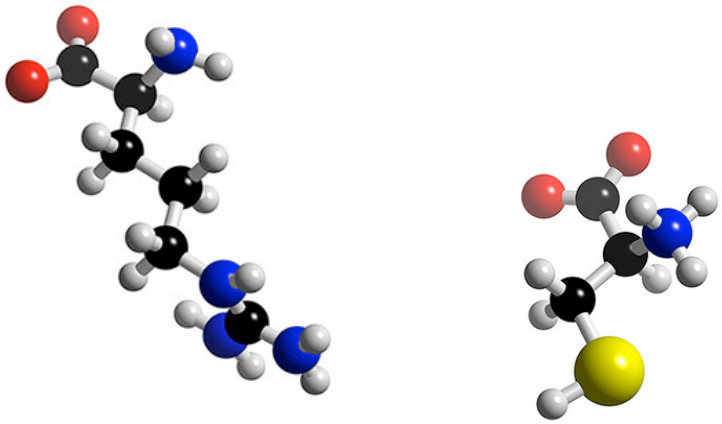
(a) (b)

Figure 1.5: Amino acid structures for arginine (a) and histidine (b)

1.2.2 R303C (Arginine to Cysteine)

The R303 amino acid is located directly in the active site. Its function in the active site is to use its positive charge for binding the phosphate group of the substrate. Cysteine is a non-polar, hydrophobic, residue that usually plays an important structural role. Because of these differences the Grantham score for arginine and cysteine is 180. If a cysteine is substituted for arginine, the positive charge is lost and the ability to bind the phosphate group of the substrates is lost. Interestingly, this mutation shows a more severe loss in activity with SAMP than with SAICAR, although the extent of disproportionality varies from study to study, perhaps reflecting differences in how the enzyme assays were carried out. Some investigators were unable to detect activity with SAMP in extracts from fibroblasts [53], while others found 3% of normal activity (8). In other cases using different recombinant ADSL constructs, 7% and 18% of wild-type (WT) activity was observed with SAMP and

44% with SAICAR [53, 67]. The reason for the non-parallel loss of activity is not yet understood.

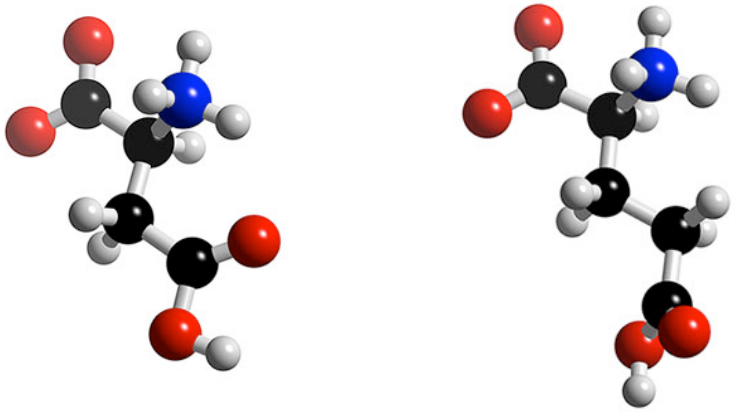


(a) (b)

Figure 1.6: Amino acid structures for arginine (a) and cysteine (b)

1.2.3 E80D (Glutamic Acid to Aspartic Acid) and D87E (Aspartic Acid to Glutamic Acid)

The E80D and D87E mutations, which these laboratories collaborated in identifying, lead to relatively mild disease [76]. Both mutations are highly conserved on the basis of evolutionary considerations for mammals and the locations of the two mutations relative to the active site histidine residue at position 86. This is surprising because the Grantham Score (45) and the frequency of the D to E and E to D mutations in evolution are high (Table 1.2). This would lead one to think the two are very interchangeable, but that is not the case for residue D87.

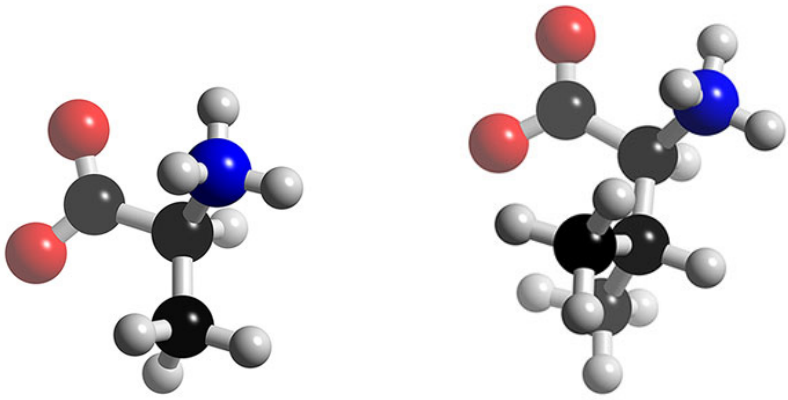


(a) (b)

Figure 1.7: Amino acid structures for aspartic acid (a) and glumatic acid (b)

1.2.4 A291V (Alanine to Valine)

All of the solved crystal structures for ADSL have a non-solvable region from amino acid 284 to 292. It is believed this region is a flexible loop that encloses the active site upon substrate binding. The mutation A291V was recently identified in our laboratory as the mutation in the ADSL deficient AdeI mutation of Chinese hamster ovary (CHO) cells. This mutation results in loss of detectable enzyme activity. Alanine and Valine are both non-polar amino acids that differ slightly in their structure with valine having two extra carbon bonds leading to a Grantham score of 64. Yet, the frequency of the substitution is the same for causing disease and for evolutionary purposes. These findings suggest that our current understanding of the functions of these amino acid residues in ADSL is incomplete.



(a) (b)

Figure 1.8: Amino acid structures for alanine (a) and valine (b)

1.2.5 Comments on Using *B. Subtilis* for Studying Disease Mutations

The use of Bacillus subtilis ADSL in particular as a model to study properties of mutant enzymes in some cases involved modification of additional amino acid residues in the *B. subtilis* ADSL to more closely resemble human ADSL [11, 76]. However, attempts to make use of a *B. subtilis* model system to replicate the ADSL mutations have proven difficult because *B. subtilis* and H. sapiens only have a sequence identity of 30% and a similarity of 63% (Table 1.4). Specifically, to replicate the R303C mutation, the sequence diversity of *B. subtilis* and H. sapiens results in multiple alterations within the active site to include but not limited to an asparagine residue in place of the corresponding R303 in human ADSL as well as an arginine residue in place of human corresponding T354. Not surprisingly, the removal of *B. subtilis*'s asparagine at the corresponding human ADSL R303 position for cysteine did not gener-

Table 1.4: Sequence Comparison of Different Species with *H. Sapien*

Type	Identity (%)	Similarity (%)
E. coli	23.3	57.7
T. Maritima	24.6	57.2
B. subtilis	30.3	62.6
M. Musulus	94.0	98.1

www.dbi.ac.uk/Tools/psa/lalign

ate a disproportional reduction in catalytic ability [76]. This observation as well as others suggest that the study of the bacterial enzyme, while useful for understanding some of the basic features of ADSL, may not be as useful for understanding the effects of various disease causing mutations on the human enzyme. On the basis of studies with the homologous *B. subtilis* enzyme, the E80D mutation severely affects the activity of ADSL, while the D87E mutation has a more mild effect on enzyme activity. Recent vitro studies of the human E80D and D87E ADSL mutants expressed in *E. coli* suggest that the homomeric D87E or the E80D mutations result in relatively active and stable enzyme and that the heteromeric form is also quite stable and active [98]. The conclusion regarding the mutations in the *B. subtilis* enzyme is that the mutation equivalent to E80D is more deleterious than the mutation equivalent to the D87E mutation in humans.

The use of *B. subtilis* was often employed prior to having a stable human ADSL construct to perform *in vitro* studies on fundamental properties such as function and stability. These studies contributed to initial understanding of ADSL, but since the development of a stable ADSL construct, it has been

seen that studies characterizing clinical mutations using *B. subtilis* do not match with the human protein. This is most likely due to the difference in sequence identity and similarity where a humanized *B. subtilis* would have to be constructed using multiple mutations besides the clinically associated mutation. These results emphasize the importance of using the human stain of any enzyme when attempting to characterize how mutations affect inherent properties of the enzyme.

1.3 Introduction to Enzymes

The discovery of enzymes began with Louis Pasteur in the 1850s when he observed that fermentation of sugar into alcohol by yeast is catalyzed by what he called "ferments". It was initially thought that these ferments were inseparable from the cell structure until Edward Buchner showed yeast cell extracts could also ferment sugar in 1897. This proved that these "ferments" were molecules that continued to have function outside of the cellular environment. Frederick W. K*ü*hne later gave the name enzymes to these molecules responsible for the fermentation of sugar. The isolation and crystallization of one enzyme, urease, by James Sumner in 1926 determined the enzyme consisted of protein. He postulated all enzymes are proteins, a theory that wasn't accepted until after other known digestive enzymes such as trypsin and pepsin, were also proven to be proteins. Since them, thousands of enzymes have been purified and with the exception of a small group of catalytic RNA molecules, all enzymes are proteins.

1.4 How Enzymes Work

Most chemical reactions are not favorable to occur within the cellular environment. The reaction must overcome a large energy barrier to spontaneously occur as depicted by the black line in Figure 1.9. To circumvent this issue, enzymes provide a specific environment to make a reaction occur more rapidly by lowering the energy barrier to enhance the reaction rates. The enzyme provides free energy by creating a favorable environment for catalysis; the active site binds the substrate in an advantageous orientation for catalysis. Catalysis is still the rate limiting step since it has the highest energy barrier than either the enzyme substrate or enzyme product complexes which are reaction intermediates in the reaction pathway with a finite lifetime.

$\Delta G^{\ddagger}$ shows the activation energy for the forward and reverse reactions between the substrate ground state and the transition state. $\Delta G^{`\circ}$ is the biochemical standard free energy change. Reaction equilibria are linked to $\Delta G^{`\circ}$ and the reaction rate to $\Delta G^{\ddagger}$. From thermodynamics, the relationship between the equilibrium constant, K_d and $\Delta G^{`\circ}$ is $\Delta G^{`\circ}$ = -RTlnK_d where R is the gas constant and T is absolute temperature. From transition state theory an expression that relates the magnitude of a rate constant to the activation energy can be derived: $k = \frac{\mathbf{k}T}{h} e^{\frac{-\Delta G^{\ddagger}}{RT}}$. Since enzymes do not alter the position and direction of equilibrium, they do not change the K_d of the reaction.

Specificity of enzyme substrate pairs is a crucial advantage for the physiological context because it allows regulation of a myriad of cellular process

in the same space and time. Specificity results from the active site being both geometrically and chemically suited for binding a certain substrate and performing a certain type of chemical catalysis. Formation of these small interactions is accompanied by the release of small amounts of free energy that is used to lower the activation barrier of the reaction. These weak interactions provide a substantial driving force for enzyme catalysis and are optimized in the transition state.

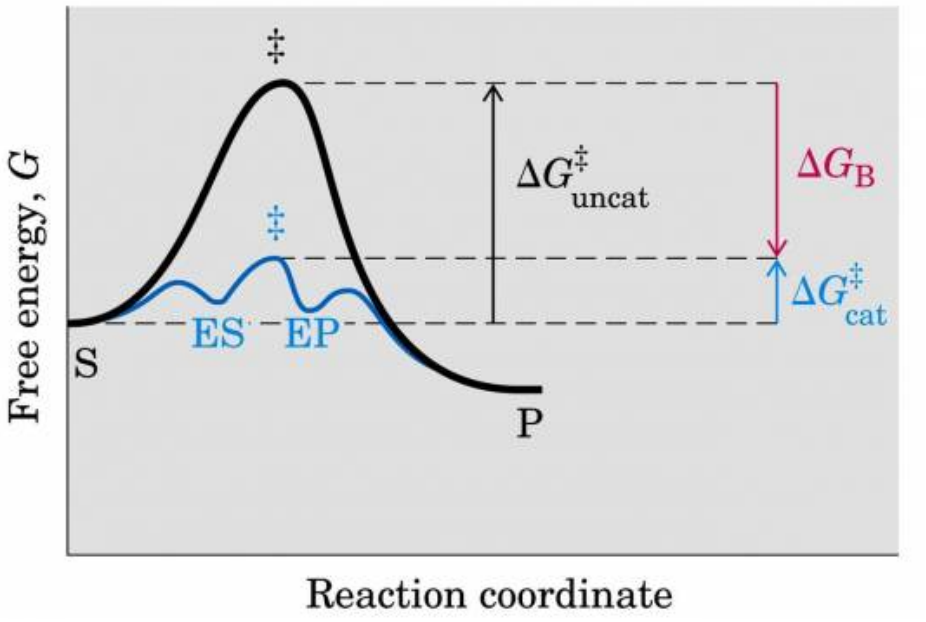


Figure 1.9: Free energy diagram of a chemical reaction with an enzyme. Figure from Principles of Biochemistry, 5th edition

1.5 ADSL Method of Catalysis

There are 5 main types of catalysis: substrate confinement, non-covalent stabilization of the transition state, covalent catalysis, metal-ion catalysis, and

general acid/base catalysis. ADSL uses the general acid/base catalysis where the transition state stabilization can sometimes be carried out by transferring a proton to the substrate or from it. When the transition state is negatively charged, protonation may stabilize it in a hydrophobic or negatively charged binding site. Protonation of the transition state requires an acid, whereas de-protonation requires a base. Enzymes often use the own chemical groups for such processes, mainly imidazole, and to a lesser extent carboxylate. In many cases, the transfer of hydrogen species occurs despite an impeding energy barrier. The transfer is made possible thanks to a quantum-mechanical phenomenon referred to as tunneling. The hydrogen atom tunnels through the barrier as opposed to crossing it using kinetic energy.

ADSL is a member of a metabolic enzyme superfamily known as the fumarase enzymes. Members of this family catalyze reactions with fumarate as one of the products. Though there are differences among the catalytic mechanisms, there is a highly conserved sequence found by all members of this family, corresponding to residues 284 - 303 in human ADSL, that is in close proximity of the active site [11, 44]. The ADSL catalyzed conversion of SAMP to AMP and fumarate proceeds via a uni-bi mechanism, where fumarate is released from the active site prior to AMP [10, 75]. The proposed chemical mechanism is β-elimination where one of the hydrogens attached to the β carbon of the succinyl moiety is abstracted by a general base; a double bond forms between the β and α carbon of the succinyl moiety; the leaving group is protonated at either the N1 of N6 position by a general acid, and the bond between the α carbon an N6 of the adenosine ring system of the leaving group is cleaved [75].

Site directed mutagenesis of *Bacillus subtilis* ADSL has revealed three histidines as critical residues of the intersubunit catalytic site, His^{68}, His^{89}, and His^{141} [11,48–50], and it has been proposed that His^{68} and His^{141} form the acid base pair [49]. Since the catalytic mechanism of the fumarase family is highly conserved, the human analogues to His^{68} and His^{141} are also proposed to play the same role in catalysis.

1.6 Review of Elementary Reaction Kinetics

1.6.1 First Order Reaction

$$[A] \xrightarrow{k_1} [P]$$

$$\frac{d[P]}{dt} = -\frac{d[A]}{dt} = k_1[A]^1 = k_1[A]$$

$$-\frac{d[A]}{dt} = k_1[A]$$

$$-\frac{d[A]}{[A]} = k_1 \mathrm{dt}$$

$$-\int_{A_o}^{A} \frac{d[A]}{[A]} = \int_{0}^{t} k_1 \mathrm{dt}$$

$$\mathrm{Ln}[A] \mid_{A_o}^{A} = k_1 \mathrm{dt} \mid_{0}^{t}$$

$$\mathrm{Ln}[A] - \mathrm{Ln}[A_o] = k_1 \mathrm{t}$$

$$\mathrm{Ln}[\frac{A}{[A_o]}] = k_1 \mathrm{t}$$

$$[A] = [A_o] \mathrm{e}^{k_1 t}$$

1.6.2 Pseudo First Order Reaction

$$[A] + [\mathrm{B}] \xrightarrow{k_1} [P]$$

$$\frac{d[P]}{dt} = -\frac{d[A]}{dt} = -\frac{d[B]}{dt} = k_1[A][\mathrm{B}] = k_1'[A]$$

where $= k_1[B] = k_1'$ if $[B] >> [A]$ ($[A]$ would be the limiting concentration)

1.6.3 Second Order Reaction

$$[A] + [A] \xrightarrow{k_1} [P]$$

$$\frac{d[P]}{dt} = -\frac{d[A]}{dt} = = k_1[A][A] = k_1[A]^2$$

$$-\frac{d[A]}{dt} = k_1[A]^2$$

$$\frac{d[A]}{[A]^2} = -k_1\mathrm{dt}$$

$$\int_{A_o}^{A} \frac{d[A]}{[A]^2} = -\int_{0}^{t} k_1\mathrm{dt}$$

$$-\frac{1}{A} \Big|_{A_o}^{A} = -k_1\mathrm{dt} \Big|_0^t$$

$$-\frac{1}{A} - -\frac{1}{A_0} = -k_1\mathrm{t}$$

$$\frac{1}{A} = \frac{1}{A_0} + k_1\mathrm{t}$$

1.6.4 Reversible First Order Reactions

$$[A] \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} [P]$$

$$\mathrm{K}_{eq} = \frac{[P]eq}{[A]_{eq}} = \frac{k_1}{k_{-1}}$$

$$[A]_o - [A]_t = [P]_o + [P]_t$$

Where o and t represent concentration of A or P at time 0 and any time t. Assuming there is no initial concentration of P

$$[P]_t = [A]_o \text{ - } [A]_t$$

At $t = \infty$, the system has reached equilibrium

$$[P]_{eq} = [A]_o \text{ - } [A]_{eq}$$

Using the definition of K_{eq}

$$[P]_{eq} = \frac{[A]_{eq} \cdot k_1}{k_{-1}} = [A]_o \text{ - } [A]_{eq}$$

Solving for $[A]_{eq}$

$$[A]_{eq} = \frac{[A]_o \cdot k_{-1}}{k_{-1} + k_1}$$

Using the principle of mass action to write the differential equation for the change in $[A]$ at time = t'

$$\frac{d[P]_{t'}}{dt'} = -\frac{d[A]_{t'}}{dt'} = k_1[A]_{t'} \text{ - } k_{-1}[P]_{t'}$$

$$-\frac{d[A]_{t'}}{dt'} = k_1[A]_{t'} \text{ - } k_{-1}([A]_o \text{ - } [A]_{t'})$$

$$-\frac{d[A]_{t'}}{dt'} = (k_1 + k_{-1})[A]_{t'} \text{ - } k_{-1}[A]_o$$

$$-\frac{d[A]_{t'}}{dt'} = \frac{k_{-1}[A]_o}{[A]_{eq}}[A]_{t'} \text{ - } k_{-1}[A]_o$$

$$\frac{d[A]_{t'}}{dt'} = \frac{k_{-1}[A]_o}{[A]_{eq}}(A_{eq} \text{ - } A_{t'})$$

$$\frac{d[A]_{t'}}{dt'} = \frac{(k_{-1}+k_1)[A]_{eq}}{[A]_{eq}}(A_{eq} \text{ - } A_{t'})$$

$$\int_0^{[A]_t} \frac{d[A]_{t'}}{[A]_{eq} - [A]_{t'}} = \int_0^t (k_{-1} + k_1)\mathrm{d}t'$$

Which results in the final solution

$$\ln\left(\frac{[A]_o - [A]_{eq}}{[A]_t - [A]_{eq}}\right) = (k_{-1} + k_1)\mathrm{t}$$

1.7 Introduction to Enzyme Kinetics

1.7.1 Rate Law

The rate of a process depends in some way on the concentrations or amounts involved. Substances can either change concentration with time (reactant, products, intermediates) or stay constant (catalysts, intermediates (steady-state), solvent. The concentration of the catalyst will not change during a single experiment, but may change from one experiment to the next which may alter the rates of reactions. The general enzyme reaction is given by:

$$[E] + [S] \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} [ES] \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} [EP] \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} [E] + [P]$$

Assuming that $[ES]$ breaks down directly to $[E] + [P]$

$$[E] + [S] \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} [ES] \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} [E] + [P]$$

Assuming the reverse reaction $[E] + [P] \xrightarrow{k_{-3}} [ES]$ is negligible

$$[E] + [S] \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} [ES] \xrightarrow{k_{cat}} [E] + [P] \tag{1.7.1}$$

The formation of $[ES]$ is a second order process and the breakdown of $[ES]$ is a first order process. The units for the rate constants are: $k_{-1} = k_{cat} = s^{-1}$ and $k_1 = M^{-1}s^{-1}$

1.7.2 Michaelis-Menten Kinetics Derivation

$$\frac{d[ES]}{dt} = k_1 \cdot [E][S] - k_{-1} \cdot [ES] - k_{cat} \cdot [ES]$$

Steady-state approximation; $\frac{d[ES]}{dt} = 0$

$$k_1 \cdot [E][S] = k_{-1} \cdot [ES] + k_{cat} \cdot [ES]$$

$$[\mathrm{E}_{total}] = [E] + [ES] \; ; \; [E] = [\mathrm{E}_{total}] - [ES]$$

$$k_1 \cdot ([\mathrm{E}_{total}] - [ES])[S] = k_{-1} \cdot [ES] + k_{cat} \cdot [ES]$$

$$k_1 \cdot [\mathrm{E}_{total}] \cdot [S] = k_1 \cdot [ES] \cdot [S] + k_{-1} \cdot [ES] + k_{cat} \cdot [ES]$$

$$[ES] = \frac{[E_{total}] \cdot [S]}{[S] + \frac{k_{-1} + k_{cat}}{k_1}}$$

$$K_m = \frac{k_{-1} + k_{cat}}{k_1}$$

Note that K_m is not a binding constant that measures the strength of binding between the enzyme and substrate. Its values include the affinity of the substrate for enzyme, but also the rate at which the substrate bound to the enzyme is converted to product. The Michaelis-Menten model is too simple for many purposes.

$$[ES] = \frac{[E_{total}] \cdot [S]}{[S] + K_m}$$

$$k_{cat} \cdot \; [ES] = \frac{k_{cat} \cdot [E_{total}] \cdot [S]}{[S] + K_m}$$

$$V_{max} = k_{cat} \cdot [\mathrm{E}_{total}]$$

$$\nu = \frac{V_{max} \cdot [S]}{[S] + K_m}$$

$$\text{At } [S] = K_m \; ; \; \nu = \frac{V_{max}}{2}$$

Assumptions of Michaelis-Menten

• The production of product is linear with time during the time interval used.

• The concentration of substrate vastly exceeds the concentration of enzyme. This means that the free concentration of substrate is very close to the concentration added, and that substrate concentration is constant throughout the assay.

• A single enzyme forms the product.

• There is negligible spontaneous creation of product without enzyme.

• No cooperativity. Binding of substrate to one enzyme binding site doesn't influence the affinity or activity of an adjacent site.

• Neither substrate nor product acts an an allosteric modulator to alter the enzyme velocity.

1.7.3 Allosteric Enzymes

Enzymes comprised of multi-subunit proteins and that have multiple active sites are commonly allosteric, where the conformation of one active site is altered due to the binding of the substrate in another active site. These allosteric interactions can be positive or negative. A positive (negative) cooperativity means a favorable (unfavorable) conformation change that increases (decreases) the probability of substrate binding. Experimentally this is observed when the plots of ν vs. $[S]$ are no longer hyperbolic (characteristic of Michaelis-Menten Kinetics), but sigmoidal, as seen in Figure 1.10. The initial velocity equation is given by the Hill Equation:

$$\nu = \frac{V_{max} \cdot [S]^n}{[S]^n + K_{0.5}^n}$$

Where n is the Hill coefficient that determines the degree of cooperativity. If n is greater than (less than) one, it denotes positive (negative) cooperativity. If n is equal to one, the Hill Equation reduces back to the Michaelis-Menten Equation.

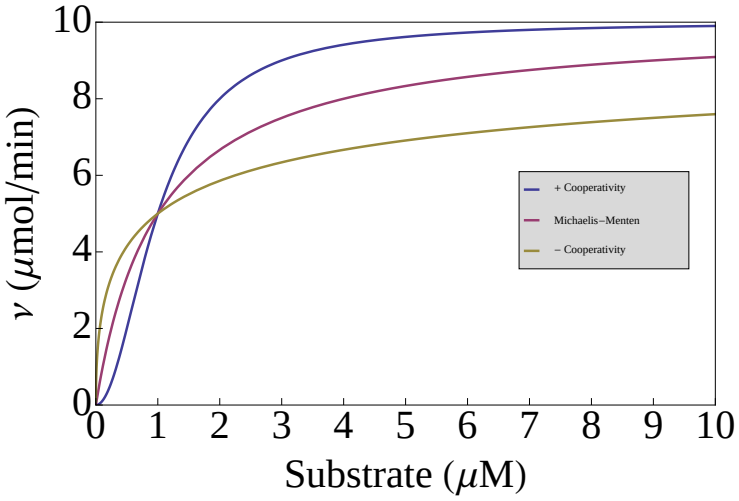


Figure 1.10: Kinetic curves of positive and negative cooperativity and Michaelis-Menten Kinetics.

1.7.4 Rapid vs. Steady-State Equilibrium

In all of the cases above, steady-state equilibrium is assumed. Steady-state equilibrium presumes that $\frac{dES}{dt} = 0$ when $[S] >>$ [Eo]. If excess $[S]$ is added to $[E]$ the $[ES]$ complex is immediately formed and the substrate is transformed to product. The importance of steady-state kinetics is within milliseconds of starting the reaction, the $[ES]$ complex concentration will not change significantly during the initial part of the reaction. Another characteristic of

steady-state kinetics is K_m, which is a complex function of all rate constants in the reaction:

$$K_m = \frac{k_{-1}+k_{cat}}{k_1}$$

As stated earlier, K_m is not a binding constant that measures the strength of binding between the enzyme and substrate. However, if k_{cat} is the limiting step in the reaction, so $k_{cat} \ll k_{-1}$, K_m reduces to a dissociation constant. This is the condition of rapid equilibrium where after the $[ES]$ complex forms, it dissociates back to $[E]$ + $[S]$, before catalysis takes place. Therefore, most of the bound $[S]$ will dissociate and not be converted to $[P]$. The resulting V_{max} is the same, $V_{max} = k_{cat}[ES]$. The assumption that $\frac{dES}{dt} = 0$ still holds true for rapid equilibrium as well. As soon as $[S]$ dissociates or is catalyzed another $[S]$ is immediately bound.

1.8 Introduction to Protein Structure

The three dimensional structure of proteins was first solved in 1960 by the British biochemist John Kendrew using X-Ray Diffraction to photograph myoglobin with 0.2 nm resolution. His colleague Max Perutz later determined the structure of hemoglobin. The two of them were awarded the 1962 Nobel Prize in Chemistry for their achievement. Since then, thousands of protein structures have been resolved and made widely available on the Internet. Proteins are very compact structures that are organized in a very specific way with an apparent hierarchy. The levels of hierarchy are primary, secondary, tertiary, and quaternary structure (Figure 1.11).

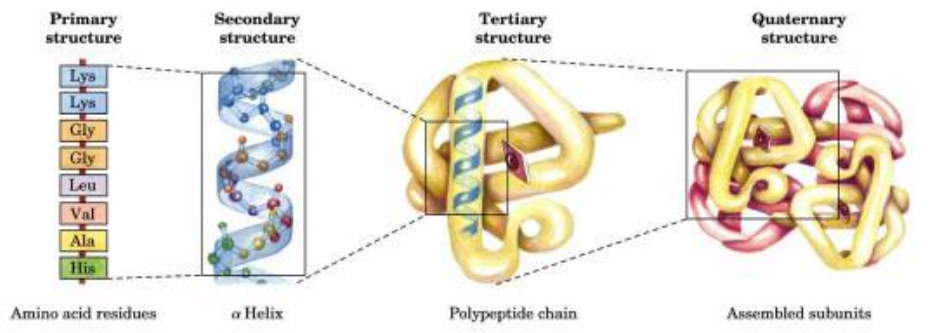


Figure 1.11: Higherarchy of a predominantly α-helical protein

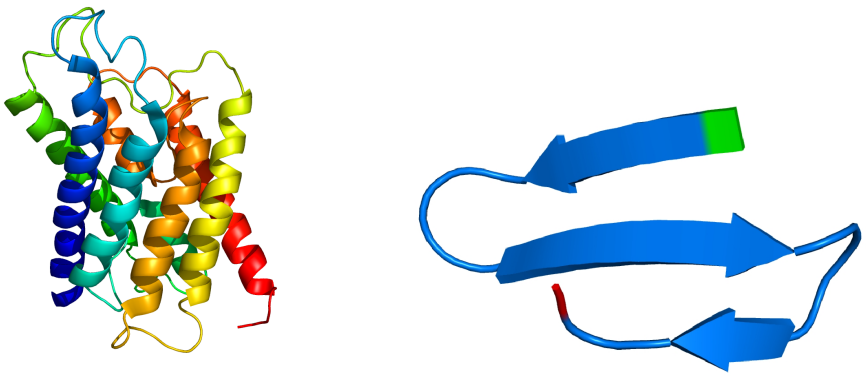
1.8.1 Primary Structure

A protein's identity is based on its amino acid sequence (Figure 1.11). The exact ordering of the amino acids is important because amino acids have different physio-chemical properties. These properties interact with each other in a specific manner to achieve the final, folded, function structure. Amino acids generally fall into one of four groups: nonpolar, polar-neutral, electrically charged, or aromatic. Nonpolar residues are hydrophobic and are usually found in the core of the protein where they interact with one another to allow for a tight compact structure. Polar amino acids are hydrophilic and are often used for interaction with other molecules, catalysis, and are important for solubility. Like polar amino acids, electrically charged amino acids are also found on the surface because of their ability to interact with molecules and ligands. However, since they are electrically charged, they can be either protonated or de-protonated, meaning they can behave as either acidic or basic amino acids depending on the pH of the cellular environment. Aromatic amino acids can assume properties of both polar and nonpolar groups because of their

large side-chains. Their ability to absorb UV radiation is also a characteristic commonly taken advantage of for fluorescence and UV assays.

1.8.2 Secondary Structure

As the amino acids start to interact via non-covalent interactions, they begin to fold into simple local structures, or also called secondary structure (Figure 1.11). We can group secondary structures into three basic groups: α-helix, β-sheet, and turns or loops. The α-helix is a very compact structure where side chains of the amino acids face the periphery while intermolecular hydrogen bonds or van der Waals contacts stabilize the core (Figure 1.12a). The β-sheet secondary structure consists of a group of β-strands that arrange themselves along each other in either parallel or anti-parallel directions connected by a short loop called a β-turn (Figure 1.12b). These are much more extended confirmations and not as compact as α-helix. Loops and turns are segments that lack a particular structure in order for flexibility.



(a) α-helix (b) β-sheet

Figure 1.12: Secondary structures

1.8.3 Tertiary Structure

As the local structures start to take shape, the entire sequence folds into a complete overall unit called the tertiary structure (Figure 1.11). Though all proteins consist of the same basic secondary structures, the overall fold of the unit is quite diverse in order to achieve a specific function. Usually the folded protein will fold into a globular or spherical shape if complex functions are necessary or it will fold into a fibrous, elongated shape to carry out simpler functions. In order to be functional in a crowded cellular environment, tertiary structures need to be compact, soluble, and properly form their active or binding sites. As the final folding structure is made, hydrophobic patches again form the core of the structure. This stabilizes the structure by a myriad of weak interactions. This is crucial because it keeps the structure stable while still allowing dynamics crucial for functionality. On the surface of the tertiary structure are usually hydrophilic secondary structures to regulate interactions. Local secondary structures can fold into separate regions characterized by its arrangement and function. These semi-independent units are known as domains.

1.8.4 Quaternary Structure

In order to acquire full functionality, it is often required for multiple chains to interact. The spatial arrangement of the chains is referred as the quaternary structure (Figure 1.11). Each chain is then called a subunit. These quaternary structures can be a construct of identical or different chains and are usually two to six chains in size but can be as big as 60 polypeptide chains. The spatial

arrangement of the subunits is often symmetrical; hence quaternary structures tend to have an even number of subunits. Proteins evolved into adopting quaternary structures because of the advantages they offer. It is common that active sties are constructed by the contribution of side chains from different subunits. This can help with the regulation of substrate specificity as well as versatility. Biochemical pathways are multi-step processes that involve more than one enzyme. For better efficiency of the pathway, enzymes evolved and create binding sites from several subunits to accommodate different substrates in the pathway. Multi-unit complexes may be more stable in the cellular environment due to the contacts at the interface of the subunits restricting motion and making the complex less accessible to proteolytic enzymes.

1.8.5 ADSL Landscape

The monomer amino acid sequence for ADSL is 484 amino acids long (Figure 1.13). The amino acids interact to form secondary structures that are depicted by the cartoons above the letters. The pink sinusoidal wave, yellow arrows, and purple horseshoes correspond to α-helix, β-sheet, and loops secondary structure respectively. The secondary structure is predominanty α-helix (62%) with a minor amount of β-sheet (4%). With the large percentage of α-helix, the ADSL monomer is more elongated than globular (Figure 1.14). Each monomer contains 3 domains: Domain 1 consists of residues 1 - 116 and is compact; Domain 2 consists of residues 117-376 and is elongated; Domain 3 consists of residues 377 - 484 and is compact. The monomeric weight of ADSL with a six-histidine amino terminal tag is approximately 57 kDa [47].

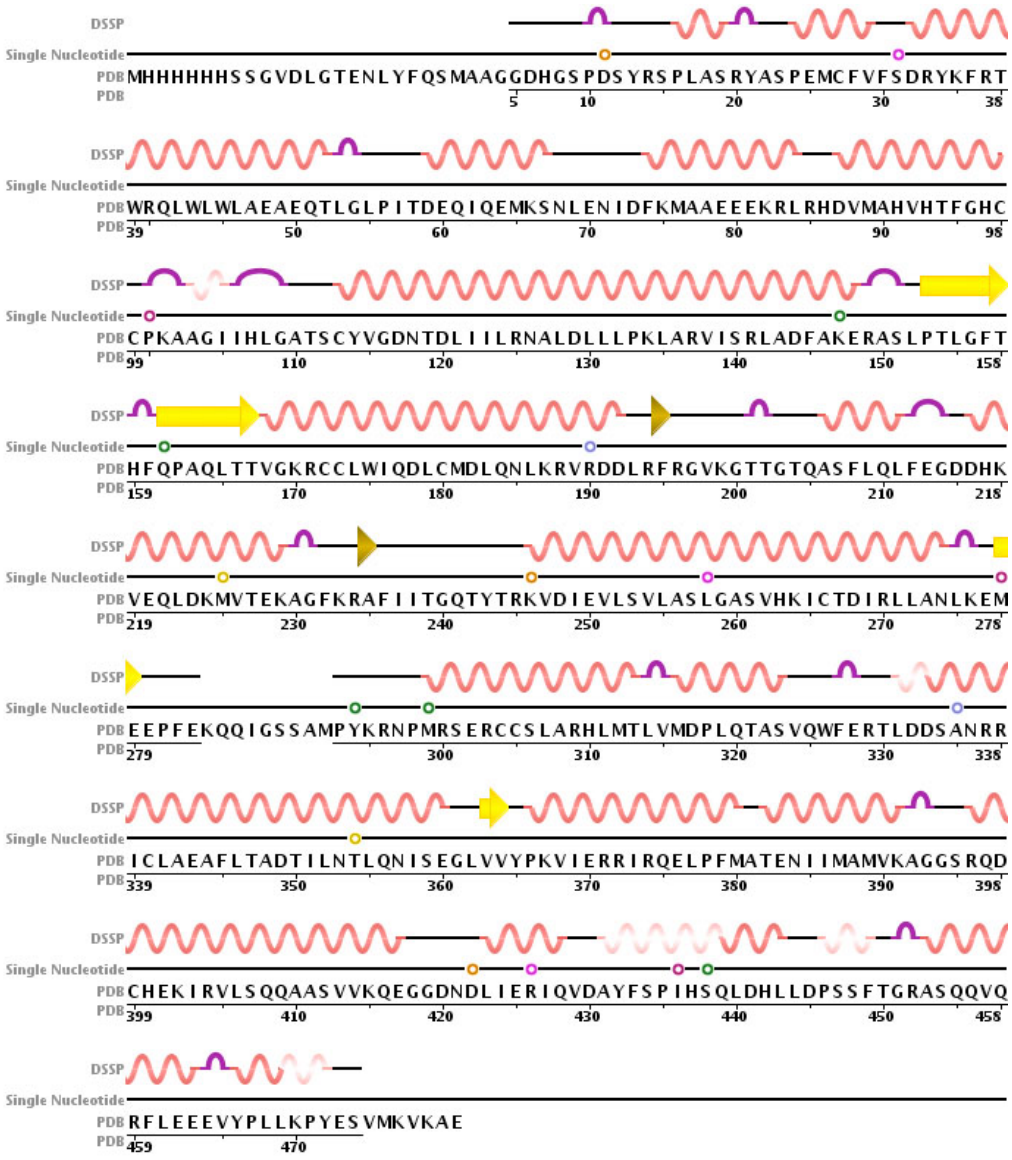


Figure 1.13: Human ADSL primary structure with labeled secondary structure

Two PDB deposited crystal structures of *H. Sapiens* ADSL, 2J91 and 2VD6, have been determined at 1.8 and 2.0 $\AA$ respectively. 2J91 was crystalized with AMP located in all four active sites. 2VD6 was crystalized with

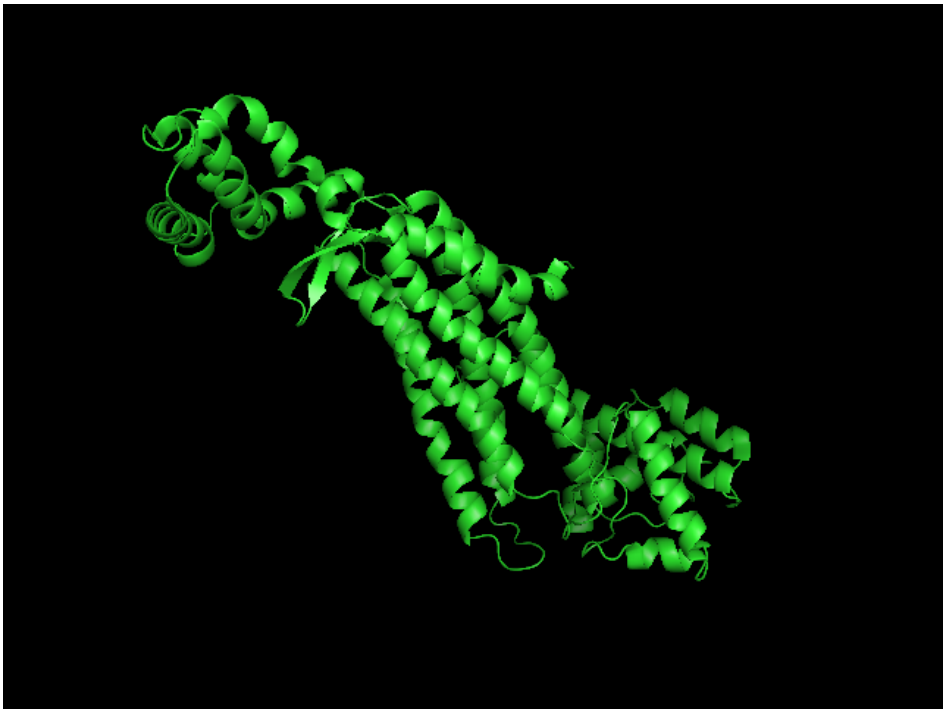


Figure 1.14: Human ADSL tertiary structure

SAMP in two active sites and AMP plus fumarate occupying the other two active sites. Both crystal structures show a tetramer structure comprised of four identical subunits, with D2 symmetry (Figure 1.15). Three monomers contribute amino acids to the active site [11], creating 4 active sites total in the homotetramer. The active sites can accommodate both substrates [77], with no necessary adaptations. This is because in SAICAR, the carbon to which the succinyl moiety is attached is able to rotate 180° prior to ring closure [82]. The D2 symmetry of ADSL is important because it tells us that ADSL is a "dimer of dimers", or a protein with two distinct interfaces. This strongly suggests the protein assembly pathway will follow a pathway com-

prised of monomers and dimers to form the active tetramer structure [65].

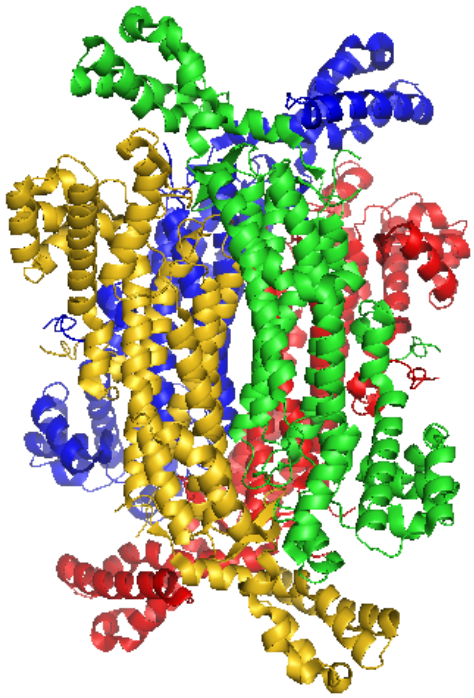


Figure 1.15: Human ADSL crystal structure

ADSL is active only in the tetrameric form, which implies the tetramer is in equilibrium with monomer subunits from which they are formed. However, the presence of dimers, trimers and higher order oligomers (aggregation number higher than four) cannot be ruled out either. The different degree of structures that exist among these different multimeric units can also vary. Thus, there could be a highly heterogeneous pool of species with different size and structure that may coexist with the functionally active tetrameric ADSL. Under normal conditions tetramers will be the most stable and dominant state and the populations of others will be negligible [3]. However the equilibrium

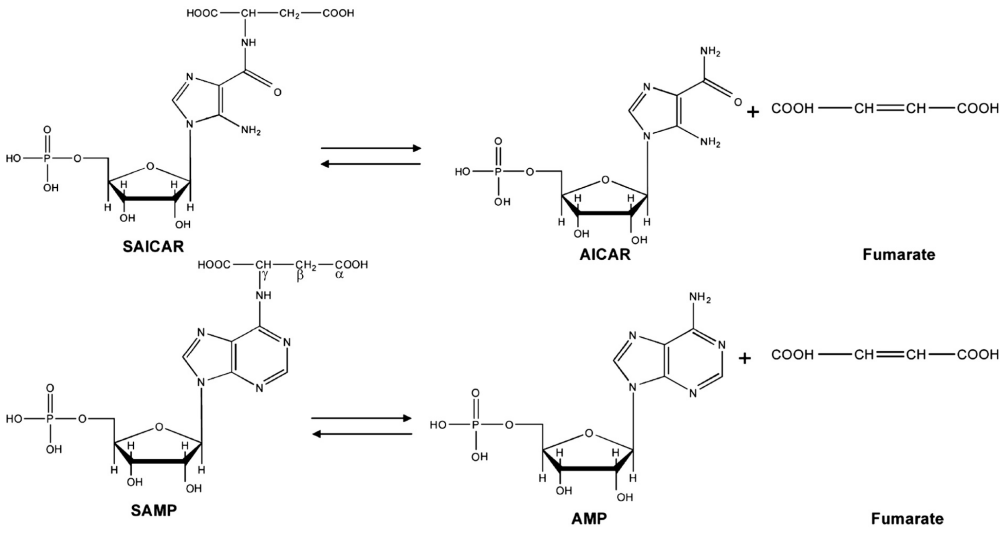


Figure 1.16: ADSL substrates and products

among these different species may get significantly altered in the presence of disease mutants and hence affect the function. The presence of aggregates due to mutations has already been reported in some studies [3]. ADSL clearly has a diverse and interesting landscape that will require carefully designed experiments and non-trivial solutions to modeling when investigating its inherent properties.

1.9 Introduction to Protein Folding

From the 1850s to the 1950s, scientists followed the perception that proteins were rigid entities that employed a "lock and key" model proposed by Emil Fischer. This model attempted to explain the substrate specificity of enzymes. The problem with this view is if the enzyme's active site is complementary

to the substrate and not the product. This would impede the reaction from taking place because the free energy barrier between the enzyme-substrate and enzyme-product complexes will be too large. The view of proteins changed from static to dynamic when Daniel Koshland proposed the "induced fit" model. It suggested proteins use small conformational changes in the active site to create communication between enzyme structure and substrates. Thus the establishment of a dynamic folded structure for functional proteins began. These dynamics range from vibrational motions that last around the order of femtoseconds to motions of side chains and secondary structures that last around the order of microseconds.

1.9.1 Driving Forces for Protein Folding

The term "unfolded protein" can take on two different definitions: 1) The state of the protein prior to folding after being transcribed by the ribosome, or 2) the state resulting from a complete protein denaturation. From a thermodynamic point of view, it is not too important which definition is used as long as the process is reversible. But the often thought idea of the unfolded protein being a "random coil" is more of a hypothetical state than an actual one because unfolded chain often retain residual structure. When done properly, protein denaturation experiments are able to quantify the free energy between states. Suprisingly, such experiments reveal protein stability to be marginal [25, 32]. Despite the thousands of non-covalent interactions existing in the structures of folded proteins, the free energy between the native and unfolded states is a mere 5-20 kcal/mol [24, 61, 66]. At first, this may seem surprising, noting that a single covalent bond is 65-175 kcal/mol. However,

when considering that proteins need to maintain their native fold and dynamic nature in order to function, it seems reasonable.

Non-covalent bonds are important for the stability of the protein while allowing enough motion for them to be functional. This leads to the belief that they are the driving forces behind protein folding. The three main non-covalent interactions are non-polar, electrostatic, and van der Waals. Thus the free energy of folding can be written as

$$\Delta G_{folding} = \Delta G_{np} + \Delta G_{elec} + \Delta G_{vdW} - T\Delta S \tag{1.9.1}$$

where the last term is the stability due to configuration entropy. Which of these free energies is the driving force for protein folding? In the unfolded state, most of the peptide chain is exposed to the aqueous environment. Most of molecules in the solvent will be water molecules. This would cause electrostatic and van der Waals interactions between the chain and water. This is also stabilized by the high configurational entropy. The folded state is stabilized by many of these same interactions, but the majority of intramolecular interactions occur in the protein core where the majority of the amino acids are non-polar. This leads to the common hypothesis that the "hydrophobic effect" is the driving force for protein folding while configurational entropy opposes it [18].

1.9.2 Kinetics and Dynamics of Protein Folding

The folding rate of protein is a quick process, where some proteins fold up on the microsecond time scale [28]. Theoretically the folding process involves a search through many different configurations, intermediate states, and a large degree of freedom. The two previous statements seem to be at odds with another and the problem was presented by Cyrus Levinthal in 1968, now known as the Levinthal's Paradox [51]. It can be summarized as follows:

Assuming a protein with a chain length of 100 amino acids long samples all of the possible conformations or the protein and each residue has at least 3 states. The number of different conformations the protein can assume is equal to $3^{100} = 5 \text{ x } 10^{47}$. Assuming the protein takes 1 ps to sample one conformation, the total time to sample all the conformations would be 1.6 x 10^{28} years. This time period is longer than the age of the universe. How is this reconciled with the fact that proteins fold up in the microsecond time scale? The answer is proteins do not randomly sample all possible configurations. Rather, they fold in a cooperative manner where each step reduces the number of folding possibilities for future steps. Protein folding cooperativity can be depicted by the folding funnel (Figure 1.17) that has been developed by Energy Landscape Theory [12, 13, 19, 20]. The shape of the folding funnel points out three key aspects of folding: 1) the folding process involves a decrease in both entropy and free energy, 2) there are many transitions state defined by local minima separated by energy barriers, and 3) there are multiple paths along which folding can occur from the unfolded to the native state.

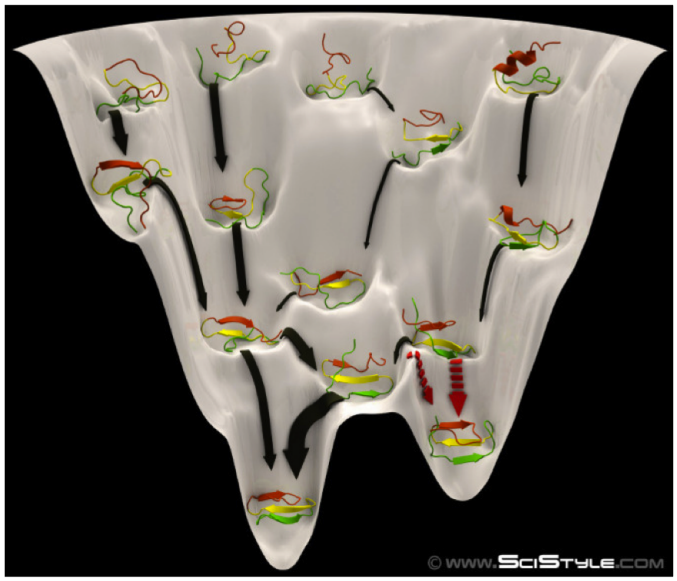


Figure 1.17: Current understanding of how a protein folds. The unfolded protein starts at the top of the funnel with maximum entropy and free energy and ends as a folded protein at the bottom of the funnel with minimum entropy and free energy.

While the folding funnel provides a good general solution to Levinthal's Paradox, the kinetics of the folding process still needs to be addressed. This is a particular issue with protein folding because of the difficulty to detect an intermediate state or a structural change that has a very short life span. From the current data, there have been three basic models proposed to describe the kinetics: framework model, hydrophobic collapse or molten globule model, and the nucleation model(Figure 1.18). The framework model proposes the secondary structures form first, followed by the final fold into the

tertiary or quaternary structure. The hydrophobic model suggest the driving hydrophobic forces cause the protein to collapse. It is loosely packed due to non-polar residues still being exposed to the solvent and it is possible some secondary folds have occurred. Eventually the non-polar residues interact and form the compact final fold. The nucleation model proposes both hydrophobic collapse and complete secondary folds occur simultaneously until the final fold is acheived. There is always the possibility of a combination of the three models as well referred to as the jigsaw model. The moral of the story is there is no single, universal mechanism for the kinetics of protein folding. This also hasn't taken into account the *in vivo* factors of protein folding such as the dense cytoplasm, translation process, and molecular chaperones and stabilizers.

It is clear that a protein needs to reach its native state to be functional. However, even when the native state is reached, the dynamics of the folded state are continuously happening. This means the protein may exist in an equilibrium between many different native conformations. From thermodynamics, we would expect the protein to spend the most of its time occupying the state with the least energy, which is what is seen in crystallographic structures. Other conformations are adopted by using the thermal energy of the system for kinetic energy This results in structures that are very similar on a macroscale where the observer would see the quaternary structure, but not on a microscale where the observer would see atoms in different positions.

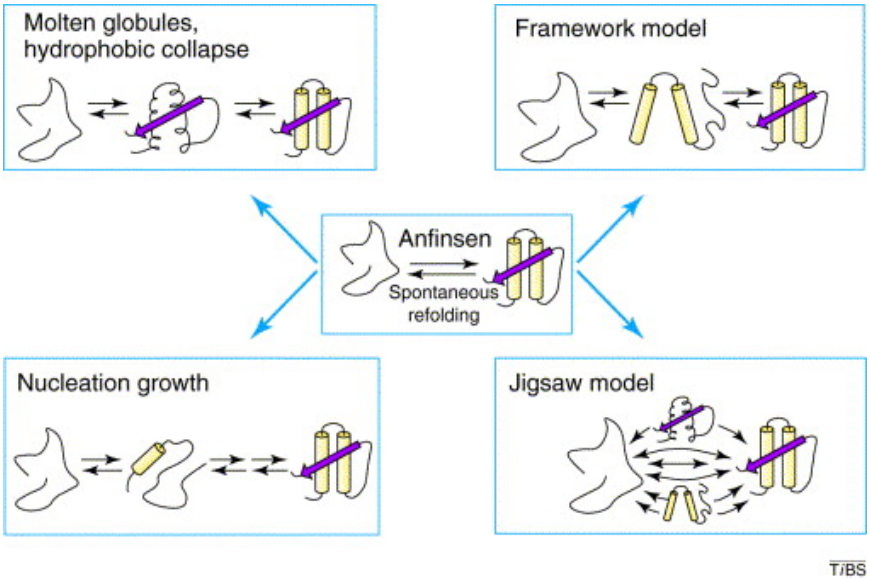


Figure 1.18: Different proposed mechanisms for protein folding. Figure was taken from [68]

1.9.3 Protein Misfolding in Human Genetic Disorders

Proteins are dynamic structures that can misfold after being synthesized or will undergo reversible structural changes as long as the structural unfolding is followed by a fast re-folding process. Occasionally, the unfolding is a permanent transition that can have detrimental effects on the cell. Misfolded proteins lose their ability to perform their biological function (loss of function disorders), however they often gain an ability to aggregate or interact inappropriately with other cellular components (gain in function disorders) leading to cell function impairment and eventually cell death. Normal proteins will experience misfolding to some degree and if the process is not reversible, the protein quality control system will tag the protein for degradation. One factor that can increase the propensity of misfolding is genetic mutations of the amino acid

sequence. Genetic mutations can destabilize the native structure of proteins, induce aggregation, and may be congenital. There are numerous human genetic disorders that are now associated with protein misfolding such as Cystic Fibrosis, Alzheimer's Disease, Parkinson's Disease, Huntingon's disease, Amyotrophic Lateral Sclerosis, as well as many others. The establishment of the misfolding theory as a pathogenic mechanism has had important implications for pursuing research characterizing the stabilty-structure-function relationship of enzymes. Particularly for therapeutic treatments to either increase the stability or enhance the functionality via pharmaceuticals.

1.9.4 Review of Thermodynamics

The get a true physical understanding of proteins, we need to characterize them in terms of forces and energies, particularly their energetics and dynamics. A biophysical approach also requires careful attention to the direction the forces are driving the process. Are they folding or unfolding the protein? The basic principles of thermodynamics can be applied to describe the free energy of the system and the probability of different states under different environmental conditions. It utilizes the dependency of energy on temperature, pressure, volume, and chemical concentration to characterize these different states. Since proteins experience a constant temperature and pressure, Gibbs free energy (G) is used to determine directionality of the process. Specifically, the change in free energy (ΔG) describes the direction as the system will proceed toward a state of minimum free energy. (ΔG) can be calculated by the following equation

$$\Delta G = \Delta G^o + RTlnK_{eq} = \Delta G^o + RTln\frac{f([products])}{g([reactants])} \tag{1.9.2}$$

where ΔG^o is the standard free energy, R is the universal gas constant, T is the temperature, and f and g are dependent on the stoichiometry and concentration of the products and reactants. Once the free energy is measured, it can then be used to determine the fraction of the population that went through the transition. The two major components to free energy are enthalpy and entropy which will be described in the next two sections.

$$P \propto e^{\frac{\Delta G}{RT}} \tag{1.9.3}$$

Enthalphy

Enthalpy (H) is related to the internal energy of the system (E) and the work from pressure (P) and volume (V) done on or by the system.

$$H = E + PV \tag{1.9.4}$$

Hence, any change in H will be due to internal energy changes of the system, or by the work done on or by the system by changes in P or V. Since the biological environment does not experience significant change in P or V, the ΔH is mostly due to ΔE

$$\Delta H = \Delta E \tag{1.9.5}$$

This ΔH is caused by three molecular phenomena: 1) formation or breaking of covalent bonds, 2) changes in electrostatic of van der Waals interactions, or 3) changes in thermally induced atomic motions. These interactions involve a heat transfer (q) between the system and the environment. In biological systems, q generally decreases the ΔH by the formation of exothermic chemical bonds. This q can be experimentally measured by calorimetric methods such as differential scanning calorimetry or isothermal calorimetry which measure heat capacity (C_p). If the process is reversible, C_p can be calculated by

$$\Delta C_p = \frac{q}{\Delta T} = \frac{\Delta H}{\Delta T} \tag{1.9.6}$$

A positive ΔH would coincide with the breaking of bonds, hence the unfolding process. This results in a positive ΔC_p. This infers the unfolded protein has a larger ΔC_p than the folded protein. This is not too surprising since the unfolded protein has more accessible surface area (ASA) and intramolecular interactions occurring. Therefore is would take more heat energy to increase the temperature of the system. By modeling the dependence of the ΔASA in terms of chain length, Myers et al found an expression of ΔC_p in terms of ΔASA [60]

$$\Delta C_p = -251 + 0.19(\Delta ASA) \tag{1.9.7}$$

Entropy

The entropy (S) of the system is related to the number of possible configurations (ω).

$$S = k_B ln\omega \tag{1.9.8}$$

A highly ordered system cannot adopt many configurations. This means the entropy and the order of the system are inversely related. A higher order system has a lower entropy than a disorders system. This is why S is associated with the amount of disorder in a given system. The change in entropy (ΔS) is represents changes in the freedom of motion in either the solute or the solvent for molecular processes. As a protein folds into secondary, tertiary, and quaternary structures, the amount of order increases causing a decrease in entropy. The entropy of the system increases because as the non-covalent bonds form, it is accompanied by an increase of S in the solvent caused by the disordering of the water molecules surrounding the non-polar regions. At constant temperature, the ΔS can be related to the ΔH by

$$\Delta S = \frac{\Delta H}{T} = \frac{q}{T} \tag{1.9.9}$$

Using equation 1.9.6 to express ΔS in terms of ΔC_p

$$\Delta S = \frac{\Delta C_p}{T} \Delta T \tag{1.9.10}$$

This relation suggests ΔS also measures the number of states the heat energy can be dispersed over at a constant temperature. As we have already discussed, the unfolded protein has a higher C_p than the folded state. Equation 1.9.10 is consistent with that conclusion by showing a higher S will require a greater amount of energy to raise the temperature. It also allows for experiments to quantify the ΔS with measure values of ΔC_p and temperature.

1.9.5 Methods of Denaturation

Protein denaturation is the common method to study the free energy between states rather than protein folding. Up to this point, unfolding and denaturation have been used interchangeably. However, denaturation is different than protein unfolding. By definition, it is the process by which it loses functionality. Unfolding will occur in the process based on the structure-function relationship and may lead to a completely unfolded protein, but the denaturation pathway is not necessarily the same at the unfolding pathway. Since thermodynamics is only concerned with the state of the peptide chain, denaturation is accepted at a means to study the free energy of the transitions induced by denaturant methods. The four standard methods of denaturation, or different knobs we can turn to change the environment, are temperature, pH, pressure, and chemical.

Temperature

Raising the temperature of the system will cause an increase in the kinetic energy of the system. The added kinetic energy boosts the configurational entropy of the unfolded state due to increasing the vibrations and motion of the peptide chain. Interestingly, at high temperatures, since the kinetic energy of the water is too high to drive the hydrophobic effect, protein folding is an enthalpy-driven process as opposed to an entropic-driven process. Denaturation can also occur at low temperatures where the dynamics of the protein are slowed down and "frozen". This lack of dynamics causes a loss of function, but not necessarily structure. As the water begins to freeze and form a crystal lattice, hydrogen-peptide water bonds are lost. This can lead to aggregation and precipitation out of solution.

pH

Enzyme have an optimum pH range where they are able to catalyze reactions. If the pH is too far out of this range, the enzyme begins to lose function. The pH can also affect the structure of the protein by changing the charge of electrically charged amino acids. A low pH would cause pronation, while a high pH would cause de-pronation. These ionic changes can have an affect on stabilizing interactions such as salt-bridges by weakening the coulomb interaction or reducing the electrostatic screening. They could also affect catalysis by altering the binding sites for substrates or the cleavage mechanism.

Pressure

Increasing the hydrostatic pressure will cause the density of the water to rise. Eventually the water molecules will begin to penetrate the protein surface, therefore increasing the probability to unfold. This is the least invasive method of protein denaturation and is one of the better approximations for protein unfolding.

Chemical

There are two types of chemicals that can be added to the solvent to induce protein denaturation, organic and polar. Organic solvents interact with the non-polar side chains eventually overcome the intramolecular interaction. This causes an unfolding process to occur since the non-polar residues are now stable in the denatured state. The mechanism by which polar denaturants, such as guanidinium chloride and urea, cause denaturation is still unclear. Being polar, they are believe to interact mainly with the surface residues. This infers they induce denaturation by an indirect effect to the solvent which requires a high concentration of the denaturing chemical. This is usually the case for urea, while guanidinium chloride is a stronger chemical that takes a lower concentration to induce denaturation.

1.9.6 Solutions to Unfolding Transitions

As stated in the thermodynamic review, the fraction of the population that went through the transition can be determined by the free energy between the states. For chemical denaturation experiments the free energy change

for unfolding in the absence of denaturant (ΔG^o) can be derived by linear extrapolation of the free energy change of the unfolding in the presence of various concentrations of denaturant

$$\Delta G = \Delta G^o + m[denaturant] \tag{1.9.11}$$

where m is the m-value of the protein. Simply, the m-value is the measure of a protein's response to the added chemical. A large response refers to a quick unfolding process while a low response indicates a slow unfolding process. There are two cases where the fraction folded (unfolded) can be solved for directly: a simple two state transition and for a dimer unfolding.

Simple Two State Transitions

$$[M_i] \rightleftharpoons [U_i] \tag{1.9.12}$$

Where M, U, and i represents the monomer state, unfolded state, and the stoichiometry respectively. Meaning a monomer or multimer could undergo a two-state transition from a folded monomer (multimer) to an unfolded monomer (multimer). For the rest of this example, the i will be dropped.

$$\Delta G = -RTlnK_{eq} \tag{1.9.13}$$

$$K_{eq} = \frac{[U]}{[M]} \tag{1.9.14}$$

$$\Delta G = -RTln\frac{[U]}{[M]} \tag{1.9.15}$$

Solving for the ratio of $\frac{[U]}{[M]}$

$$\frac{[U]}{[M]} = e^{-\beta\Delta G}; \beta = RT \tag{1.9.16}$$

If the fraction of interest is the fraction of proteins that are unfolded (F_u), solve for $[M]$ in terms of $[U]$. If the fraction of interest is the fraction of proteins that are folded (F_f), solve for $[U]$ in terms of $[M]$. In this case, F_u will be our fraction of interest.

$$[M] = [U]e^{\beta\Delta G} \tag{1.9.17}$$

If the fraction of interest is the fraction of proteins that are unfolded (F_u)

$$F_u = \frac{[U]}{[P]} \tag{1.9.18}$$

Where $[P]$ is the total protein concentration

$$[P] = [M] + [U] \tag{1.9.19}$$

Substituting equations 1.9.17 and 1.9.19 into equation 1.9.18

$$F_u = \frac{[U]}{[U]e^{\beta\Delta G} + [U]} = \frac{1}{e^{\beta\Delta G} + 1} = \frac{K_{eq}}{K_{eq} + 1} \tag{1.9.20}$$

Where $K_{eq} = \mathrm{e}^{-\beta\Delta G}$

Dimer Unfolding

The only higher order unfolding process that can be solved for directly without the use of iterative methods is the transition from folded dimer to unfolded monomer

$$[M_2] \rightleftharpoons [U_1] \tag{1.9.21}$$

$$\Delta G = -RTlnK_{eq} \tag{1.9.22}$$

$$K_{eq} = \frac{[U_1]^2}{[M_2]} \tag{1.9.23}$$

$$K_{eq} = \frac{[U_1]^2}{[M_2]} \tag{1.9.24}$$

Solving for $[M_2]$

$$[M_2] = [U_1]^2 e^{\beta \Delta G} \tag{1.9.25}$$

$$F_u = \frac{[U_1]}{[P]} \tag{1.9.26}$$

Writing the total protein concentration in terms of monomer concentration

$$[P] = 2[M_2] + [U_1] = 2 \cdot [U_1]^2 e^{\beta \Delta G} + [U_1] \tag{1.9.27}$$

Using the quadratic formula to solve for $[U_1]$

$$2[M_2] + [U_1] = 2 \cdot [U_1]^2 e^{\beta \Delta G} + [U_1] - [P] = 0 \tag{1.9.28}$$

where $a = 2{\cdot}\mathrm{e}^{\beta\Delta G}$; $b = 1$; -$[P]$

$$[U_1] = \frac{-1 \pm (1 + 8e^{\beta\Delta G} \cdot [P])^{\frac{1}{2}}}{4e^{\beta\Delta G}} \tag{1.9.29}$$

Choosing "+" because the final $[U_1]$ concentration cannot be negative and substituting ${K_{eq}}^{-1}$ for $\mathrm{e}^{\beta\Delta G}$

$$[U_1] = \frac{(1 + 8K_{eq}^{-1} \cdot [P])^{\frac{1}{2}} - 1}{4K_{eq}^{-1}} = \frac{(K_{eq}^2 + 8K_{eq} \cdot [P])^{\frac{1}{2}} - K_{eq}}{4} \tag{1.9.30}$$

Substituting equation 1.9.30 into equation 1.9.26

$$F_u = \frac{[U_1]}{[P]} = \frac{(K_{eq}^2 + 8K_{eq} \cdot [P])^{\frac{1}{2}} - K_{eq}}{4[P]} \tag{1.9.31}$$

Which can be solved by a minimization function in most mathematica softwares.

Chapter 2

Inherent Properties of Adenylosuccinate Lyase Could Explain S-Ado/SAICAr Ratio Due to Homozygous R426H and R303C Mutations

2.1 Introduction

Enzymatic studies provide the specific activity and the maximum product formation rate (V_{max}) as a result of enzyme catalysis of the substrate. They also measure the equilibrium constant (K) between the free enzyme state and the enzyme substrate complex formation. The importance of the V_{max} is that it gives a quantitative measure of the extent to which the enzyme acts on

its designated substrates, SAICAR and SAMP. Numerous studies attempt to relate V_{max} values with mutations associated with various types of ADSL deficiencies and the corresponding S-Ado/SAICAr ratio [3,4,44,62,63,67,76,98]. However, previously reported human ADSL kinetic parameters have been inconsistent. The wild-type (WT) SAICAR/SAMP activity ratio varies from 0.53 - 1.7 while SAMP and SAICAR K_m values vary from 1.9 - 4.9 and 1.8 - 3.6 μM, respectively [4, 44, 70]. Recent studies using each substrate independently have identified only one disease associated mutation, R303C, resulting in a nonparallel reduction of enzyme activity favoring SAICAR over SAMP [3,4,44,70,98]. An early study of fibroblasts taken from patients showing reduced ADSL activity showed that synthesis of completed adenine and guanine nucleotides remains possible in both cell types [86, 87]. The affected patients used in these studies were originally reported in 1988 by Jaeken and Van den Berghe and the disease causing mutations reported later [38, 54, 67]. The fibroblasts used were from severely affected individuals including two patients homozygous for the R426H mutation and one mildly affected patient homozygous for the R303C mutation. R426H is the most common mutation occurring in 30% of all known patients; homozygous R426H mutations generally lead to severe disease type and result in a S-Ado/SAICAr value of 1.0 - 1.6 [41, 78]. The homozygous R303C mutation leads to the mildest form of the deficiency and gives rise to a S-Ado/SAICAr ratio of 3.0 - 3.7 [78]. Interestingly, only fibroblasts from the individual with the R303C mutation accumulated SAICAR or SAMP as measured by radiolabeled formate incorporation despite having a similar reduction in SAICAR activity as fibroblasts from the R426H patient. This may help explain the elevated ratios of S-

Ado/SAICAr seen in the extracellular fluid of the R303C patients, and led to the hypothesis that SAICAR accumulation may be toxic. A larger reduction in SAMP activity would result in more accumulation of intracellular SAMP, and subsequently more S-Ado. However, alternative hypotheses to explain this difference in S-Ado/SAICAr ratios have been proposed, including the age of the patients when tested, the dephosphorylation rate of each substrate (likely required for extracellular transport), or the rate of extracellular transport itself [98].

As discussed above, most kinetic studies of ADSL mutant proteins have been done with each substrate individually. This does not reflect the in vivo situation, in which the presence of both substrates would be expected. Findings reported for the R303C fibroblasts by Van den Bergh et al. suggest that for the in vivo situation, at least for the R303C mutation, the elevated S-Ado/SAICAr is due to the intrinsic properties of the enzyme [86, 87]. Under in vivo conditions both substrates bind with the ADSL enzyme. Competition due to resource sharing can mediate an indirect communication between the two substrates that do not interact otherwise. This hidden coupling can alter the activities of the substrates from the non-competitive scenario, i.e. in the absence of the other substrate. Emergence of complex patterns due to resource sharing has been reported in other biochemical networks [56, 74].

In this study, we begin to explore the relationship between in vitro and in vivo environments and the possibility of hidden coupling and resource sharing. We describe a novel electrochemical-detection (EC) method to obtain the

V_{max} for ADSL in solution with its two substrates at varying concentration ratios. Using this method we are able to more accurately determine the affects of hidden substrate coupling and resource sharing in dual substrate environments on the bi-functional ADSL enzyme and disease causing mutations. For this purpose we chose two mutations that represent extremes of the phenotype, R303C representing the mild form (Type II) and the R426H representing the severe form (Type I). Additionally, studies were done on the A291V ADSL mutation found in the CHO-K1 AdeI mutant. This mutant provides a useful control since its activity has previously been reported to be extremely low or undetectable with either substrate [7, 84, 90]. Finally an enzyme kinetic model was developed to predict SAMP and SAICAR activity from a given ratio of substrate concentration.

2.2 Enzyme Kinetics by UV

Enzyme kinetic parameters for WT and R303C have been previously reported [70]. V_{max} values for the R426H mutant have been published for SAMP and SAICAR activity [44, 98]. K_m values have not been reported in the literature. In light of this, an initial investigation of enzyme activity using SAMP and SAICAR individually was performed by standard UV kinetic assays for the WT and for A291V, R303C, and R426H enzymes. All ADSL mutants with the exception of A291V have measureable activity that could be evaluated by monitoring the UV absorbance of SAMP or SAICAR. It has been shown previously that WT ADSL does not follow simple Michaelis-Menten kinetics [3]. Therefore, kinetic curves were analyzed with the Hill Equation (Figure 2.1).

SAMP and SAICAR kinetic parameters can be found in Table 2.1 and Table 2.2 respectively. WT specific activity on SAMP and SAICAR is 13.9 and 23.3 μmol/min/mg respectively. The ratio of activity of SAICAR/SAMP is 1.7. This ratio is in agreement with the previously reported ratio of activities using the UV assay [4, 67]. The R303C mutation reduces specific activity to 5% of SAMP and 38% of SAICAR relative to WT ADSL. This agrees with previous reports the R303C mutation displays a nonparallel decrease in activity in vitro and in vivo [70, 87, 98]. The measured specific activities for R426H with SAMP and SAICAR respectively are 6.7 and 6.4 μmol/min/mg. These correspond to a reduction to 45% for SAMP activity and 27% for SAICAR activity relative to WT and a gives SAICAR/SAMP ratio of 0.96. This initially suggests the R426H mutation also causes a nonparallel decrease in activity, though not to the degree as R303C. Although the mutants reduce the activity on both substrates, R303C increases the $K_{0.5}$ while R426H decreases the $K_{0.5}$. The decrease in $K_{0.5}$ may indicate an increase in substrate affinity in order to maintain R426H enzyme efficiency at a level close to the WT ADSL. WT and R426H ADSL appear to have equivalent cooperative binding on both substrates, while R303C demonstrates no cooperativity. This may indicate that cooperative binding does not play a critical role in retaining enzyme activity and other factors such as substrate binding or alterations of the catalytic mechanism are more important to retain activity [70].

Table 2.1: SAMP UV Kinetic Parameters of Human ADSL at 25°C

Enzyme	V_{max} (μmol min^{-1} mg^{-1})	$K_{0.5}$ (μM)	Hill Coeff.	$\frac{k_{cat}}{K_{0.5}}$
WT	100 ± 4	2.05 ± 0.08	1.4 ± 0.1	2.52 x 10^7
A291V	ND	ND	ND	-
R303C	5 ± 3	2.7 ± 0.2	0.97 ± 0.07	9.15 x 10^5
R426H	45 ± 3	1.4 ± 0.1	1.5 ± 0.2	1.78 x 10^7

Kinetic parameters for SAMP activity, V_{max}, k_{cat}, $K_{0.5}$, and the Hill coefficient, were determined by standard UV assays by varying substrate concentration and fitting the data to the Hill equation in Mathematica. Protein was reconstituted for 2 h at 25 C prior to measurements. Activities are relative to WT SAMP activity. ND is Not Detectable. Values are shown along with their standard error.

Table 2.2: SAICAR UV Kinetic Parameters of Human ADSL at 25°C

Enzyme	V_{max} (μmol min^{-1} mg^{-1})	$K_{0.5}$ (μM)	Hill Coeff.	$\frac{k_{cat}}{K_{0.5}}$
WT	100 ± 2	1.74 ± 0.08	1.3 ± 0.1	5.09 x 10^7
A291V	ND	ND	ND	-
R303C	38 ± 6	9 ± 0.9	0.9 ± 0.1	3.64 x 10^6
R426H	27 ± 3	0.7 ± 0.1	1.3 ± 0.3	3.75 x 10^7

Kinetic parameters for SAICAR activity, V_{max}, k_{cat}, $K_{0.5}$, and the Hill coefficient, were determined by standard UV assays by varying substrate concentration and fitting the data to the Hill equation in Mathematica. Protein was reconstituted for 2 h at 25 C prior to measurements. Activities are relative to WT SAICAR activity. ND is Not Detectable. Values are shown along with their standard error.

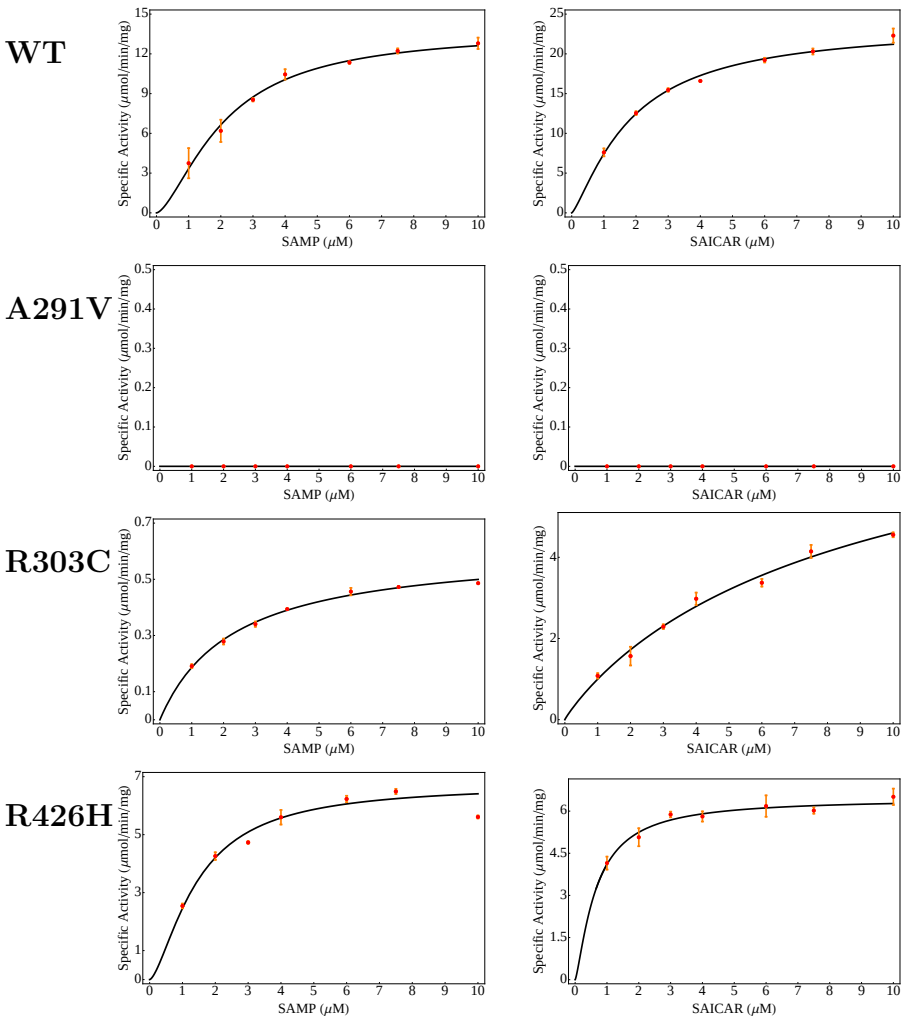


Figure 2.1: Specific Activity vs [Substrate] plots.

2.3 Enzyme Kinetics by HPLC-EC

To explore the effect on the specific activity when both substrates are present in solution we measured the levels of both substrates and products after separation by HPLC-EC and UV detection. The identities of the substrates and products were verified by co-chromatography with known compounds. To test the method, kinetic measurements on SAMP and SAICAR were run independently. WT specific activity on SAMP and SAICAR is 13.4 and 7.4

μmol/min/mg respectively. This results in a SAICAR/SAMP activity ratio of 0.55 and is similar to results previously reported using an HPLC based assay [42, 44]. Interestingly, even though the V_{max} for SAICAR is different between the UV and HPLC-EC assays, R303C activities maintained the same percentage drop in activity, 5% SAMP and 40% SAICAR. The HPLC-EC data for R426H show reduction in activities for SAMP and SAICAR to 47% and 51% respectively. As a further validation of the assay, we attempted to measure the levels of ADSL activity in the A291V mutant, previously found to be extremely low or not detectable with either substrate. This was found to be the case using the HPLC-EC assay as well.

Having validated the HPLC-EC assay for each individual substrate, we then assayed the ADSL activity for WT, R303C, and R426H mutant enzymes in the presence of varying ratios of SAICAR and SAMP. Chromatograms showed a preserved separation of products and substrates as well as the growth of product and conservation of substrate from 0 - 5 minutes (Figure 2.2). Accumulation of products was linear for the five time points as shown in Figures 2.3. A summary of the results are shown in Table 2.3 and Table 2.4 and represented graphically in Figure 2.4. For WT and R426H, as SAICAR concentration increases, SAICAR activity increases as SAMP activity decreases as expected. For the R303C mutation, SAMP activity remains fairly constant as SAICAR concentration increases. To quantitatively analyze the data, we calculated the ratio of activity rates (SAMP activity/SAICAR activity) for ADSL as a function of substrate mixture (Table 2.5 and Table 2.6). For SAMP/SAICAR mixtures of 3:1, 1:1, and 1:3, the ratio of activity rates for

WT decrease from 3.4, to 1.3, to 0.39 respectively. R426H follows a similar decrease in ratio of activity rates: 3.2, to 0.85, to 0.36. R303C ratios also decrease, but have values smaller than WT or R426H: 1.7, to 0.44, to 0.29. This implies that mutants can behave similarly or differently to WT in solution with both substrates and still result in succinylpurine accumulation. Interestingly, the relative rates of activity (mutant activity/WT activity) of R426H remain fairly constant while R303C relative rates change for any given substrate ratio (Table 2.5 and Table 2.6). Both SAMP and SAICAR relative rates for the R426H mutation remain about 50% of WT activity. The R303C mutation on the other hand has an increase in SAMP relative rate, 5 - 21%, as well as in SAICAR relative rate, 18 - 43%.

Table 2.3: Rate of Formation Values for Mixing Ratios for WT and R303C

	WT		R303C	
SAMP:SAICAR	SAMP	SAICAR	SAMP	SAICAR
1:0	13 ± 3	-	0.702 ± 0.003	-
3:1	9 ± 3	2.7 ± 0.3	0.79 ± 0.02	0.47 ± 0.08
1:1	7 ± 3	5.1 ± 0.5	0.44 ± 0.06	1.0 ± 0.2
1:3	2.4 ± 0.6	6.2 ± 0.8	0.5 ± 0.2	1.7 ± 0.5
0:3	-	7.4 ± 0.9	-	3.2 ± 0.3

Rate of product formation (dP/dt) for different mixing ratios from HPLC-EC enzyme assays. 1:0 refers to 100% SAMP, 0:1 refers to 100% SAICAR.

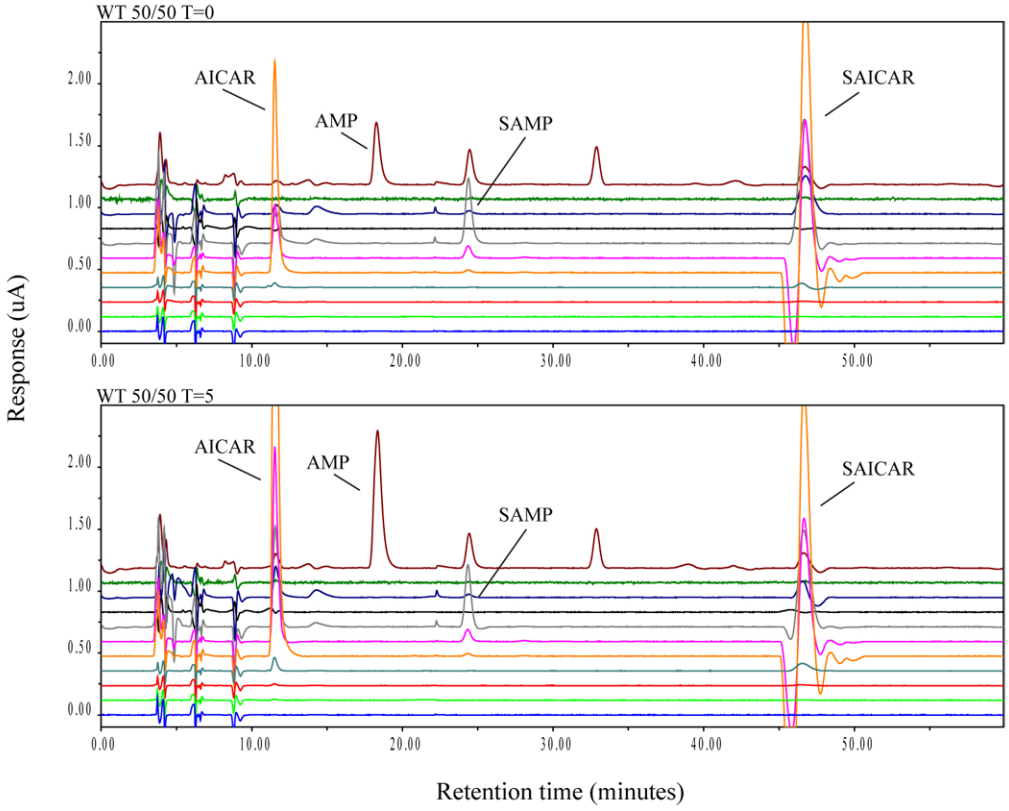


Figure 2.2: Representative chromatograms showing separation of both substrates (SAICAR, SAMP) and products (AICAR, AMP) of the wild-type ADSL kinetic assay at T=0 (A). The formation of product is visually evident as well as the conservation of substrate at reaction time T=5 (B). The EC channels shown are in 100 mV intervals from 0-900 mV. Data were taken from WT 1:1 mixing ratio of SAMP and SAICAR.

2.4 Prediction of the Specific Activity when both Substrates are Present

We carried out an enzyme mixing kinetics analysis to predict the rate of product formation when two substrates catalyzed by the same enzyme are present in solution at different concentrations. The analysis is based on the coupling of two independent enzymatic reactions via resource sharing. By

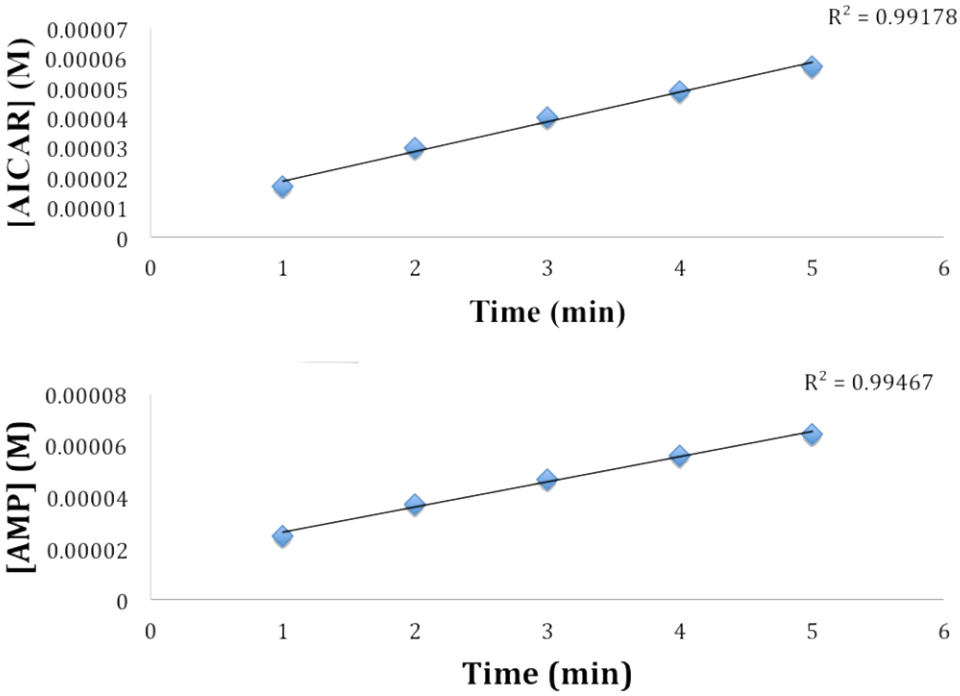


Figure 2.3: Representative linear trends of product (AICAR, AMP) growth vs time for different mixing ratios of SAMP and SAICAR using the HPLC-EC assay. Data were taken from WT 1:1 mixing ratio of SAMP and SAICAR. The activity was given by the slope of the line in units of Molarity*min^{-1} and changed into units of μmol*min^{-1}*mg^{-1}.

independent reactions, we mean two separate reactions for each substrate in the absence of the other. However, resource sharing induces a coupling due to the fact that if one enzyme molecule is occupied by substrate one, it is not available for substrate two. Although the enzyme has multiple binding sites, we assume on a given enzyme each site is occupied by substrates of the same type thus avoiding any other possible sources of coupling that may alter the equilibrium constant (K_m). Thus competitive binding has been enforced in the simplest form by conserving the total number of resources between the two substrates. This is the only source of interaction between the two

Table 2.4: Rate of Formation Values for Mixing Ratios for WT and R426H

	WT		R426H	
SAMP:SAICAR	SAMP	SAICAR	SAMP	SAICAR
1:0	13 ± 3	-	6.3 ± 0.4	-
3:1	9 ± 3	2.7 ± 0.3	4.1 ± 0.1	1.3 ± 0.1
1:1	7 ± 3	5.1 ± 0.5	2.2 ± 0.1	2.6 ± 0.3
1:3	2.4 ± 0.6	6.2 ± 0.8	1.1 ± 0.2	3.1 ± 0.1
0:3	-	7.4 ± 0.9	-	3.8 ± 0.2

Rate of product formation (dP/dt) for different mixing ratios from HPLC-EC enzyme assays. 1:0 refers to 100% SAMP, 0:1 refers to 100% SAICAR.

reactions in our resource sharing model. Based on work by Ray, Deaton et al, cooperativity may play a small role in substrate binding and catalysis. Therefore, for simplicity, we disregard cooperativity in the present model [70]. From these assumptions, the rate of product formation for each substrate in mixing conditions can be predicted. The resulting equations 2.4.1 and 2.4.2 are predictions for the rate of product formation when both substrates are present in solution.

$$\frac{d[P_1]}{dt} = \frac{V_{max,1}}{\frac{K_{m,1}}{K_{m,2}} \cdot x + 1} \tag{2.4.1}$$

$$\frac{d[P_2]}{dt} = \frac{V_{max,2}}{\frac{K_{m,2}}{K_{m,1}} \cdot \frac{1}{x} + 1} \tag{2.4.2}$$

Table 2.5: Activity Ratios and Relative Rates for Mixing Ratios for WT and R303C

	WT	R303C	R303C/WT (%)	
SAMP:SAICAR	SAMP/SAICAR	SAMP/SAICAR	SAMP	SAICAR
1:0	-	-	5.2	-
3:1	3.4	1.7	8.7	18
1:1	1.3	0.44	6.6	20
1:3	0.38	0.29	21	28
0:3	-	-	-	43

WT and mutant ADSL activity ratios of the rate of product formation and relative rates of activity from values in Table 2.3. WT has decreasing activity ratios while R303C always has a lower activity ratio. The R303C mutation results in a changing relative ratio.

where dP/dt is the rate of product formation of the substrate at a given ratio of substrates, $x = S2/S1$. The $V_{max,i}$ and $K_{m,i}$ values refer to the pure reaction when there is only one substrate (i=1 or 2) present.

To test the prediction of the rate of product formation for a given substrate ratio for WT and mutant ADSL, V_{max} values from the HPLC-EC assay and K_m values from the UV assay were used in equations 2.4.1 and 2.4.2. Specific activities from the HPLC-EC assay were used since it was the method employed to measure rate formation under mixed substrate conditions. The overlay of the prediction to the experimental data can be seen in Figure 2.4. The black line is the calculated rate of product formation for AMP while the red line represents the rate of AICAR formation. The black and red data points and error bars are the experimental data for AMP and AICAR respectively. The blue line is the V_{max} value for SAMP activity. Even though 2.4.1

Table 2.6: Activity Ratios and Relative Rates for Mixing Ratios for WT and R426H

	WT	R426H	R426H/WT (%)	
SAMP:SAICAR	SAMP/SAICAR	SAMP/SAICAR	SAMP	SAICAR
1:0	-	-	47	-
3:1	3.4	3.2	46	48
1:1	1.3	0.85	33	50
1:3	0.38	0.35	48	50
0:3	-	-	-	51

WT and mutant ADSL activity ratios of the rate of product formation and relative rates of activity from values in Table 2.4. WT and R426H have equivalent decreasing activity ratios. R426H results in a constant relative ratio for all substrate mixtures.

and 2.4.2 were generated assuming no change in equilibrium constants due to the presence of the other substrate, the predictions are in good agreement with the experimental data. These results suggest the rate of product formation is mainly dependent on the ratio of K_m values for pure SAMP and SAICAR and the stoichiometry of the two substrates (equation 2.4.1 and 2.4.2).

2.5 HPLC vs. UV Enzyme Assay

The discrepancy between the HPLC and UV assay is interesting. Both methods agreed with the respective previously reported specific activity ratio of SAICAR:SAMP. The two methods had equivalent SAMP specific activity, but contrasting SAICAR activity. The HPLC method utilizes a quantitative analysis of the increase in product and relatively constant substrate. The amount of product and substrate in solution is based on a set of known stan-

dards. The UV method is a standard assay that is dependent on the difference extinction coefficient, ϵ, which is determined by the difference in absorbance between the substrate and product at a given wavelength. The ϵ for SAMP-AMP is 10,000 M^{-1} cm^{-1} whereas it is 700 M^{-1} cm^{-1} for SAICAR to AICAR. This low ϵ leads to a less sensitive assay that could cause error in the measurement. Given the HPLC method equipped with EC detectors provides increased sensitivity and quantitative analysis of the growth of the products and the lack of sensitivity in the SAICAR UV assay, this provides confidence the HPLC method is the more robust and accurate assay for measuring specific activity. However, if the ϵ is significantly high, then both assays are equally reliable. This verifies the use of specific activities from HPLC and the K_m from the UV assays.

2.6 Quantitative Analysis of Mixing Kinetics

Upon mixing substrates in solution at different ratios, the specific activity depends on the amount of substrate present in a non-linear fashion. Equations 2.4.1 and 2.4.2 reduced pool of enzyme molecules available for the second reaction. This is the only allowed source of coupling between the two reactions accounted by equation A.2.8 in the appendix and the basis of our simple resource sharing model. The reasonable agreement between prediction and data supports the validity of our assumption as well as our assumption that cooperativity can be disregarded. Although modified reaction constants and hill coefficients due to mixed binding of two different substrates on a given enzyme cannot be ruled out, they may be smaller effects compared to the primary ef-

fect of sharing limited resource modeled by equation A.2.8. It is also worth noticing, as a result of competitive binding, the R303C relative activity (defined as a ratio of mutant V_{max} to WT V_{max}) is no longer a constant but a function of SAMP/SAICAR ratio whereas R426H is constant for any given ratio of SAMP/SAICAR.

It has been observed that there is inconsistency in enzyme activities reported for ADSL, even from the same lab using the same assay. The only discrepancy reported here is the specific activity on SAICAR from two different assays. Since in the case of the HPLC based assay we quantify the growth of the product AICAR, which is identified by both retention time and EC-profile, we believe this is a more direct method to measure enzyme activity. The HPLC-EC data also indicates R426H does not cause a nonparallel decrease in activity in agreement with the results from Race et al, Kmoch et al, and Zikanova et al [44, 67, 98].

2.7 Potential Relevance to S-Ado/SAICAr Ratios Observed in ADSL Deficiency

The current clinical diagnosis of ADSL deficiency is based on the detection of the succinylpurines in body fluids. It has been questioned whether the ratio of these dephosphorylated derivatives of the substrates is predictive of phenotypes or not, but rather depends on the life stage of the cell or patient that determines the S-Ado/SAICAr ratio [98]. Since the in vivo condition involves presence of both the substrates it is imperative to consider the relative drops

in activity of the enzymes under mixed conditions as opposed to independent conditions. This has been experimentally assessed in whole cells for the R303C mutation, but not for WT or R426H. Our present study provides quantitative analysis of this scenario in vitro, complementing earlier in vitro or cell extract studies reporting the percentage drop in activities when only one substrate is present. Thus, with the whole kinetic picture available, it is possible to determine if the S-Ado/SAICAr ratio could be due to intrinsic properties of mutant ADSL. If the amount S-Ado to SAICAr in body fluid were due to mutations of ADSL, there would be an inverse relationship between the ratio of activities (SAMP activity/SAICAR activity) and the accumulated substrates (SAMP/SAICAR) within the cell. Using the mixing kinetics model, it is possible to predict the intracellular concentrations of SAMP and SAICAR. R426H causes a S-Ado/SAICAr ratio of $\sim$1. From the prediction curves, we would expect the SAMP/SAICAR ratio to be 13:10 where the SAMP and SAICAR activities are equal. Whereas R303C causes a S-Ado/SAICAr ratio of $\sim$4. This corresponds to a SAMP/SAICAR ratio of 2:5 where the SAMP activity is four times less than SAICAR. These are quite different predicted intracellular ratios. Quantitative accumulation of these substrates has not been studied in mammalian cells [34, 71]. Normal levels of purine nucleotides and variable levels of residual enzyme activity from various tissues have been reported in tissues of ADSL deficiency patients [37, 88]. However, human WT and mutant ADSL were notoriously unstable at the time of the study.

One of the leading hypotheses for the pathogenic mechanism of ADSL deficiency is the toxic effect of accumulating succinylpurines, specifically SAICAR.

Therefore, it is important to understand how SAICAR might accumulate and to understand the consequences of this accumulation. Similarly, it has been hypothesized that the accumulation of SAMP might ameliorate the toxic effects of SAICAR accumulation. The observation that both SAICAR and SAMP accumulate in fibroblasts derived from an ADSL patient homozygous for the R303C mutation is consistent with this hypothesis.

Recently, SAICAR has been shown to activate pyruvate kinase M2 (PKM2) under conditions of limiting glucose, thus enhancing energy (ATP) production [43]. SAMP also activates PKM2 but at a much higher concentration. PKM2 is expressed in human brain and glucose uptake in the brain of ADSL deficient patients may be reduced [71]. Interestingly, ADSL deficient cells were shown to produce more ATP and higher pyruvate kinase activity than cells with normal levels of ADSL [43]. PKM2 has been considered to be present primarily in cancer cells, embryonic tissues, and in cells with a high proliferation rate [59]. However, there is evidence that PKM2 is the prominent isoform in most tissues [9]. If PKM2 is present in most tissues, then the relative amounts of accumulation of excess SAICAR and SAMP may cause aberrant regulation of PKM2 leading to toxic effects. Different relative amounts may regulate PMK2 differently, which could play a role in the heterogeneity of ADSL deficiency. Alterations in the activity of PKM2 in embryonic tissues may also contribute to defects in brain maturation and overall development in early stages of life when purines are mainly supplied by purine biosynthesis [1, 91]. If aberrant glucose metabolism occurs because of SAICAR accumulation and leads to toxicity, it may be possible to ameliorate this effect by modulating glucose metabolism

or by lowering accumulation of SAICAR. Clearly the ability to control the amount of SAICAR activity would be useful to maintain a low intracellular SAICAR concentration. If enzyme activity is measured in healthy individuals, then equation 2.4.1 and 2.4.2 could be used to introduce appropriate inhibitors to alter the concentration of the substrates in the cell to gain desired activities for SAMP and SAICAR.

2.8 Conclusions

In this study, we provide a model that predicts the dependency of product rate formation (dP/dt) as a function of substrate mixing that was verified by experimental HPLC-EC measurement for WT ADSL and two clinically associated mutations, R303C and R426H. This strongly suggests competition due to resource sharing mediates indirect communication between the two substrates. This hidden coupling causes a non-linear dependency of the activities on the substrate ratios. Electrochemical detection is a robust method for this purpose because of its ability to detect low concentrations of chemical compounds and two degrees of separation: retention time and electrochemical profile. From in vitro results, it is clear that mutations behave differently when mixtures of substrates are present. R426H retains a constant relative activity ratio while R303C does not. We propose that making inferences about intrinsic properties of ADSL should be based on how it functions in the presence of both substrates as opposed to two isolated activities may more accurately reflect the in vivo situation. Currently, the only mutation that has been experimentally verified to accumulate succinylpurines in whole cells is R303C while accumulation in

R426H was undetectable. HPLC-EC could provide a more sensitive assay to quantify intracellular concentrations of succinylpurines. The prediction model could then be used to calculate ADSL relative activity to determine if the S-Ado/SAICAr ratio is caused by inherent properties of mutant ADSL.

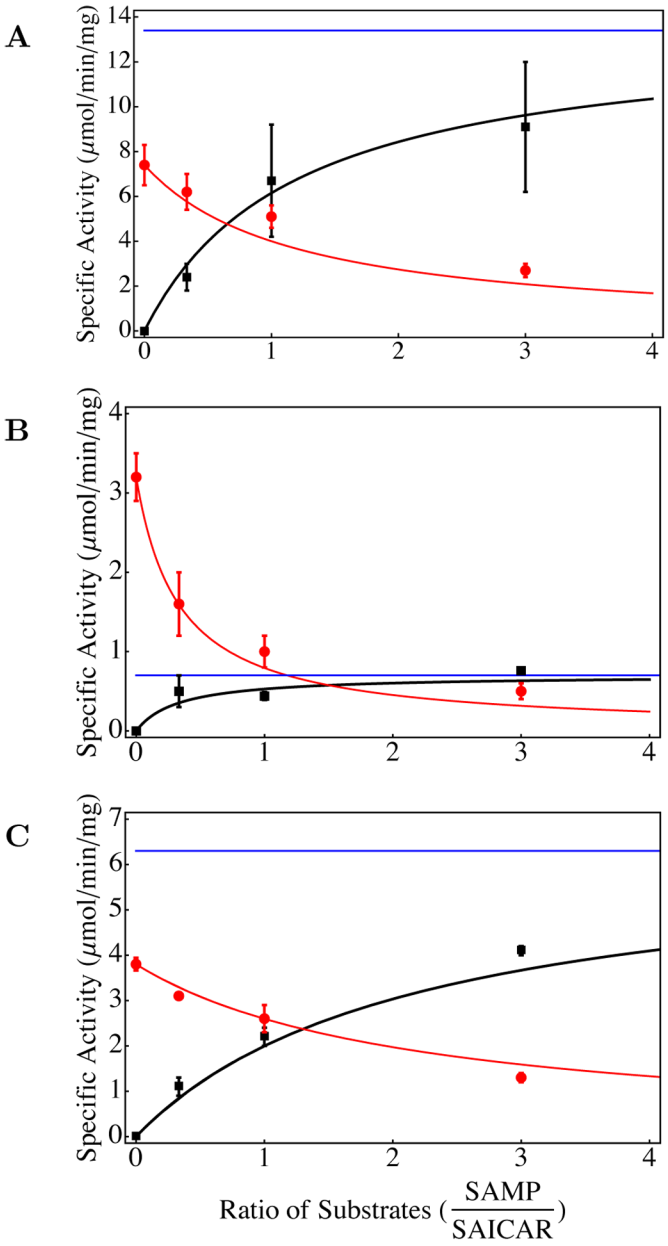


Figure 2.4: Graphical representation of mixing kinetic data from HPLC-EC analysis and prediction curves: A) WT, B) R303C, C) R426H. Plots are specific activity of SAICAR (circle, red) and SAMP (square, black) vs. substrate ratio (SAMP/SAICAR). The V_{max} value for SAMP activity is depicted by a blue line. Plots show a non-linear dependence of product rate formation vs substrate ratio.

Chapter 3

Structural and Biochemical Characterization of Human Adenylosuccinate Lyase (ADSL) and the R303C ADSL Deficiency Associated Mutation

3.1 Introduction

Crystal structures of the bacterial ADSL and other studies have provided insight into the active site and the catalytic mechanism; however these studies predominately use bacterial strains of ADSL, such as *Thermotoga maritima* or *Escherichia coli* [82, 83]. *T. maritima* and *Homo sapiens* have a sequence identity of 25% and similarity of 57%. *E. coli* and *H. sapiens* have a se-

quence identity of 23% and a similarity of 57%. ADSL was found in these bacterial species to function in the pathway as a homotetramer (Figure 3.1). Interestingly, three monomers contribute to each of the four active sites in the tetramer [11, 77]. The use of *Bacillus subtilis* ADSL in particular as a model to study properties of mutant enzymes in some cases involved modification of additional amino acid residues in the *B. subtilis* ADSL to more closely resemble human ADSL [62, 76]. However, attempts to make use of a *B. subtilis* model system to replicate the R303C phenomenon have proven difficult because *B. subtilis* and *H. sapiens* only have a sequence identity of 30% and a similarity of 63%. Specifically the sequence diversity of *B. subtilis* and *H. sapiens* results in multiple alterations within the active site to include but not limited to an asparagine residue in place of the corresponding R303 in human ADSL as well as an arginine residue in place of human corresponding T354. Not surprisingly, the removal of *B. subtilis*'s asparagine at the corresponding human ADSL R303 position for cysteine did not generate a disproportional reduction in catalytic ability [62, 76]. This observation as well as others suggest that the study of the bacterial enzyme, while useful for understanding some of the basic features of ADSL, may not be as useful for understanding the effects of various disease causing mutations on the human enzyme. Only limited structural information exists of human ADSL active site in the unpublished SAMP bound and S-AMP/AMP-fumarate bound PDB deposited entries (2J91, 2VD6). Current studies on disease-associated mutations of human ADSL have focused on correlation of substrate activity with clinical phenotypes, thermal stability of ADSL, activity of hybrid WT ADSL and mutants, and global changes in structure [3, 4, 44, 67, 98]. They have not inves-

tigated local structural changes or binding properties to the products within the ADSL active site. In this study, we present structural and biochemical characterization data of WT and mutant R303C ADSL by enzyme kinetics, product binding by isothermal titration calorimetry (ITC), and X-ray crystallography to reveal the effects of the R303C mutation that results in nonparallel reduction in enzyme activity. This work was completed in collaboration with the Pegan Lab at the University of Denver. Michelle Deaton completed X-Ray Crystallography. Glenn Capodagli collected Isothermal Titration Calorimetry data. Lucas Sawle performed energy minimization for SAICAR active site modeling.

3.2 Peculiarity of R303C Enzyme Kinetics

To ensure that the previously observed reduction in activity of the ADSL possessing the R303C mutation was not due to aberrant tetramer formation or outright global instability of the mutated protein in solution, static light scattering was employed to determine the polymeric distribution of the R303C and WT ADSL. As expected, WT ADSL is predominantly found to be a tetramer, 92.7 $\pm$ 2.6%, with aggregates contributing the final percentage. Similarly, R303C ADSL is predominantly a tetramer, 94.3 $\pm$ 0.1%, with aggregates contributing the final percentage. Observed weights of the WT and R303C ADSL tetramers were 225.0 $\pm$ 2.9 kDa and 214.5 $\pm$ 12.2 kDa, which is in close agreement with the His-tagged WT and R303C ADSL predicted tetrameric weights, suggesting that the R303C mutation results in no degradation of the ADSL

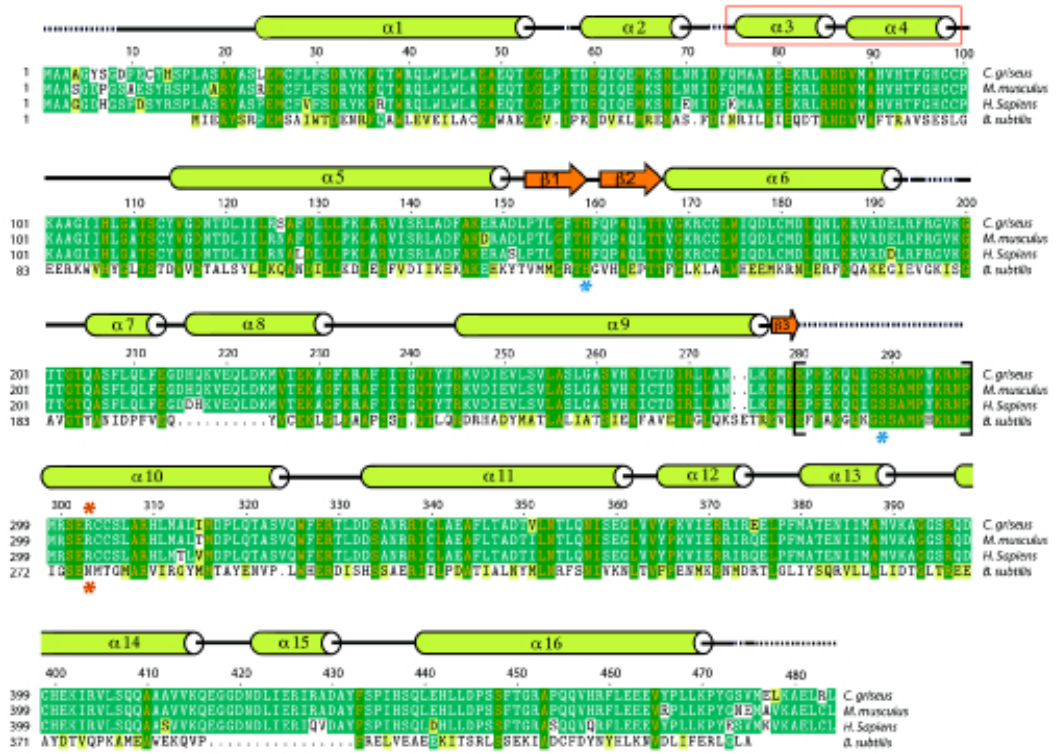


Figure 3.1: Sequence alignment of ADSL from various species. ADSLs are from *H. sapiens* (GenBank AAC83935.1), *M. musculus* (GenBank AAB60684.1), *B. subtilis* (NCBI Reference Sequence: YP 003865018.1), and C. griseus (NCBI Reference Sequence: NP 001230974.1). Secondary structure of hADSL according to Defined Secondary Structure of Proteins (DSSP) is represented by green cylinders (helical regions), orange arrows (β-sheet regions), black lines (loops), and black dashes (unstructured regions). Mutation site 303 is marked with an asterisk. The location of helices α3 and α4 is highlighted with a red box. The C3 loop is enclosed in black brackets. Catalytic residues H159 and S289 are marked with blue asterisks. Domains 1, 2, and 3 are indicated with bars in shades of brown.

tetramer.

Previously, a report noted that the ADSL carrying the R303C mutation displayed non-parallel decrease in activity at a single concentration [98]. Addi-

tionally, the kinetic parameters for human ADSL utilization of SAICAR have never been fully determined. Interestingly, only one earlier report, Stone et. al. 1993, details a K_m and k_{cat} of ADSL for SAICAR [79]. Unfortunately, the boundaries of ADSL were not well established at the time of that study, resulting in their use of a 25 amino acid N-truncated version of ADSL and not accounting for cooperativity. A subsequent study by Kmoch et al. also did not account for ADSL cooperativity nor did they report on the R303C mutant [44]. To further investigate the R303C phenomenon and determine the K_m and k_{cat} of ADSL for SAICAR, kinetic assays were performed on both WT and R303C ADSL using SAICAR and SAMP as substrates [98]. Both WT and R303C ADSL had measurable activity that could be evaluated by monitoring the UV absorbance of either SAMP or SAICAR. As WT ADSL was previously reported not to follow simple Michaelis-Menten kinetics for SAMP and perhaps be cooperative, the Hill equation was initially employed for calculating WT ADSLs kinetic parameters (Table 3.1; Figure 2.1) [3]. The resulting Hill coefficients of WT ADSL for SAMP and SAICAR are both significantly above one, indicating cooperativity for the substrates. The resulting k_{cat} and $K_{0.5}$ of WT ADSL for SAMP was in agreement with a prior study that utilized the full ADSL protein [3]. For WT ADSLs utilization of SAICAR, its $K_{0.5}$ was 1.8 $\pm$ 0.1 μM, which is only slightly higher than SAMPs. However, WT ADSLs k_{cat} for SAICAR was 1.6 fold higher than that for SAMP suggesting that ADSL is more kinetically efficient for the catalysis of SAICAR over SAMP. This is confirmed by WT ADSLs $k_{cat}/K_{0.5}$ for SAICAR being 2 fold higher than SAMPs. Additionally, the k_{cat} for both substrates was found not to be diffusion limited.

Table 3.1: Enzymatic Activities of ADSL with SAMP and SAICAR at 25°C

	SAMP		SAICAR	
	WT	R303C	WT	R303C
$k_{cat,H}$ (1/sec)	52 ± 2	2.33 ± 0.07	90 ± 2	27.8 ± 0.8
$k_{cat,H}$ (%)	100 ± 3	4.5 ± 0.1	100 ± 2	31 ± 1
$K_{0.5}$ (μM)	2.1 ± 0.2	2.3 ± 0.1	1.8 ± 0.1	8.4 ± 0.8
Hill Coeff.	1.5 ± 0.2	1.03 ± 0.08	1.22 ± 0.09	0.94 ± 0.06
$k_{cat,H}/K_{0.5}$ ($\text{sec}^{-1}\mu\text{M}^{-1}$)	25	0.99	50	3.4
$k_{cat,M}$ (1/sec)	N/A	2.36 ± 0.04	N/A	27.2 ± 0.4
$k_{cat,M}$ (%)	N/A	4.5 ± 0.1	N/A	30 ± 1
K_m (μM)	N/A	2.3 ± 0.1	N/A	8.4 ± 0.8
$k_{cat,M}/K_m$ ($\text{sec}^{-1}\mu\text{M}^{-1}$)	N/A	0.99	N/A	3.5

The k_{cat}, $K_{0.5}$, K_m, and Hill coefficient were determined by varying substrate concentration and fitting data to the Hill Equation (H) or the Michaelis-Menten Equation (M) in Sigma Plot. Protein was reconstituted for 2 hours at 25°C prior to measurements. R303C activity is relative to % WT. The values are shown along with their standard errors. N/A stands for not applicable.

Introduction of the R303C mutation into ADSL has multiple effects. Initially expecting similar cooperativity observed in WT ADSL, the Hill equation was utilized (Table 3.1). Unlike WT ADSL, the R303C ADSL demonstrates no cooperativity, with a Hill coefficient for both substrates of one. As a result, the simple Michaelis-Menten equation was employed for R303C. The R303C ADSL mutation reduces k_{cat} to 4.5% and 30% of that relative to WT ADSL for SAMP and SAICAR, respectively. Interestingly, this 7-fold difference in residual in vitro activity is in line with cell free extract studies of a Type II patients fibroblasts carrying the R303C mutation, which had 3% of normal SAMP activity and 30% of normal SAICAR activity [44, 87]. As a result, the bias of ADSL for SAICAR:SAMP is 1.7, which is in agreement with a previously published ratio of 1.6, but shifts to that of 11.5 in R303C ADSL [4]. Interestingly, K_m values of R303C ADSL increased for SAICAR compared to that of WT ADSL. The change observed in K_m of R303C ADSL for SAMP is almost negligible, with the K_m for SAICAR increasing more than 4-fold. This may initially suggest the mutation affects the ability of R303C ADSL to bind SAICAR more than SAMP resulting in the disparity in their cleavage. However, the rate of catalysis of R303C ADSL for SAICAR is 5 times larger than SAMP at R303C ADSLs K_m of SAMP. Therefore, the effect of the R303C mutation on ADSL may not solely be reflective in its divergent ability to bind SAICAR and SAMP. It suggests that R303C may affect the catalytic mechanism of SAMP and SAICAR in addition to binding of the substrates.

3.3 SAICAR Forms an Extra Hydrogen Bond Compared to SAMP

To explore the possible thermodynamic factors involved in the non-parallel reduction in activity resulting from the R303C mutation and in general ADSL-substrate interactions, ITC was performed on WT and R303C ADSL with their products AMP and AICAR. By measuring the amount of heat liberated per injection as a function of the molar ratio of the substrate and protein, thermodynamic parameters for R303C and WT ADSL's interaction with their products were calculated (Table 3.2; Figure 3.2). Intriguingly, product binding does not show any evidence of cooperativity, as the data fit to an independent model rather than a multiple binding site model. This does not necessarily rule out the possibility that the substrate binding is cooperative, but suggests that cooperativity may require inclusion of interactions between the fumarate component of the substrates and ADSL, which is lacking in the products.

Beyond the absence of cooperativity observed, the K_d values for WT ADSL were 54 and 34 μM for AMP and AICAR, respectively (Table 3.2; Figure 3.2). The thermodynamic properties that are responsible for AICAR and AMP binding to WT ADSL globally are similar with the enthalpic component, ΔH, being the predominant driving force and the entropic component, ΔS, being unfavorable. Although similar, the thermodynamic properties are not identical. AICAR binding liberates an additional 18 kJ/mol, suggesting an additional hydrogen bond may be formed between AICAR and WT ADSL that is absent in a AMP and WT ADSL complex. The lack of this additional

enthalpic contribution is partially offset in the AMP and WT ADSL complex by the reduced unfavorable entropic component compared to the AICAR and WT ADSL complex. Overall, this leads to a difference in ΔG of AMP and AICAR to WT ADSL to be only 1.2 kJ/mol in AICARs favor.

Intriguingly, the R303C mutation impacts the ability of ADSL to form a complex with AMP and AICAR in two distinct manners. The enthalpic component for both products decreases by a similar degree of $\sim$30 kJ/mol suggesting the loss of one to two hydrogen bonds, whose identity could be common between the two products and ADSL. Additionally, the R303C mutation shifts the ADSL complex formation with AMP and AICAR from an unfavorable to favorable entropic event. Taken together, these thermodynamic factors result in a similar decrease of ΔG and increases of 76 and 88 μM in the K_d of AMP and AICAR respectively. As a result, the R303C mutation appears to negatively affect ADSLs interaction with AMP and AICAR almost indistinguishably.

3.4 Origin on ADSL Cooperativity

The elucidation of the ADSL-SAMP structure has previously been suggested to support cooperativity in human ADSL, as AMP and fumarate are found in two active sites and with the other two occupied by SAMP. However, previously limited biochemical data existed to support the origins or form of the cooperativity found in ADSL. Interestingly, the kinetic and thermodynamic properties of the WT ADSL and the R303C mutant suggest that

Table 3.2: ITC Thermodynamic Parameters of ADSL at 20°C

	AMP		AICAR	
	WT	R303C	WT	R303C
K_d (μM)	54 ± 3	130 ± 8	34 ± 2	121.9 ± 0.9
ΔH (kJ mol^{-1})	-35 ± 2	-5.9 ± 0.2	-53 ± 1	-16.7 ± 0.5
$T\Delta S$ (kJ mol^{-1})	-12 ± 1	15.92 ± 0.06	-28 ± 1	5.3 ± 0.5
ΔG (kJ mol^{-1})	-23.9 ± 0.1	-21.8 ± 0.2	-25.1 ± 0.1	-21.95 ± 0.02
n	0.99 ± 0.01	1.01 ± 0.04	1.00 ± 0.02	1.00 ± 0.02

Data sets were collected in duplicate and analyzed with NanoAnalyze and fit to an independent model concurrently with a blank constant model to adjust for the heat dilution. All measurement were from 25 injections of 2.5 mM AMP into 170 μL of protein in 25 mM HEPES (pH 7.0), 150 mM KCl, and 2 mM DTT at 20°C

the involvement of all three monomers that comprise an ADSL active site is necessary to achieve ADSLs positive cooperativity. Specifically, no cooperativity was observed when binding AICAR, or AMP, to WT ADSL. Binding of these products form the majority of their interactions with only two of the monomers. For example, in active site 1, AMP interacts with six residues from monomer A, and one from monomer B. Additionally, the phosphate group of AMP is anchored with two additional residues from a relatively immobile portion of monomer D. (Figure 3.3a). This leaves the fumarate product to span interactions between monomer As mobile α-helices 2-4, monomer B, and the near α3-α10 catalytic loop region of monomer D, suggesting that it is necessary for cooperativity to be observed. In line with this hypothesis, mutation of R303 to cysteine, which removes two hydrogen bonds, thus removing the majority of the interactions between monomer D and SAMP, results in no

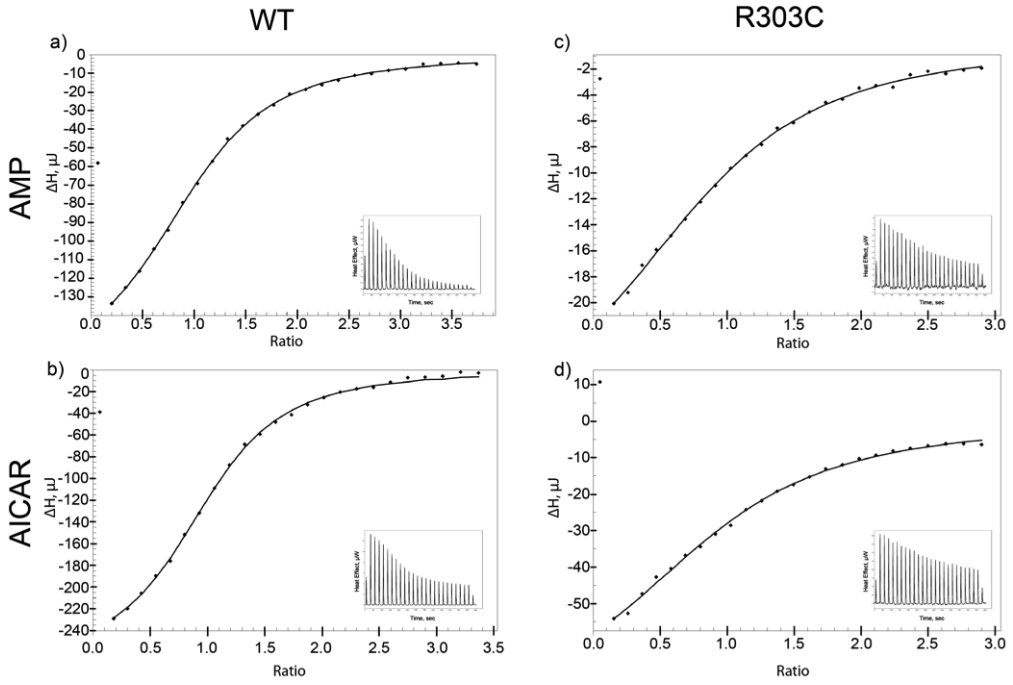


Figure 3.2: Calorimetric titration of ADSL with AMP and AICAR. Integrated heat peak areas against the molar ratio of AMP (a and b) and AICAR (c and d) added to ADSL-WT and ADSL-R303C. The inset shows raw heat data gained from 25 injections of 2.5mM AMP into 170μL protein in 25 mM HEPES, pH 7.0, 150 mM KCl, 2 mM DTT at 20°C. Data sets were collected in duplicate and analyzed with NanoAnalyze software and fit to an independent model concurrently with a blank constant model to adjust for heat of dilution. The line shows the best fit to an independent model. Figures generated using the NanoAnalyze Software provided by TA Instruments.

cooperativity (Figure 3.3b).

Beyond the necessity for the substrate, or products, to involve all three monomers when binding to achieve an allosteric response, comparison of the WT-ADSL-apo and ADSL-SAMP active sites provided a glimpse into the model of allosteric regulation that ADSL undergoes. Specifically, the WT-ADSL-apo active sites are considerably more open to the bulk solvent then

those of the ADSL-SAMP and ADSL-AMP structures (Figure 3.4a). The 0.27 nm shift of α-helices 2, 3 and 4 observed between the WT-ADSL-apo active sites and those bound with AMP, or AMP and fumarate, illustrates a clamping down of the active site over SAMP, or AMP. With AICAR and SAICAR having homologous phosphate, fumarate, and ribose moieties to that of AMP, the same shift is likely to occur upon the binding of these ligands as well. This closure of the active site around SAMP suggests that a concerted model of allosteric regulation would not be possible (Figure 3.4c). In other words, if this shift occurred in a neighboring active site as a result of binding SAMP and SAICAR for the reaction running in the forward direction, SAMP attempting to bind to the neighboring active site would be sterically impeded. This would not be reflective of the positive cooperativity observed. Applying sequential allosteric regulation to the reverse reaction is less certain. The binding of AMP and AICAR would also be sterically impeded from binding to an active site that already underwent a shift of α-helices 2, 3 and 4. However, as illustrated by the thermodynamic information, the binding of AMP and AICAR is does not illustrate cooperativity. With fumarate binding at the top of the active site, which is exposed to the bulk solvent, a concerted model cannot be ruled out in the reverse direction (Figure 3.4b).

Beyond the sequential model of allosteric modulation of ADSL, in what order, or the extent to which the active sites fill with substrates remains an open question. The ADSL-SAMP structure has its two active sites on the same end of ADSL occupied by SAMP and other distal active sites bound with fumarate and AMP. The absence of SAMP in all of the active sites could

be due to restraints of the enzyme imposed by the crystal lattice, but alternatively might propose that substrates bind to adjacent active sites first, or only can bind to them. The latter would infer that the active sites at one end of the ADSL tetramer communicate to its distal neighbors. Although the mild loss of cooperativity observed by Ariyananda et al. [3] by an ADSL mutation L311V at the center of the tetramer would support this possibility, additional structurally guided mutagenesis along monomers interface and molecular dynamics simulations will be required to completely tease out the global allosteric nature of ADSL.

3.5 Proposed Catalytic Mechanism

The loss of cooperativity because of the R303C mutation removal of hydrogen bonds with one of the three monomers comprising the active site as reflected in the structures of ADSL-SAMP, or ADSL-AMP, would not itself explain the divergent catalytic properties of ADSL observed towards SAICAR and SAMP. However, the combination of the kinetic and thermodynamic data coupled with the structural evidence suggests that the R303C mutation illuminates SAICAR as a better substrate for ADSL. At first glance, the unparallel reduction in k_{cat} between SAMP and SAICAR upon the R303C mutation could suggest that the divergence in catalytic ability of ADSL between SAMP and SAICAR is due to a lack of substrate binding that favors SAICAR, or AICAR binding over that of SAMP. However, the K_m for SAICAR is slightly elevated relative to that of SAMP indicating that binding of SAICAR might be slightly weaker (Table 3.1). Moreover, the ADSL ligand bound structures

have R303 forming hydrogen bonds to the phosphate group that is conserved between SAICAR and AICAR (Figure 3.3). Also, the thermodynamic data reflects that the mutation causes a loss of enthalpy at almost equal magnitude for both AICAR and AMP, reflective of a loss of the same ADSL-phosphate group interactions. Despite the loss of the bond to the phosphate group, the overall ΔG_{bind} for both products remains similar to WT ADSL. This is accomplished by shifting the product binding from an unfavorable to favorable entropic event. The resulting K_d's for both substrates are nearly equally decreased by approximately 2.5-3 fold. This near equal drop in K_d but not k_{cat} suggests that the R303C mutation is indirectly affecting the catalytic ability of ADSL divergently for its two substrates. In other words, this thermodynamic, kinetic, and structural evidence suggests that the substrate can bind to the active site at a significant affinity despite the R303C mutation. As a result, unlike WT ADSL, the R303C mutant is likely operating at rapid equilibrium state, where if the substrate binds to ADSL, most substrate will dissociate, and only a small amount will be converted to product.

This would suggest that ADSL is more efficient at cleaving the fumarate-AICAR/AMP bond of SAICAR than SAMP. The inequality of catalysis might suggest that SAICAR and SAMP do not follow the same mechanistic pathway. The reaction mechanism for conversion of SAMP and SAICAR by ADSL has been previously described as a general acid-base mechanism resulting in β-elimination of fumarate (Figure 3.5) [83]. Although, the conversion of SAICAR to AICAR and fumarate could follow the same steps as the SAMP to AMP mechanism, due to the presence of an additional carbonyl in SAICAR, there is

another possibility. In this SAICAR selective pathway, the first two steps are the same: the C^{β}-proton is removed, and the negative charge is stabilized by the δ-carboxyl group. However, in the third step, as the electron density from the double-negatively charge carboxyl group is shifted down and the C^{α}-N6 bond is broken, a double bond is formed between N6 and the carbonyl carbon, and electron density is pushed up onto the carbonyl oxygen. Then, in the last step, this negative charge moves back down to form a C-O double bond as N6 becomes protonated by H159. This negatively charged oxygen may be stabilized by nearby positively charged residue, R235. This additional stabilization of the intermediate may be a factor in the faster reaction rate of SAICAR compared to SAMP.

An alternative possibility, or potential additional factor, for the inequality of bond cleavage between SAMP and SAICAR is the ability of SAICAR to be positioned with the active site in a more catalytically effective orientation than SAMP. The loss of the hydrogen bonds that R303 contributed to substrate binding potentially affected the ability to orient SAMP and SAICAR ideally for cleavage. Unlike SAMP, which possess a pyrimidine ring, SAICAR has a single rotatable bond that has been previously suggested to infer greater flexibility to adopt different conformations (36). This also might explain the lack of unparallel reduction in catalytic ability of *B. subtilis* ADSL. In that case, *B. subtilis* ADSL has an arginine in place of human ADSL T354 that may provide a similar influence. Unfortunately in the Palenchar et al. study that looked at *B. subtilis* ADSL corresponding human ADSL R303C mutation (*B. subtilis* ADSL N276C), only performed a double mutation to mimic the

human active site configuration and did not include a single mutation for the corresponding arginine at human ADSLs T354 [62]. This prevented observing whether elimination of anchoring ADSLs substrates by a third arginine resulted in unparallel catalytic ability.

Additionally, the thermodynamic data suggests that AICAR forms an additional hydrogen bond not present in the ADSL-AMP complex. The enthalpic divergence between the two substrates remains after the introduction of the R303C mutation. The ability of SAICAR to possess an additional point of contact within the ADSL active site could ensure that SAICAR properly oriented within the active site more often then SAMP, contributing to the unparallel k_{cat} between SAMP and SAICAR observed. Based on the ADSL-AICAR-fumarate model, the additional hydrogen bond could be reflective of an interaction between AICAR and the highly conserved S334 (Figure 3.6a). Curiously, even in human ADSLs distant homologue of *B. subtilis*, this serine is conserved. This serine has not previously been implicated in substrate binding or catalysis, as it does not interact directly with SAMP or AMP within the ADSL-SAMP and ADSL-AMP structures. Naturally, additional site-directed mutagenesis efforts within the active site including S334, or an X-ray structure of ADSL with SAICAR, or AICAR, will be necessary to fully identify the ADSL residue acceptor that forms the additional hydrogen bond with AICAR and bring final clarity to the unparallel catalytic activity observed.

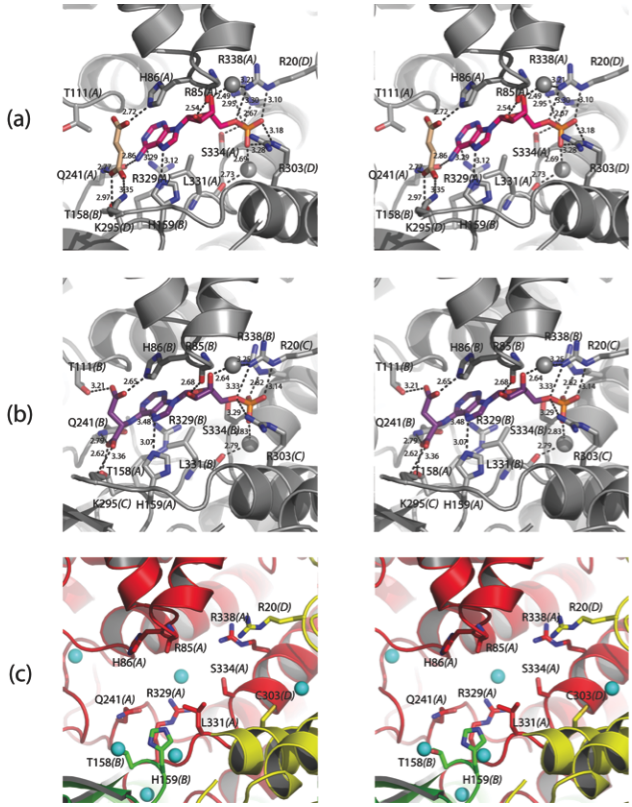


Figure 3.3: Comparison of ADSL active sites for WT, WT with substrate, and R303C mutant. (a) Wall-eyed stereo view of a ADSL-SAMP active site with AMP and fumarate occupying the active site. AMP is rendered pink and fumarate is rendered tan. Heteroatoms are colored according to their element. Water molecules (grey) are depicted as spheres and are scaled to 50% for clarity. The monomer each residue belongs to is given in parentheses after the residue number. (b) Wall-eyed stereo view of a ADSL-SAMP active site with SAMP occupying the active site. SAMP is rendered purple, and heteroatoms are colored according to their element. Water molecules (grey) are depicted as spheres and are scaled to 50% for clarity. The monomer each residue belongs to is given in parentheses after the residue number. (c) Wall-eyed stereo view of R303C ADSL active site. Monomers A, B and D are rendered red, green, and yellow, respectively. Heteroatoms are colored according to their element. Water molecules (cyan) are depicted as spheres and are scaled to 50%, and the side chain of R85 is hidden for clarity. The monomer each residue belongs to is given in parentheses after the residue number.

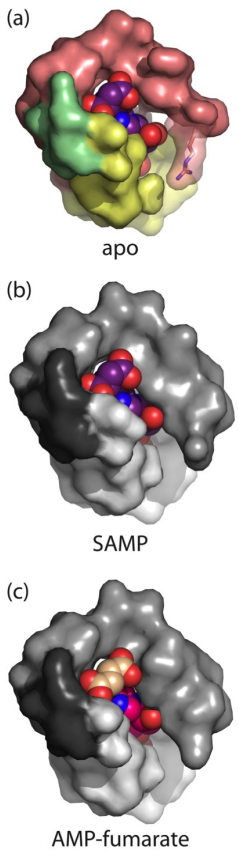


Figure 3.4: Constriction of the ADSL active site upon substrate binding. (a) Surface rendering of WT-ADSL-apo active site 1. WT active site depicted in salmon, light yellow and pale green corresponding to monomer A, B, and D respectively. R85 side chain is shown with transparency to reflect the lack of electron density for the side chain in all monomers. AMP and fumarate from active site 1 of ADSL-SAMP was placed in the apo active site for scaling purposes. (b) Surface rendering of SAMP and fumarate bound active site 1 of ADSL-SAMP. Coloring of light grey, medium grey, and dark grey correspond to monomers A, B, and D respectively. (c) Surface rendering of SAMP and fumarate bound active site of ADSL-SAMP. Coloring of light grey, medium grey, and dark grey correspond to monomers A, D, and C respectively. All waters were removed from surface renderings in (a-c).

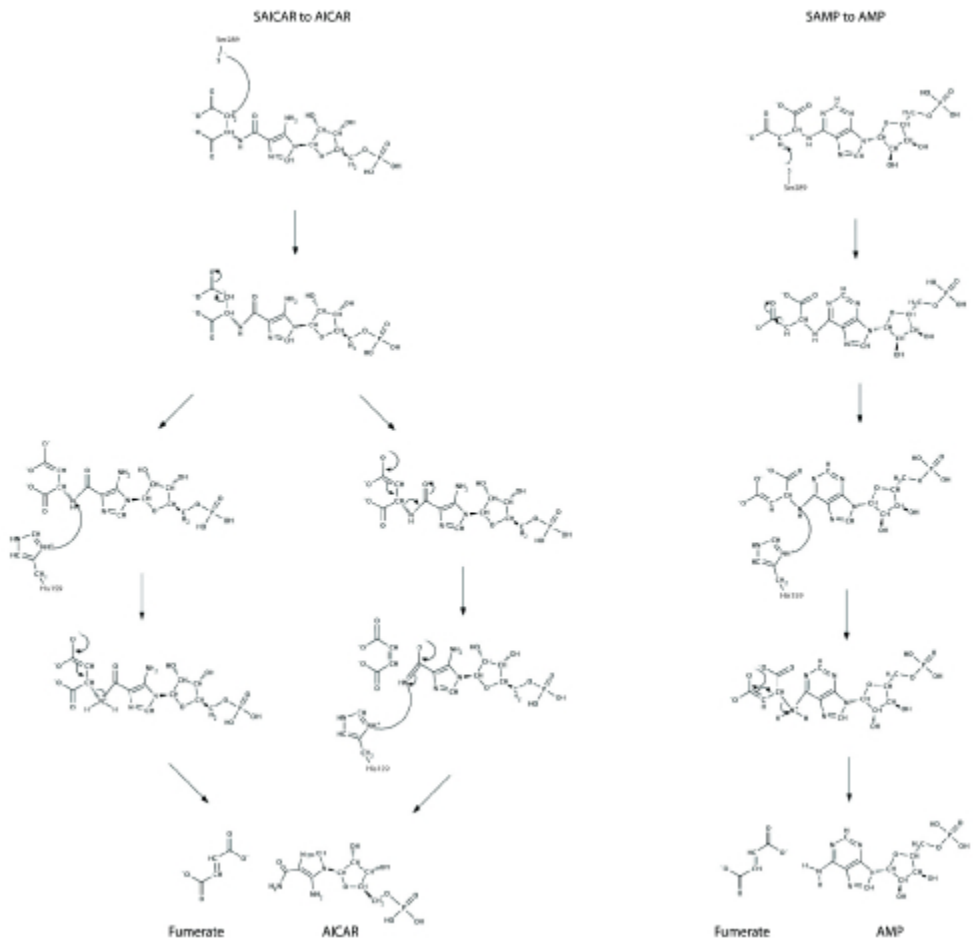


Figure 3.5: Catalysis mechanism of ADSL on SAICAR and SAMP

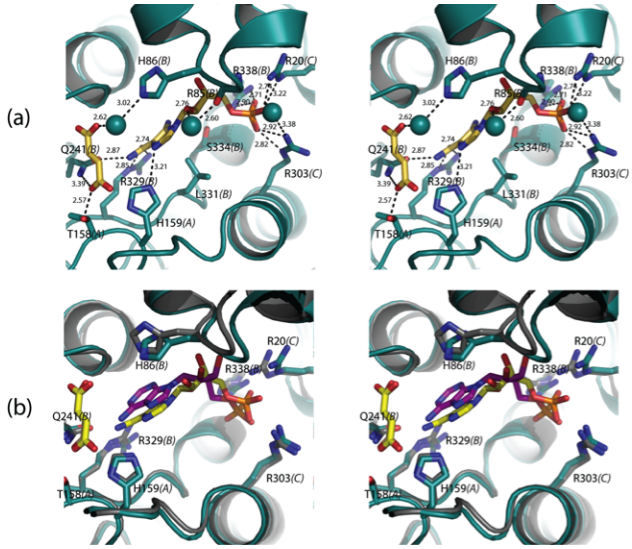


Figure 3.6: ADSL-AICAR-fumarate model. (a) Wall-eyed stereo view of model of WT active site with AICAR and fumarate occupying the active site. AICAR and fumarate are rendered yellow. Heteroatoms are colored according to their element. The monomer to which each residue belongs is given in parentheses after the residue number. (b) View of model of WT active site with AMP and fumarate overlaying model of WT active site with AICAR and fumarate. WT with AMP is rendered grey with AMP in pink and fumarate in tan, while WT with AICAR residues are teal with AICAR and fumarate in yellow.

Chapter 4

Protein Folding and Stability of ADSL and Potential Connection to Pathogenic Mechanisms

Protein denaturation, by definition, is characterized by the loss of enzyme activity due to an induced stress on the system. A common biochemical method to characterize protein stability is by thermal inactivation of the enzyme. Generally, this is how stability of ADSL has been studied [3]. Recently the use of blue native gels have been used to view whether the ADSL maintains its tetramer structure after incubation at higher temperatures [98]. While these studies are informative, they only provide an answer to the question, "Does the mutation cause instability or a loss of function at a raised temperature?" A mutated enzyme can appear to be in the same state at raised temperature, when in reality, it is in a different final folded state with residual activity. Therefore, a more interesting question to ask is, "How much

does the stability decrease (or increase) due to a clinically associated mutation at standard conditions?" The answer to this question requires a quantitative analysis of the free energy between states of the denaturation process. The change in free energy between states would provide insight as to where the protein ends in the folding funnel, whether it has reached its minimum energy state, or if it has folded into a less (or more) favorable state. The studies presented in the this chapter begin the process of revealing which states are present during ADSL denaturation (Figure 4.1), quantifying the free energy between those states, and determining how the extreme mutations of ADSL change the stability. This fundamental research can assist in directing therapeutics by suggesting whether the pathogenic mechanism could be due to a "loss of function" or "gain in toxic function" as a result of protein misfolding.

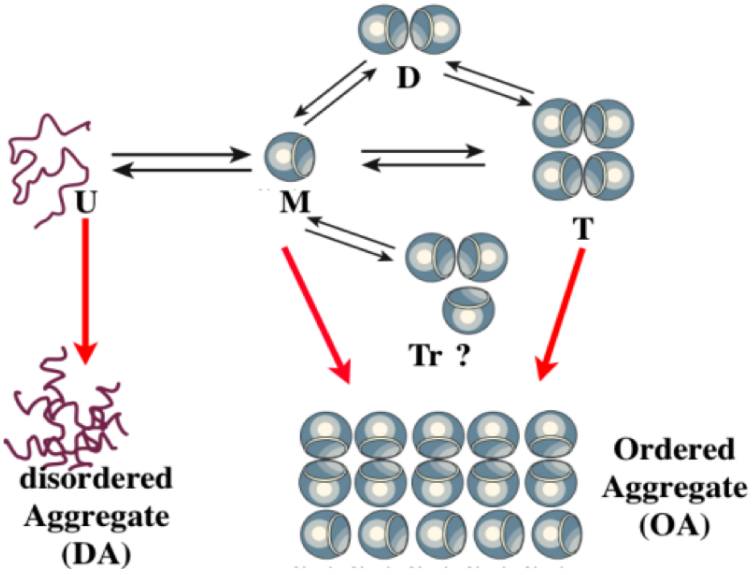


Figure 4.1: Schematic of the ADSL landscape. The are a myriad of different states available for a tetrameric protein.

Table 4.1: ADSL Population Distribution

Enzyme	Tetramer Weight (kDa)	Tetramer (%)	Aggregate (%)
WT	220.8 ± 0.2%	96.2 ± 0.6	3.6 ± 0.6
R303C	219.2 ± 0.2%	97.8 ± 0.1	1.8 ± 0.1
R426H	225.4 ± 0.2%	94.4 ± 1.6	5.6 ± 1.6

4.1 WT, R303C, and R426H Are Identical at Standard Conditions

Prior to denaturation experiments, it was verified that WT, R426H, and R303C begin in the same state with the same amount of secondary structure. Upon measuring the population distribution using Static Light Scattering, WT, R303C, and R426H were all found have similar distributions with tetramer as the dominant state and a small percentage of aggregates (Table 4.1). Circular Dichroism was then employed to measure the signal from secondary structure and found near identical curves for WT and mutant ADSL (Figure 4.2). Therefore, assuming the aggregation is negligible, a folded tetramer can be considered the final state for thermodynamic modeling.

The absence of aberrant tetramer formation initially rules out the "gain of toxic function" for these two ADSL mutations. Experiments such as sedimentation equilibrium may be useful to investigate this further. This merely rules out R303C and R426H causing a "gain in toxic function". It does not rule out "gain in toxic function" as a pathogenic mechanism for other clinically associated mutations.

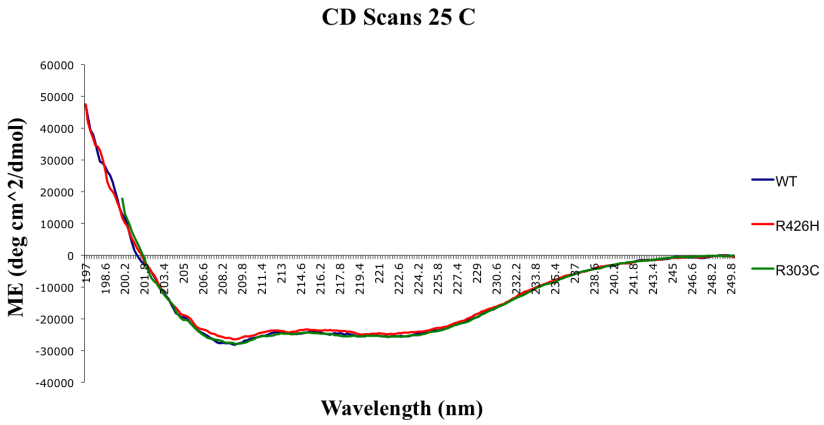


Figure 4.2: CD Spectra of WT, R303C, and R426H ADSL.

4.2 Protein Unfolding by Thermal Denaturation is Irreversible

Thermal denaturation of proteins is a powerful tool to use. From differential scanning calorimetry measurements, values of ΔG, ΔH, and ΔC_p can be quickly quantified by equation 4.2.1 and equation 4.2.2.

$$\Delta G(T) = \Delta H(1 - \frac{T}{T_m}) + \Delta C_p(T - T_m - Tln[\frac{T}{T_m}]) \tag{4.2.1}$$

$$\Delta S(T_m) = \frac{\Delta H(T_m)}{T_m} \tag{4.2.2}$$

DSC measurements were initially taken on a VP-DSC at the biophysics core at the University of Colorado Anschutz Medical Campus. Following the transition peak around 60 °C, instead of an expected rise in the baseline, there is a "crash" in signal at 65 °C that likely represents protein aggregation (Figure 4.3a). Unfortunately, as is the case with the majority of proteins, thermal denaturation for ADSL is irreversible. Defining a baseline does allow for characterization of the enthalpy needed to unfold the system up to a point of aggregation. This can be useful information in limiting parameters for other modeling methods.

The VP-DSC sample cell is referred to as a cylindrical cell. The protein molecules in solution are free to move throughout the entire cell as it is heated. TA Instruments' Nano DSC offers a new design called the capillary cell that limits the movement of proteins to increase the chance of reversibility. The unfolding transition of ADSL was resolved in the Nano DSC as there is a clear baseline following the transtion (Figure 4.4). It appears from the non-Gaussian shape that an intermediate form may exist or that the protein consists of two separate processes. With the information known about ADSL, it is likely that the slow initial pre-transition is the dissociation of the tetramer into its monomeric subunits and the unfolding of the protein is faster, as seen visually in the sharp post-transition. However, ADSL was also found the be irreversible in the Nano DSC as well. Information from this method confirmed the presence of an intermediate state. While it can be presumed to be tetramer dissociation, further investigation is needed to confirm this assumption. While the process is not reversible, since the unfolded transition was resolved, this

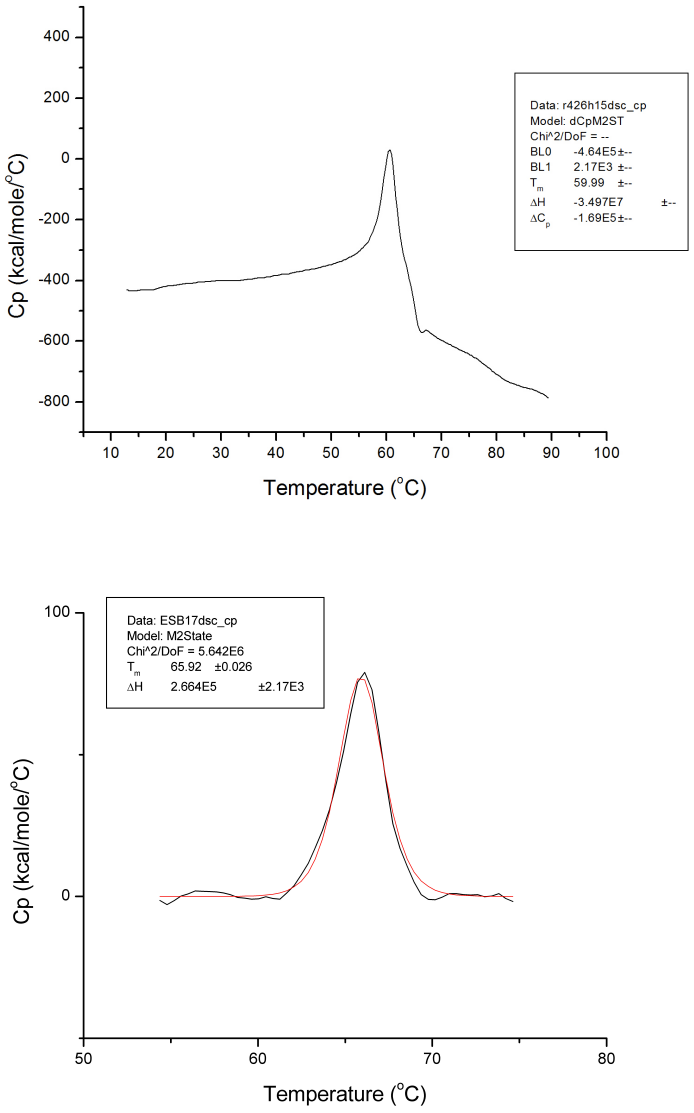


Figure 4.3: Representative plots for measuring excess heat capacity as a function of temperature with VP-DSC from MicroCal. As depicted in a) instead of an expected rise in the baseline after the melting temperature, there is a "crash" in signal that likely represents protein aggregation. Defining a baseline allows characterization of the enthalpy needed to unfold the system up to a point of aggregation.

provides better estimates for the enthalpy, heat capacity, and melting temperature of the ADSL than the VP-DSC. These values can be substituted into equation 4.2.1 for an approximate free energy of folding curve (Figure 4.5).

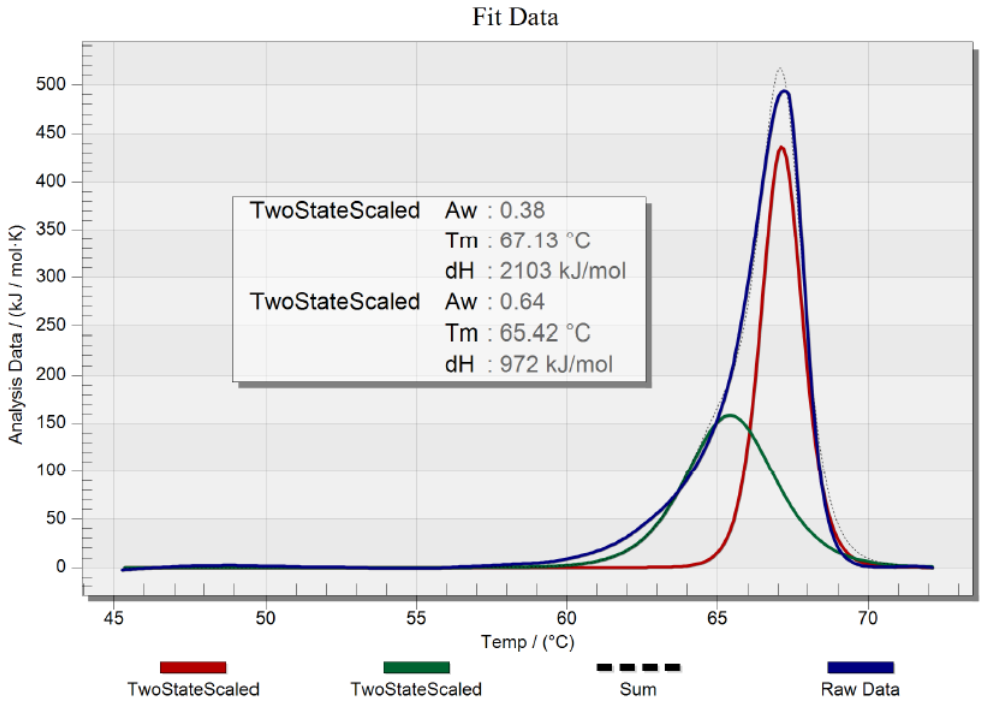


Figure 4.4: DSC curve for WT ADSL from TA Instruments Nano DSC.

Thermal denaturation by CD measurements for R303C and R426H were performed as well (Figure 4.6). Interestingly, R303C, the mild phenotype mutation, appears identical to WT while R426H has a much lower melting temperature. Although this is interesting, it still does not indicate whether R426H is less stable in standard conditions at physiological temperatures. The free energy of folding curve from DSC can be used to predict the fraction folded curve from CD. A three state model was used to calculate the prediction

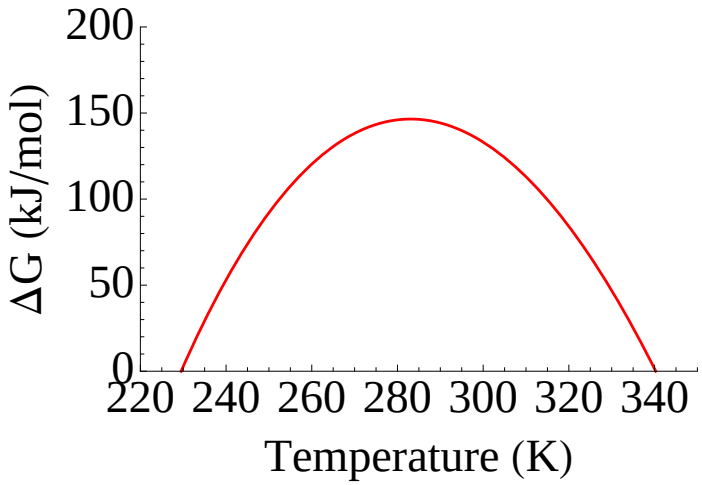


Figure 4.5: Free Energy of Folding vs. Temperature based on thermodynamic parameters values from the Nano DSC. This is only an estimate for the free energy curve since the unfolding process was resolved, but not reversible.

since it was suggested ADSL has one intermediate in the unfolding process. The prediction was compared to a fraction folded curve obtained by thermal denaturation from CD. The agreement between experiment and prediction is good as can be seen in Figure 4.7. Unfortunately, since the process is irreversible, we cannot make any conclusions regarding protein stability from thermal denaturation. If conditions are found where reversibility is achieved, it would pave the way for quick stability comparisons between WT and disease associated mutations.

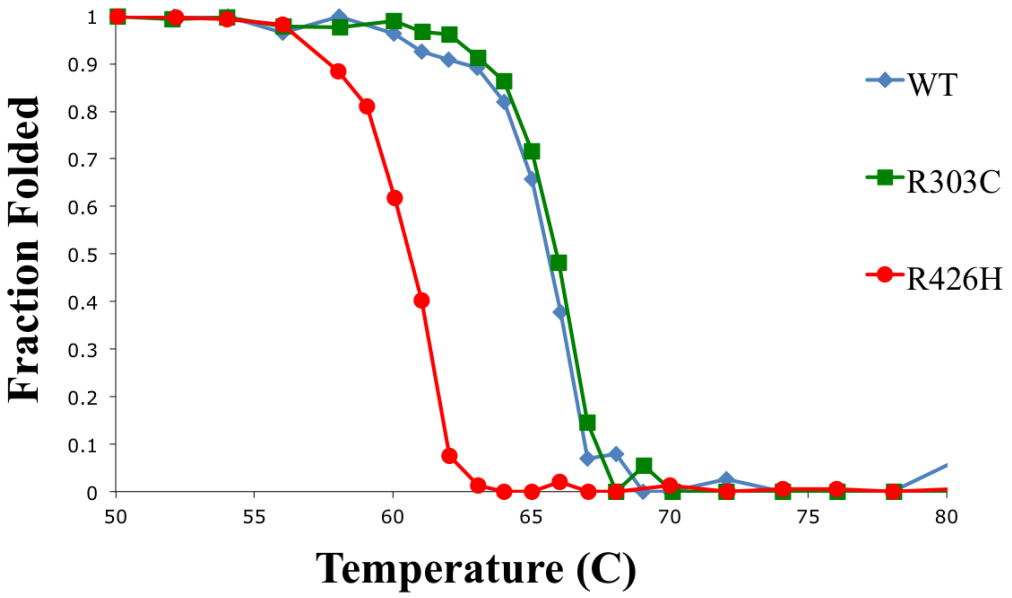


Figure 4.6: Thermal melting curves for WT, R303C, and R426H ADSL.

4.3 ADSL Unfolding is Reversible by Chemical Denaturation

Since it is clear that thermal denaturation, in the conditions tested, is not reversible, chemical denaturation was attempted next. Fraction folded curves for WT (a), R426H, (b), and R303C (c) exhibit a reversible process shown in Figure 4.8. Unfortunately, a complete refolding signal was not achievable. Refolding was tested by dialyzing protein at 6 M denaturant in 0 M denaturant buffer. The protein always precipitated out of solution, most likely due to the shock of being exposed to 0 M denaturant solution. It is noted that R426H shows some hysteresis from 1-2 M denaturant. There is no current explanation

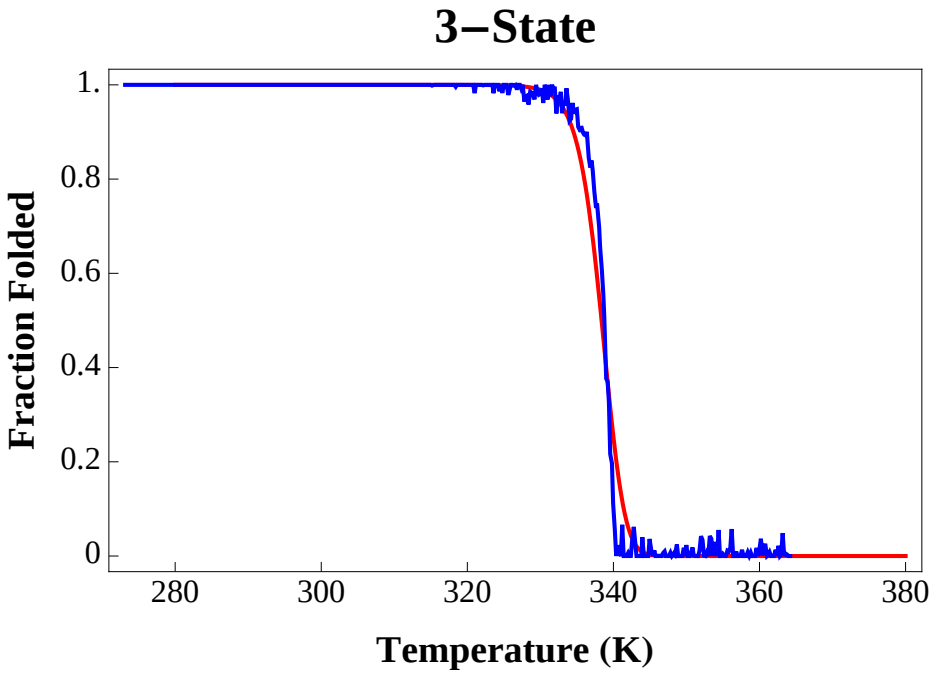


Figure 4.7: CD melting curve with 3-state fraction folded prediction. The blue line is experimental data calculated from CD. The red line is the prediction from DSC thermodynamic parameters.

for the hysteresis. Similarly to thermal denaturation, R426H appears to be less stable while R303C is very similar to WT.

4.4 ADSL Dissociates Into Monomers

The landscape of ADSL is very diverse. Starting from a folded tetramer, there are many different pathways that can be taken, meaning there are a number of states that could be occupied. Differential scanning calorimetry suggested at least one detectable intermediate is present. To detect what states are present, samples from the fraction folded curve were ran on blue native gels. The results for WT showed a main band at 0 M denaturant the

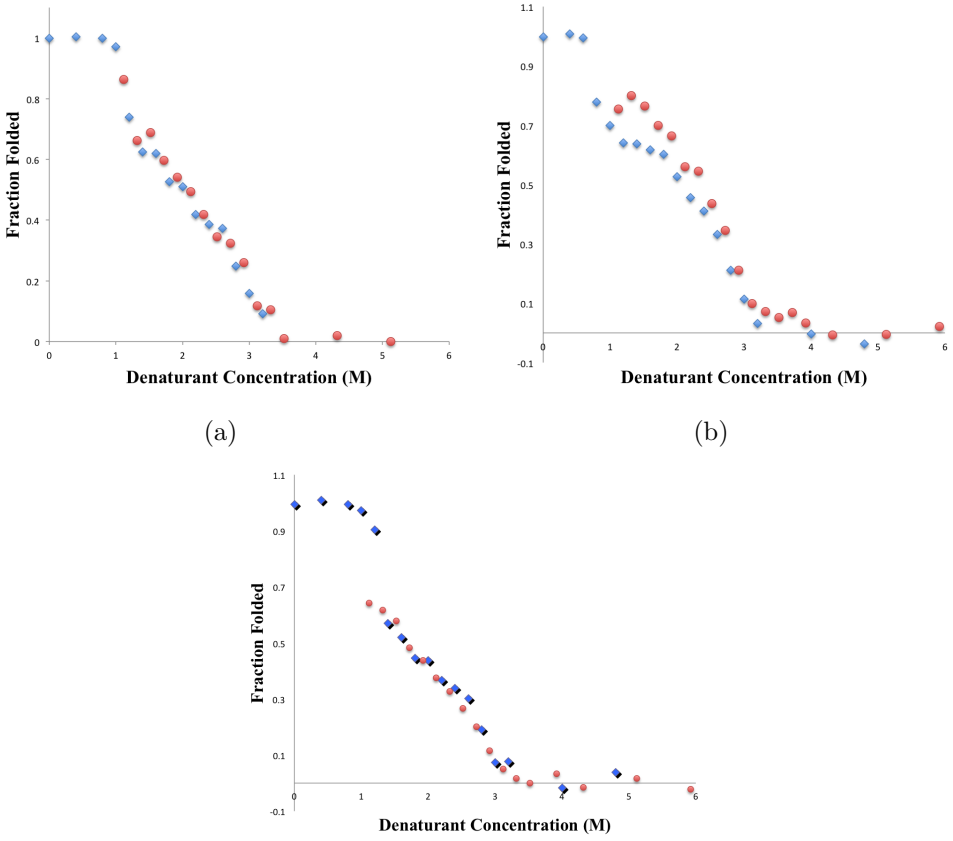
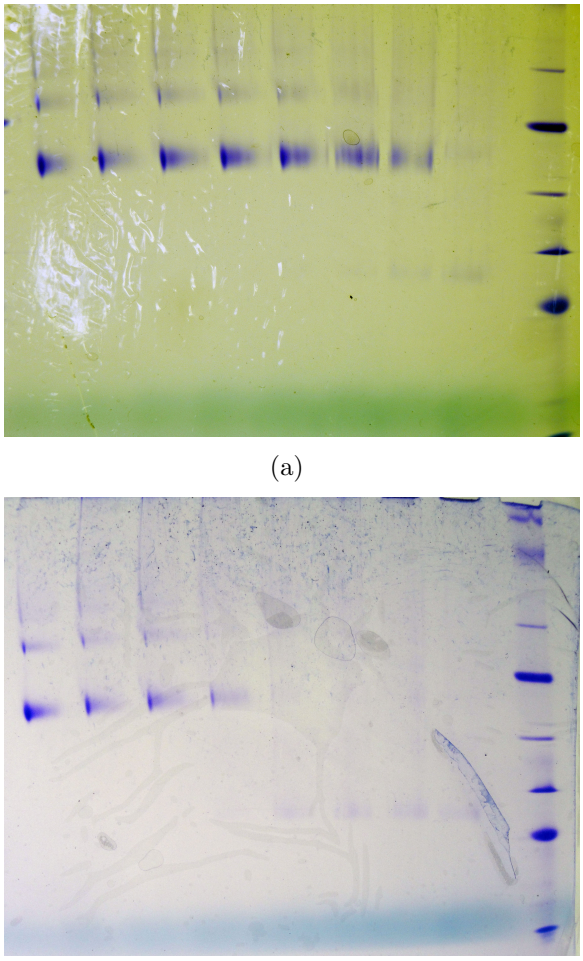


Figure 4.8: Chemical Denaturation plots of Fraction Folded vs. Denaturant Concentration (M) for WT (a), R426H (b), and R303C (c). Denaturation plots are in blue while renaturation plots are in red.

corresponds to the tetramer (Figure 4.9a). The first transition around begins at 1.2 M denaturant, in agreement with CD data. A faint, new band begins to appear half way down the gel as the main band fades and potentially starts to move upward on the gel. It is also noted that the new band decreases in intensity as denaturant concentration increases. A new band appearing, and the tetramer band fading, is interpreted as the tetramer unit dissociating into smaller subunits. A possible explanation for the low band intensity for the new band is that the large concentration of guanidinium chloride is not allowing

the stain to bind to the protein. This pattern is also true for R426H (Figure 4.9b). It is also confirmed in the blue native gels that the R426H mutation causes the tetramer to fall apart sooner.



(a)

(b)

Figure 4.9: Chemical denaturant experiments on blue native gels. Concentration of denaturant increases toward the right.

The molecular weight of the new band falls between two standards around the molecular weight of the ADSL monomer and dimer. The resolution was

not good enough to confirm it was one, or the other. Therefore, a secondary and more quantitative analysis using sedimentation velocity was used on WT ADSL to determine the state of the second band. It showed at 0 M, tetramers were mainly present and as denaturant concentration increased, molecules with molecular weight of 57 kDa began to appear. This experimentally shows the tetramer is indeed dissociating into monomers as denaturant concentration increases (Figure 4.10). Aggregation is also present throughout the unfolding process. It is fairly consistent around 4% of the total concentration fraction. Sedimentation velocity also calculates the average hydrodynamic radius (R_h) of the different molecular states. Average R_h values for tetramer and monomer molecules are shown in Figure 4.11 overlaying the fraction folded curve for WT. The fraction folded curve (black, circles) corresponds to the left scale and average R_h values for the tetramer (blue, triangles) and monomer (red, squares) correspond to the right scale (Figure 4.11).

Theoretical boundaries for folded and unfolded ADSL can be calculated based on random walk (R_g) [92]. Based on chain length, the R_h for folded protein is

$$R_h = (4.75 \pm 1.11)N^{0.29\pm0.02} \tag{4.4.1}$$

The R_h for a completed unfolded protein is

$$R_h = (2.21 \pm 1.07)N^{0.57\pm0.02} \tag{4.4.2}$$

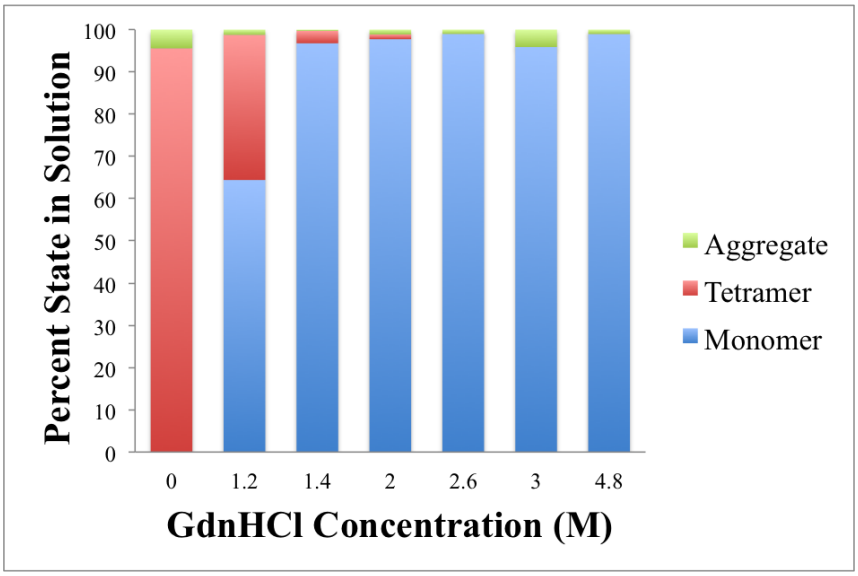


Figure 4.10: Fractions of protein states with different concentrations of denaturant.

This particular study included more α-helical proteins, so it is thought it will give a better estimate for ADSL monomer and tetramer R_h. The boundaries can be seen in Figure 4.12. The folded state R_h prediction is depicted by a solid line and a double line for the unfolded state. The R_h experimental data for the monomer fits well between the calculated boundaries and appears to start at the folded boundary and increase linearly to the unfolding boundary. Accordingly, it can be inferred the monomer is undergoing an unfolding process. The boundaries for the tetramer were not as clear. The calculated folded R_h was 4.2 nm too small. Since at 0 M denaturant, the enzyme is known to be in its native state, the average of the first two points was used as the boundary value. The completely unfolded boundary is very high, leading to the conclusion that most of the tetramer must be dissociating, but because of the noticeable increase in radius, there may be some unfolding occurring

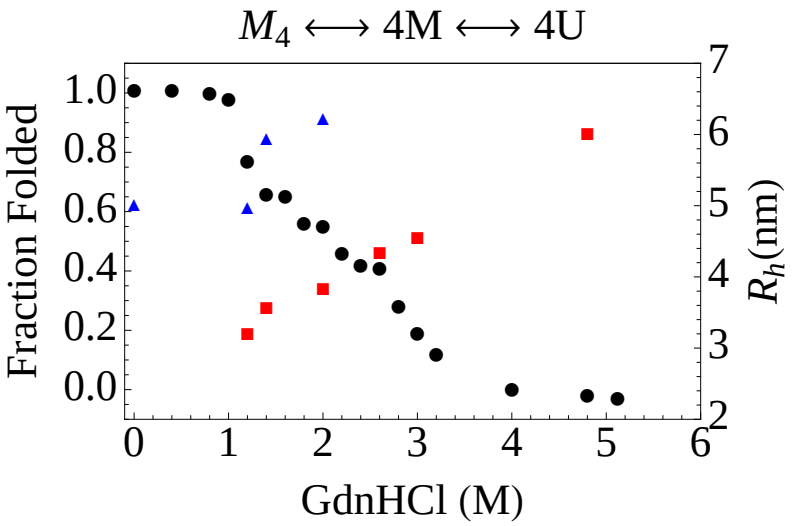


Figure 4.11: Overlay of fraction folded curve (black, circles) corresponding to the left scale and average R_h values for the tetramer (blue, triangles) and monomer (red, squares) corresponding to the right scale.

prior to breaking apart. It is unclear whether the tetramer is being completed unfolded, partially folded, or "swelling".

4.5 Thermodynamic Modeling of ADSL Stability

4.5.1 Validation of Code

Lactose repressor protein is a well characterized homotetrameric protein of molecular weight 150 kDa. Its monomer has three structurally separated domains and it is a dimer of dimers. Three papers have been published on the

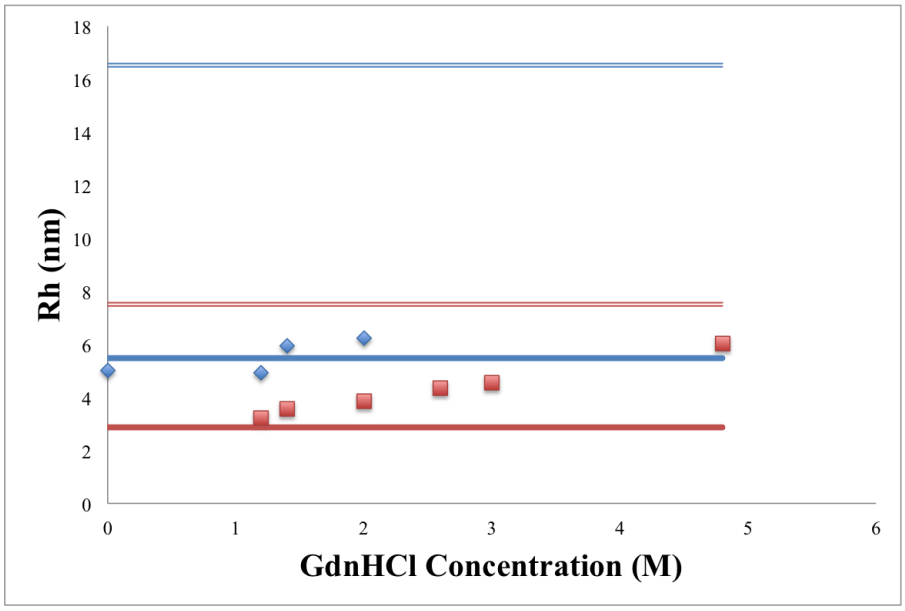


Figure 4.12: Theoretical boundaries for folded and unfolded tetramer (blue, diamonds) and monomer (red, squares). Solid lines are the folded boundary while the double lines are the unfolded boundary.

chemical stability of lactose repressor [6, 15, 69]. This makes lactose repressor a good system to test a prediction model for ADSL chemical denaturation. Lactose repressor is an interesting tetramer that under high concentration of urea, does not undergo dissociation [6]. Instead it transitions from a folded tetramer to an unfolded tetramer. Since there is no dissociation, the transition was treated as a two state problem [6,69]. To gain a complete characterization of the pathway, Chen et. al employed two alterations of lactose repressor to make the system form dimers instead of tetramers, a deletion mutant missing 11 amino acids (-11 aa) at the C-terminus that has diminished apparent operator binding and a mutant designated as R3 where the substitution of the C-terminal leucine heptad repeat region of the normally tetrameric lactose repressor by the leucine heptad repeat dimerization domain of GCN4 [15]. Since

Table 4.2: Comparison of Free Energy for Dimer Unfolding

Mutation	ΔG_{D-U} - Chen (kcal/mol)	ΔG_{D-U} - Ray (kcal/mol)
-11 aa	19.3 ± 1.4	19.2 ± 0.9
R3	23.9 ± 1.0	24.4 ± 0.5

Table 4.3: m-value for lactose repressor

Mutation	m-value - Ray (kcal/mol/M)
-11 aa	-4.2 ± 0.2
R3	-6.1 ± 0.3

dimer unfolding has an exact solution, a numerical, iterative solution could be tested with this data set.

The reported values for dimer unfolding were 19.3 ± 1.4 kcal/mol and 23.9 ± 1.0 kcal/mol, for -11 aa and R3 respectively (Table 4.2). Using PlotDigitizer to estimate the values of the fraction unfolded curves from ref [15], thermodynamic parameters were found by fitting with the exact solution (equation 1.9.31) in Mathematica 8.0. The solved parameters were very close to the published values (Table 4.2). The m-values were also determined. It is assumed they are correct because of the matching free energies because they were not included in the publication.

These values were then used for the numerical solution for a dimer unfolding. The graph was visually identical to the exact solution for both mutants as seen in Figure 4.13. A better way to test the numerical solution would have

been to run an iterative process to find the best fit values. Unfortunately, time did not permit such a test. However, this gives enough confidence that the method used for arriving at a numerical solution for dimer unfolding will also work for tetramer unfolding.

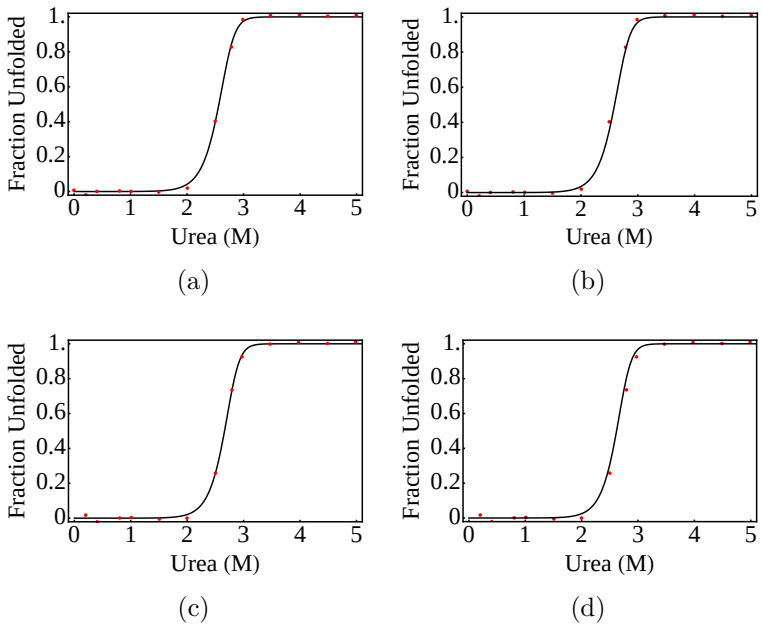


Figure 4.13: Graphs from exact solution of dimer unfolding (a and c) and numerical solution (b and d) for the -11 aa (a and b) and R3 (c and d) mutations. The two solutions provide almost identical graphs given the same thermodynamic parameters.

4.5.2 Two State Model Fails

All data suggests that there is an intermediate state present in ADSL denaturation. This is also confirmed by the following two state models for a folded tetramer to an unfolded monomer.

Boltzmann Distribution

The Boltzmann Distribution is a quick check for two state models. The only limitation is it does not account for concentration dependence due to tetramer dissociation.

$$P_f = \frac{1}{1 + exp[\beta \Delta G]} \tag{4.5.1}$$

The ΔG for the transition can be written as a linear dependence on denaturant concentration; $\Delta G = \Delta G_o$ + m[d]

$$P_f = \frac{1}{1 + exp[\beta(\Delta G_o - m[d]]} \tag{4.5.2}$$

The two-state model fails to capture the features of the fraction folded graph (Figure 4.14).

Numerical Solution

Unlike the dimer unfolding, tetramer unfolding does not have an exact solution that is dependent on concentration. Therefore the following numerical model was built.

$$N_{total} = N_{f,t} + N_{u,m} \tag{4.5.3}$$

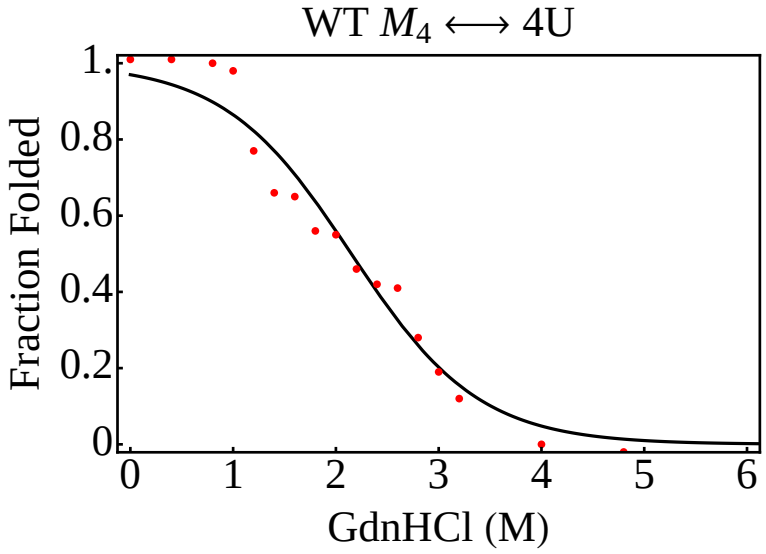


Figure 4.14: Two state prediction for the Boltzmann Distribution.

Where N_{total} is the initial monomer concentration, f and u denote folded and unfolded state, and t and m represent the tetramer and monomer state. Since the solution is in a state of equilibrium, we can relate $N_{f,t}$ and $N_{u,m}$ by

$$N_{f,t} = N_{u,m}^4 exp[-\beta \Delta G] = N_{u,m}^4 exp[-\beta(\Delta G_{f,t} - 4\Delta G_{u,m})] \tag{4.5.4}$$

The ΔG for the transition can be written as a linear dependence on denaturant concentration; $\Delta G = \Delta G_o + m[d]$

$$4\Delta G_{u,m} - \Delta G_{f,t} = \Delta G_o - m[d] \tag{4.5.5}$$

where ΔG_o and $m > 0$

$$N_{f,t} = N_{u,m}^4 exp[\beta(\Delta G_o - m[d])] \tag{4.5.6}$$

This allows equation 4.5.3 to be written in terms of $N_{u,m}$.

$$N_{total} = N_{u,m}^4 exp[\beta(\Delta G_o - m[d])] + N_{u,m} \tag{4.5.7}$$

From here, an iterative process can be coded to find the ΔG_o and m-values that result in the best prediction of the fraction folded curve.

$$P_f = \frac{N_{u,m}^4 exp[\beta(\Delta G_o - m[d])]}{N_{total}} \tag{4.5.8}$$

The two state model again fails to capture the features of the fraction folded graph (Figure 4.15). This confirms at least one intermediate state is present during denaturation.

4.5.3 Three State Model Improves Prediction

A similar method for a numerical solution can be applied for a three state model. The three state model was based on data supporting a folded monomer intermediate state ($M_{f,t} \rightarrow M_{f,m} \rightarrow M_{u,m}$). The numerical solution can attained as follows:

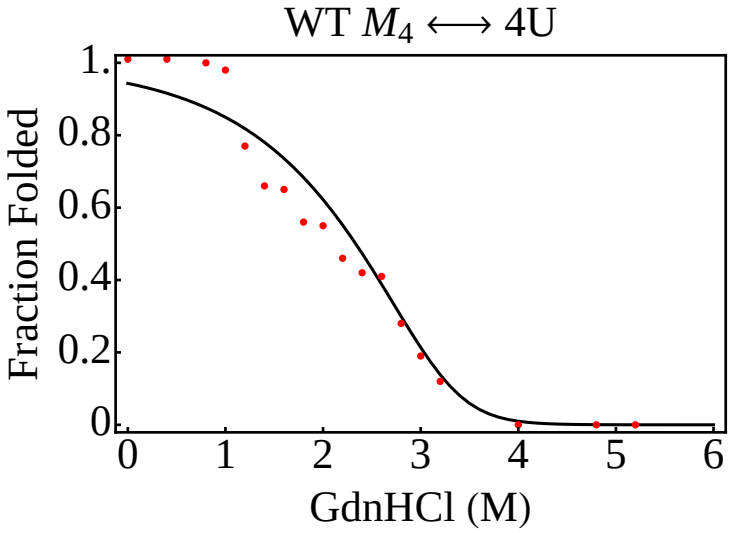


Figure 4.15: Two state predictions for the numerical solution.

$$N_{total} = N_{f,t} + N_{f,m} + N_{u,m} \tag{4.5.9}$$

$$N_{f,t} = N_{f,m}^4 exp[-\beta(\Delta G_{f,t} - 4\Delta G_{f,m})] \tag{4.5.10a}$$

$$N_{u,m} = N_{f,m} exp[-\beta(\Delta G_{u,m} - \Delta G_{f,m})] \tag{4.5.10b}$$

Substituting the linear expression of ΔG in terms of ΔG_o and m.

$$N_{f,t} = N_{f,m}^4 exp[\beta(\Delta G_{o,1} - m_1[d])] \tag{4.5.11a}$$

$$N_{u,m} = N_{f,m} exp[-\beta(\Delta G_{o,2} - m_2[d])] \tag{4.5.11b}$$

Writing equation 4.5.9 in terms of $N_{f,m}$

$$N_{total} = N_{f,m}^4 exp[\beta(\Delta G_{o,1} - m_1[d])] + N_{f,m} + N_{f,m} exp[-\beta(\Delta G_{o,2} - m_2[d])] \tag{4.5.12}$$

The fraction folded can then be predicted by

$$P_f = \frac{N_{f,m}^4 exp[\beta(\Delta G_{o,1} - m_1[d])] + N_{f,m}}{N_{total}} \tag{4.5.13}$$

where $\Delta G_{o,1}$ and $\Delta G_{o,2}$ are the free energies of the tetramer dissociation and monomer unfolding respectively. The iteration was run with unbound and bound parameters. Executing the code with unbound parameters should lead to a global minimum. The resulting values are summarized in Table 4.4. There is a large contrast in the free energies. The ΔG for dissociation is 119.3 kT/mol while the monomer unfolding is 3.1 kT/mol. Since four monomers are unfolded, the total free energy is 12.4 kT/mol. This value is still a small contribution to the total free energy of unfolding ($\sim$10%). Based on chain length, the expected m-values for unfolding of a tetramer and monomer would be 68-81 and 18-21 respectively [27, 60]. The m-value for dissociation is close to the expected m-value for a tetramer that undergoes unfolding. The m-value for the monomer is much lower than expected. Compared to lactose repressor, which is similar in structure and chain length, there is a significant difference in the total free energy of unfolding and m-values. The graph resulting from

Table 4.4: WT 3 State Parameters for Bound and Unbound Parameters

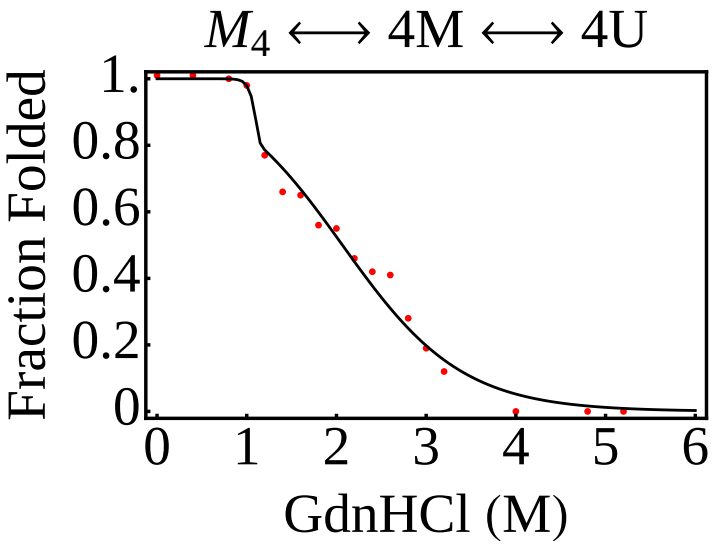
Parameter	Bound	Unbound	Theoretical
ΔG_{M_4-4M} (kT/mol)	119.3	65.0	-
ΔG_{M-U} (kT/mol)	3.1	3.1	-
m_{M_4-4M} (kT/mol/M)	74.0	23.0	68-81*
m_{M-U} (kT/mol/M)	1.5	1.5	18-21

*Theoretical range for tetramer unfolding

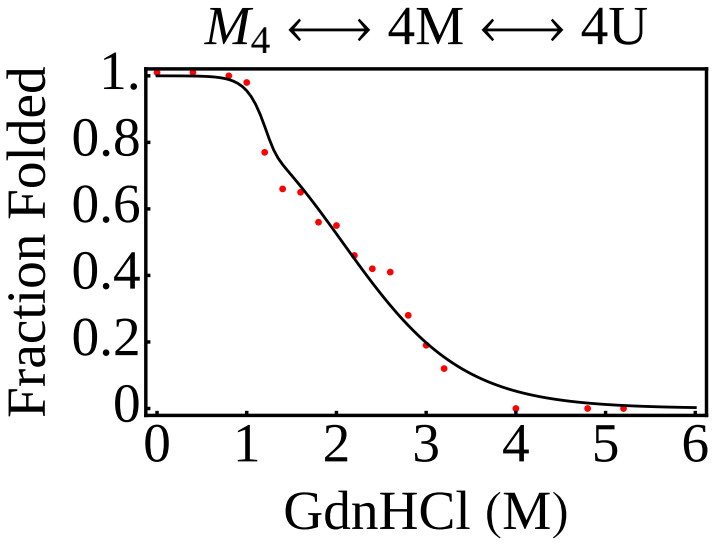
the fit parameters is in good agreement with the data (Figure 4.16a).

The bound parameters only set a limit on the total free energy of the unfolding process based on DSC results. The thermodynamic parameters for the monomer unfolding were identical to the unbound results. The values for the tetramer unfolding were lower, but the resulting graph is clearly not as good of a fit (Figure 4.16b). While the unbound fit may be the better of the two graphs, the bound fit is closer to physically relevant values for ΔG_o. It is also difficult to compare the theoretical m-value for the first transition because the modeled transition was dissociation and not unfolding.

Parameters from the bound numerical solution were used to predict the average R_h. The upper and lower limit for folded and unfolded monomer were set to 2.9 nm based on reference [92] and 6.1 nm respectively. The unfolded boundary came from the experimental data at 4.8 M denaturant concentration. Since the data point on the fraction folded curve is on the unfolded baseline, we believe this to be a good estimate for the final transition. The predictions (lines) for the average R_h displayed an overestimation of the monomer unfold-



(a)



(b)

Figure 4.16: Three state predictions for unbound (a) and bound (b) parameters.

ing and failed to capture the increase in tetramer radius (Figure 4.17). The fraction folded curve is calculated based off the loss of secondary structure. In the current three state model, the only cause for a loss of structure is due to

the monomer unfolding. Thus, the overestimation is not a surprise. What this indicates is that a fourth state is necessary to accurately describe the system.

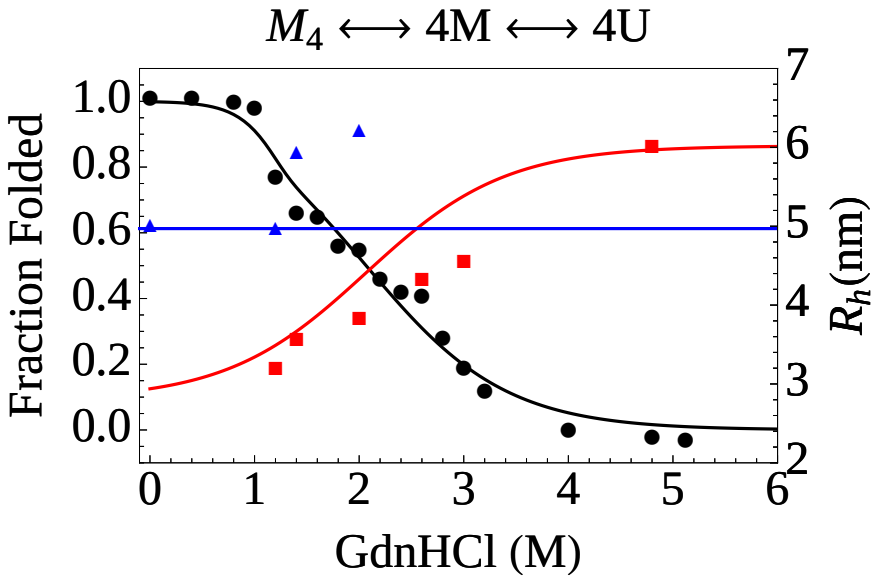


Figure 4.17: Overlay of fraction folded curve (black, circles) and average R_h values for the tetramer (blue, triangles) and monomer (red, squares). The fit for the 3 state model is the black line. The prediction of the R_h for the tetramer and monomer species are blue and red respectively. Evident from the monomer prediction, the unfolding of the monomer is over-predicted.

4.5.4 Four State Model is the Next Step

Though the three state model has a reasonable fit to the CD data, it is unable to predict the AUC measurements. The major flaw was it over predicted the fraction of molecules in the monomer state and under predicted the tetramer state. It also fails to capture the increase in radius of the tetramer. This motivates the need for a fourth state in the model. Based on observations

in our data and on the homotetramer lactose repressor protein, the fourth state will initially be consider an unfolded tetramer [6].

Chapter 5

Future Work

5.1 Research Directions for Enzyme Activity

5.1.1 Further Development of HPLC-EC Enzyme Assay

A current limitation to the HPLC-EC Enzyme assay is we have only developed an experimental method for measuring the V_{max} kinetic parameter. It would be beneficial to develop an experimental method that captures the whole kinetic curve that can be modeled with either the Michaelis-Menten or Hill Equation. The challenge is measuring the initial velocities at low concentration of substrate. According to the UV assay, the linearity of the initial velocity lasts only a few seconds at low concentrations before the substrate is no longer saturated or the substrate is completely undetectable by UV. Using the current method to stop the reaction would not give enough resolution to determine the initial velocity. An automated stop-flow system would have to be engineered to achieve good resolution. Although, It is possible with the

sensitivity of the HPLC-EC that the linearity actually lasts longer than what the UV enzyme assay is able to detect.

5.1.2 Finish Characterization of D87E and E80D Mutations

The E80D and D87E mutations are an enigma in regards to why they cause ADSL deficiency. They both have relatively high activities and appear to be qualitatively as stable as WT (Figures 5.1, 5.2 and 5.3; Table 5.1 and 5.2). The stability of each protein was checked by thermal melting curves using circular dichroism (CD). WT ADSL melts at 65.5 , E80D at 63.9, and D87E at 67.7 ^{o}C(Figure 5.2). Though these differences are small, the higher melting temperature of D87E is consistent with our finding that this mutation stabilizes the equivalent *B. subtilis* mutant ADSL. We also checked this by performing chemical denaturation for WT and mutant ADSL and found similar transition patterns (Figure 5.3). We assessed the molecular weight of each mutant and WT ADSL by light scattering. In each case, the predominant species has a molecular weight around 228 kDa, as would be expected for the tetrameric form of ADSL. About 3.6% of WT ADSL was present as aggregates larger than tetramer, with slightly larger amounts for E80D (4.6%). Importantly, 13.1% of the D87E protein was present in large aggregates (Table 5.3). This could point to the mutations E80D and/or D87E causing a "gain in toxic function" as the pathogenic mechanism for ADSL deficiency. Since they are both present in the same patient, it may be pertinent to perform experiment on

Table 5.1: SAMP UV Kinetic Parameters of Human ADSL at 25°C

Enzyme	V_{max} (μmol min^{-1} mg^{-1})	$K_{0.5}$ (μM)	Hill Coeff.	$\frac{k_{cat}}{K_{0.5}}$
WT	13.9 ± 0.2	2.05 ± 0.08	1.4 ± 0.1	2.52 x 10^7
E80D	9.2 ± 0.2	2.49 ± 0.06	1.6 ± 0.1	2.35 x 10^7
D87E	8.0 ± 0.2	1.09 ± 0.07	1.3 ± 0.1	2.80 x 10^7

The SAMP kinetic parameters were determined by varying the concentration of SAMP and fitting thedata to the Hill equation using Mathematica. The values are shown along with their standard error.

hybrid tetramers as well [4].

In vivo studies by other members of the Patterson Lab show WT and E80D transfectants grew well in the absence of adenine. Surprisingly, the D87E transfectants were unable to grow in the absence of adenine in the growth medium. RT-PCR analysis of the transfectants revealed that all made abundant hADSL mRNA (data not shown). However, while WT and E80D transfectants made significant amounts of hADSL protein, this protein was not detectable in D87E transfectants. Consistent with these observations Spiegel has shown that P1 genomic clones containing either WT or E80D genomic hADSL can allow growth of AdeI cells, whereas P1 clones containing D87E hADSL cannot. More work needs to be done to fully characterize these two mutations in order to fully understand how they change the properties of WT ADSL.

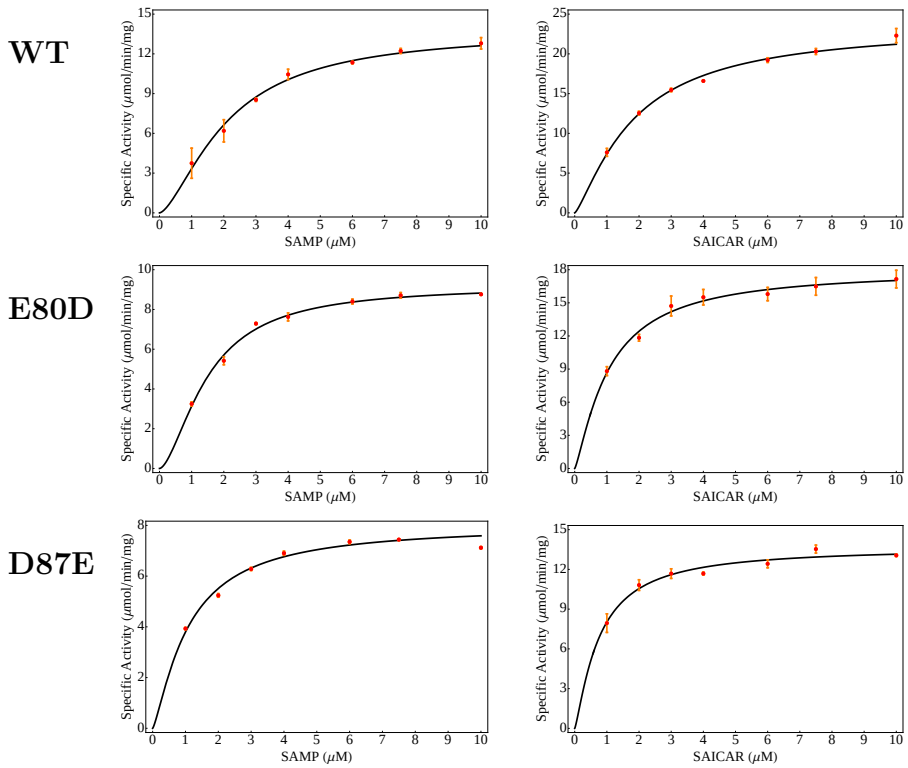


Figure 5.1: Specific Activity vs [Substrate] plots.

5.1.3 Further Investigate the Flexible Loop Region

The current understanding of the functions of the amino acid residues from 284 to 292 in ADSL is incomplete. They are hypothesized to function as a flexible loop that enclosed the active site upon substrate binding. The A291V mutation results in an inactive ADSL enzyme. To test the flexibility hypothesis, a collaboration with Banu Ozkan via Kingshuk Ghosh was initiated to use a harmonic oscillator model predict the loss of flexibility due to the A291V mutation (Figure 5.4).

Table 5.2: SAICAR UV Kinetic Parameters of Human ADSL at 25°C

Enzyme	V_{max} (μmol min^{-1} mg^{-1})	$K_{0.5}$ (μM)	Hill Coeff.	$\frac{k_{cat}}{K_{0.5}}$
WT	23.3 ± 0.5	1.74 ± 0.08	1.3 ± 0.1	5.09 x 10^7
E80D	18.0 ± 0.4	1.06 ± 0.07	1.3 ± 0.1	6.45 x 10^7
D87E	13.7 ± 0.4	0.75 ± 0.08	1.3 ± 0.2	6.88 x 10^7

The SAICAR kinetic parameters were determined by varying the concentration of SAICAR and fitting the data to the Hill equation using Mathematica. The values are shown along with their standard error.

Table 5.3: ADSL Population Distribution

Enzyme	Tetramer Weight (kDa)	Tetramer (%)	Aggregate (%)
WT	220.8 ± 0.2%	96.2 ± 0.6	3.6 ± 0.6
E80D	221.8 ± 0.3%	95.4 ± 0.8	4.6 ± 0.8
D87E	231.5 ± 0.2%	86.7 ± 0.2	13.1 ± 0.1

The results indicate a predicted decrease in the regions flexibility supporting the current hypothesis. More work needs to be done on the complete loop region to provide a better understanding of these amino acids.

5.1.4 Incomplete Models Relating to Enzyme Activity

Two Site, Two Substrate Model

The mixing ratios and cooperativity led to an attempt to derive a model based on of the mass action principle to predict relative rates of formation of the products at any mixing ratio of substrate. It is based on the idea that the "top" two active sites and "bottom" two active sites are too far away from each other to induce a conformational change. Under this assumption, the

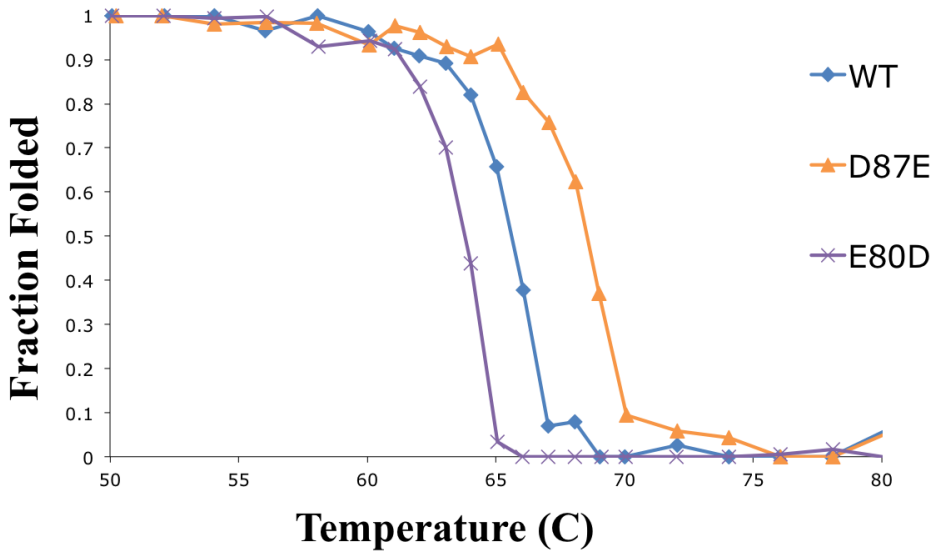


Figure 5.2: E80D and D87E Thermal Melts

only cooperativity that could occur would be between two active sites and only models the activity of the top (or bottom) two active sites as depicted in Figure 5.5. This model attempts to predict the relative rates of formation of the products at any mixing ratio of substrate. Preliminary results can be seen in Figure 5.6 for WT (a), R303C (b), and R426H (b). Nothing new came from this early prediction, but it could be that it may need to be developed more.

Purine Biosynthesis Pathway

A limitation to the *in vitro* experiments is they do not take into account the upstream and downstream activities of the other enzymes involved in *de novo* purine biosynthesis. It is a very basic model that does not take into account anything but the relative catalytic constants of the enzymes that were all assumed to be equal to 1 except ADSL. It does not account for purinosome

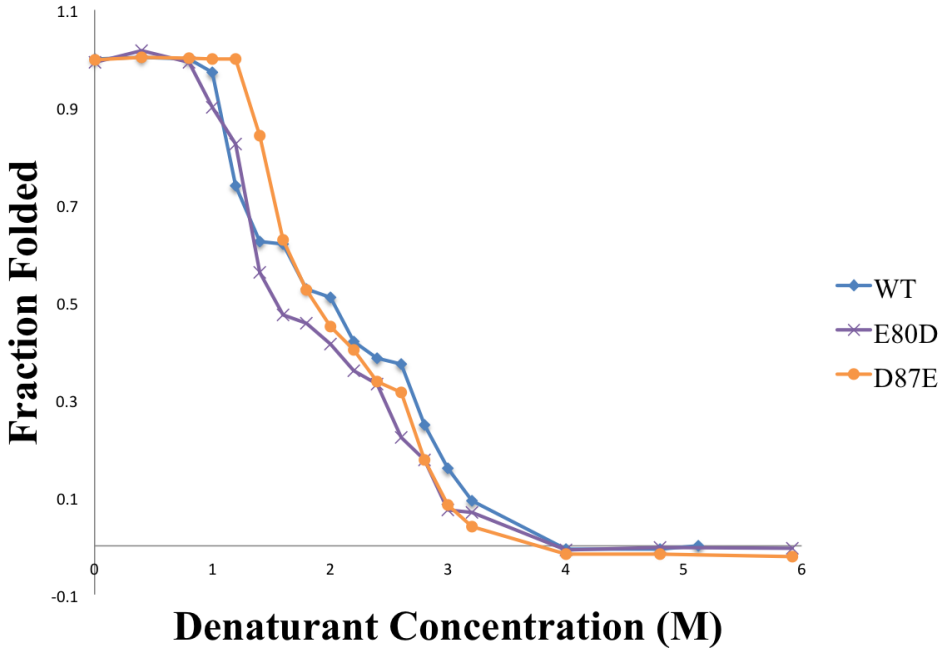


Figure 5.3: Chemical denaturation of E80D and D87E

formation, potential substrate trafficking, feedback loops, etc. Preliminary results can be seen in Figure 5.7 using k_{cat} values for WT (a), R303C (b), and R426H (c). The most dramatic difference is with R303C where the AICAR (red line) peak is significantly lower and is broader. This causes a rise in both substrates, but a more prominent rise in SAMP (green line). More work is required to come to any inferences or conclusions.

5.2 Investigate Crowding Effect

A shortcoming of current biochemical studies is they are performed under ideal conditions with pure enzymes stored in dilute buffers that may not accurately reflect the actual intra-cellular environment. It has been shown

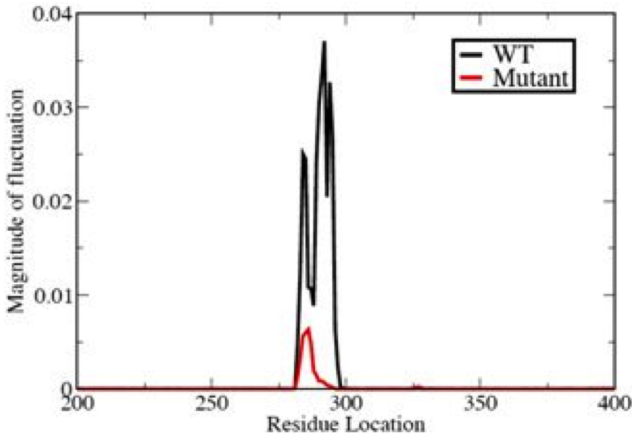


Figure 5.4: Harmonic oscillator model for the flexible loop of ADSL that is thought to cover the active site to speed up catalysis. The mutation A291V severely affects the flexibility of the loop.

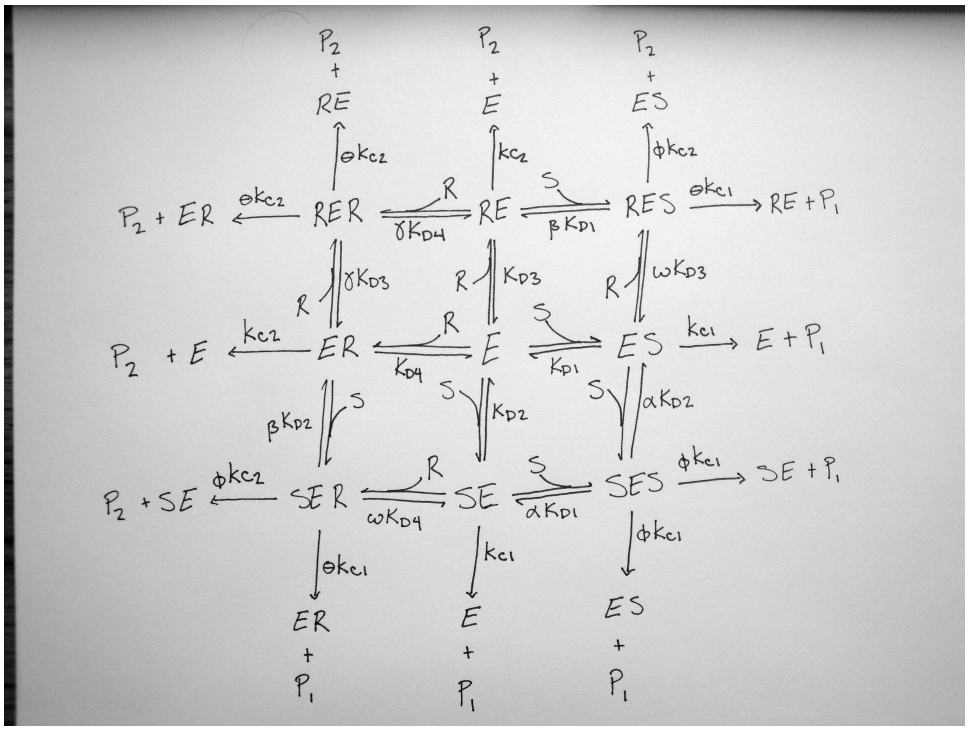


Figure 5.5: Pictorial representation of the enzyme reactions for a two site, two substrate model that accounts for cooperativity

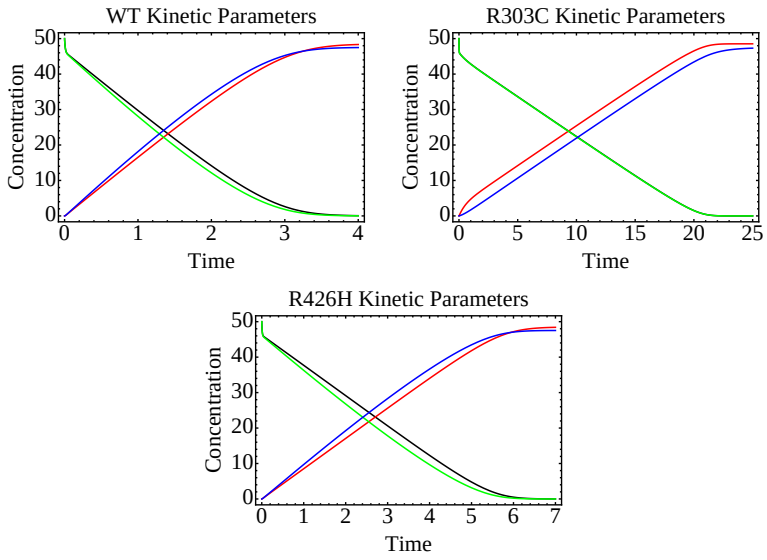


Figure 5.6: Prediction of ADSL kinetics by a model of a system with 2 active sites and 2 substrates present. Graphs represent a 50/50 mixture of substrate concentration. Black - SAICAR, Red - AICAR, Green - SAMP, Blue - AMP.

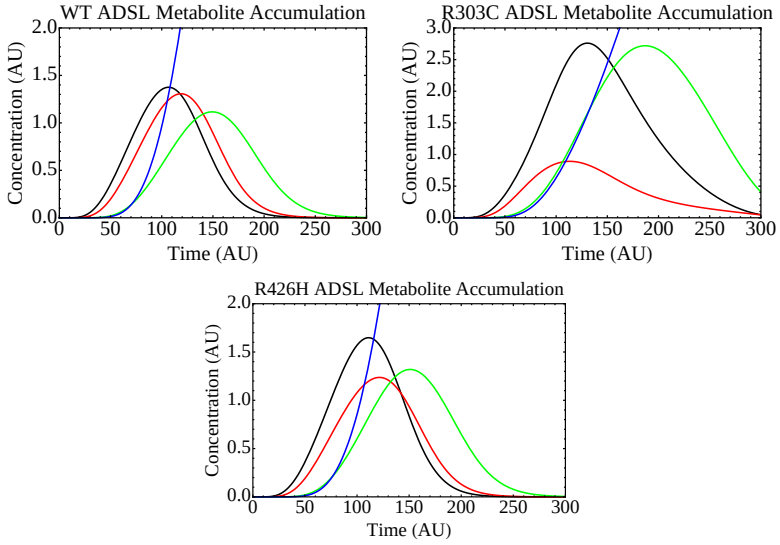


Figure 5.7: Prediction of accumulation of ADSL metabolites in the *de novo* purine pathway. Black - SAICAR, Red - AICAR, Green - SAMP, Blue - AMP.

that the dynamics and stability of biomolecules *in vitro* is modified by crowding [16, 26, 35, 39]. The majority of cellular functions take place in a crowded

environment due to the presence of macromolecules that occupy up to 5% to 40% of the total cell volume [23]. Due to synergistic effects of volume exclusion by macromolecules, even less space becomes available to other functional macromolecules such as enzymes [23, 96]. Thus crowding can significantly alter biochemical properties of these enzymes. For example, crowding usually promotes processes that stimulate increase in the accessible volume such as aggregation [23]. A reduced accessible volume available for the proteins may effectively enhance the probability of association of different monomers into higher order protein aggregates. However, this is not the only effect of crowding; there are other competing effects as well. For example, protein association often occurs by simple diffusion and the presence of crowding may significantly reduce the diffusion constant thus kinetically hindering the aggregation process. This may adversely affect the rate of formation of tetrameric ADSL. Presence of confinement and crowding can also induce steric hindrance reducing stability of disordered or flexible conformations that may affect the enzyme activity as well. It is thus imperative to carry out structure function analysis of ADSL enzyme in the presence of crowding to better mimic in vivo conditions. The peculiar sensitivity of ADSL and several atypical activities measured under in vitro conditions that does not correlate well with expected function of the enzyme could be attributed to the absence of crowding in such experiments which perhaps play a significant role in vivo and hence in disease pathogenesis.

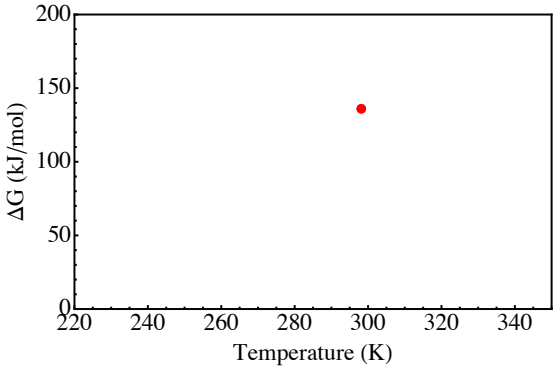
5.3 Research Directions for Protein Stability

5.3.1 Immediate Projects for Chemical Stabilty

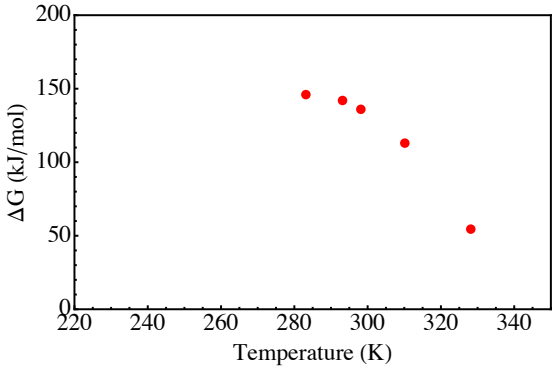
Once analysis is complete for the first ΔG at 25 °C (Figure 5.8a), it can be applied systematically to chemical denaturant experiments at different temperatures and concentrations. Measuring the ΔG at different temperatures will add more data points to the free energy of folding vs. Temperature curve (Figure 5.8b). If a sufficient number of data points are collected, thermodynamics can be used to model the free energy of folding curve to extract thermodynamic parameters ΔH, ΔS, and ΔC_p (Figure 5.8c). Measuring the unfolding transition at different concentrations will provide more information about whether or not a dissociation. A concentration dependent transition means there has been a dissociation that has taken place. This could be useful in complimenting sedimentation velocity in identifying whether the tetramer begins to dissociate before unfolding.

5.3.2 Investigate DSC Conditions for Reversibility

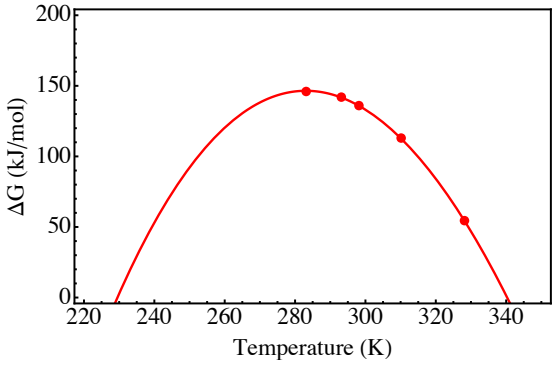
Thermal denaturation has been shown to promote protein aggregation following unfolding and hence reversibility is difficult to achieve with a monomeric protein. Hence it is not surprising ADSL, a tetrameric protein, was not reversible by thermal denaturation. Papers have shown that additions of small amounts of denaturant (such as Guanidine Hydrochloride) or macromolecules (such as TMAO), make solution conditions favorable for reversibility. The logic is the denaturant and macromolecules act as crowding agents that do not allow unfolded units to interact with each other. This would be useful to



(a) After analysis at 25 °C is complete there will be one point on the ΔG vs T plot



(b) If chemical denature experiments are run at different temperatures, more points can be plotted



(c) When a sufficient number ΔG's are measured at different temperatures, thermodynamics can model the free energy of folding curve

Figure 5.8: Future Directions for Chemical Denaturation Experiments.

try with ADSL protein. Thermodynamic reversibility on the DSC would best proven by showing ΔC_p is independent of scan rate as opposed to seeing the second scan providing the same, or similar, ΔH. The later is more of a test of thermodynamic repeatability as opposed to reversibility [73]

5.3.3 *In Vivo* Fluorescence Spectroscopy

Simulating the intracellular environment is not a simple task for *in vitro* experiments. Currently, the best we can do is to insert inert crowding agents into the solution. But this does not take into account the myriad of interactions that could occur with cellular transport, chaperones. To reveal the full functionality of proteins and other biomolecules, they must be studied in the living cell [22].

Fluorescently Tagged Proteins

We have successfully tagged all ADSL protein with RFP and GFP tags (Figure 5.9). They express the protein and glow under the microscope. It also appears co-transfections are possible. Very preliminary experiments have been done on the cells to create stable cell lines and to make sure the fluorescent tags do not change how the transfected ADSL behaves without the tag. Basic starvation experiments were performed for cell lines with RFP and GFP tagged ADSL (Figure 5.10). Interestingly, D87E never grew and GFP-R426H was able to rescue the cell lines while RFP-R426H was not. The case was vise-versa for R303C which was able to rescue the cell line when tagged with RFP and was unable to rescue the cell line with GFP.

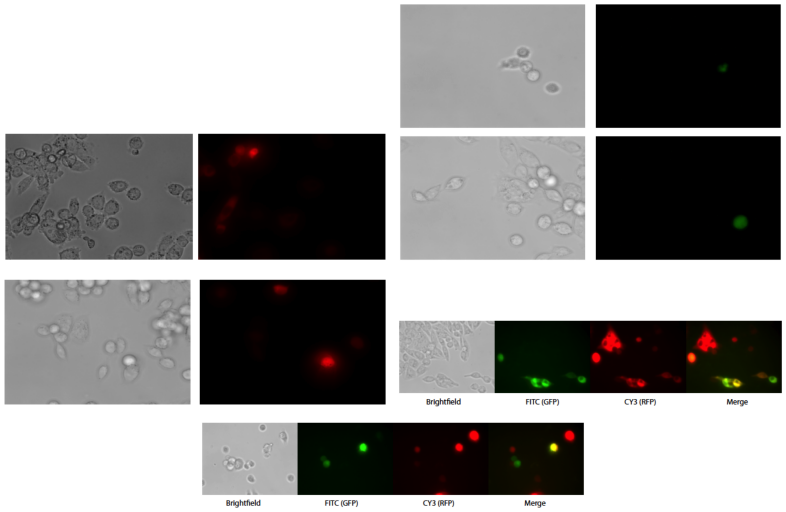


Figure 5.9: Transient transfections

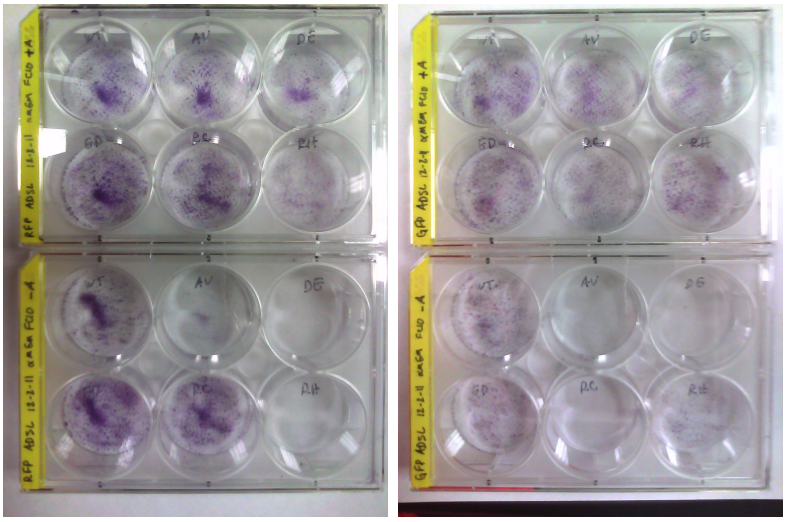


Figure 5.10: Starvations

These abnormalities caused by the fluorescent tags as well as the inability to create stable cell lines with both tags (Figure 5.11 bring up the concern the tags themselves are altering the functionality of ADSL. Overall, since the transient cell lines seem to express fluorescent protein as well as co-transfected ADSL, it provides promise for future experiments for ADSL trafficking, purinosome

formation, and *in vivo* fluorescence stability measurements. it is possible this method will not provide any useful information and a different method such as immunoflourescence should be used concurrently to compare results.

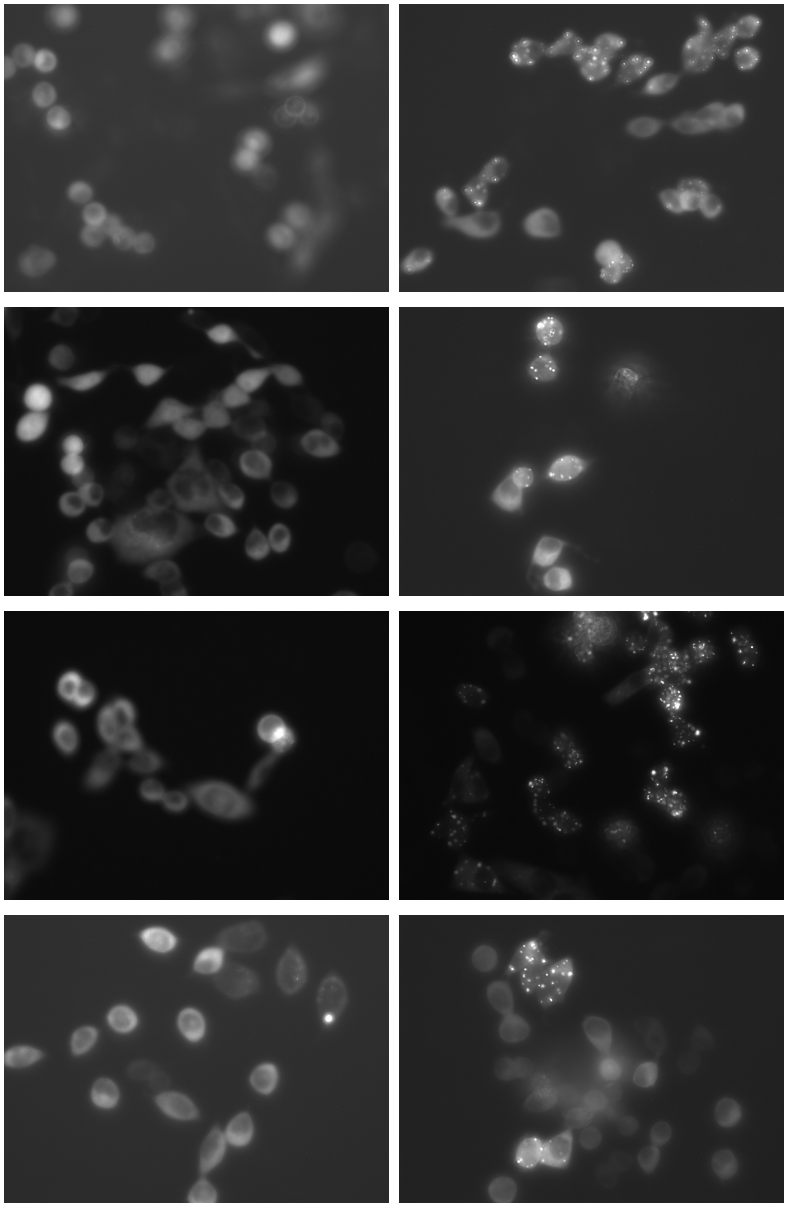


Figure 5.11: Stable cell lines

Fast Relaxation Imaging

Fast relaxation imaging (FReI; Figure 5.12) is a technique that can be used to image fast dynamics of biomolecules *in vivo* [22]. It couples time-resolved fluorescence imaging with fast temperature jump induced kinetics. It utilizes FRET signaling to determine whether a protein is folded or unfolded Figure 5.13. ADSL is tagged on the N-terminus for both RFP and GFP tags. This was quite fortunate because it the only two terminuses that come in close enough contact for FRET to occur is two N-terminus of the A and D chains (3.3 nm; Figure 5.14). Setting up this sort of experiment would be a big task for either multiple undergraduates or new graduate student. It could also start up collaboration with Dr. Mark Siemen's optics lab if a FReI system need to be "rigged up" as opposed to purchasing all new equipment.

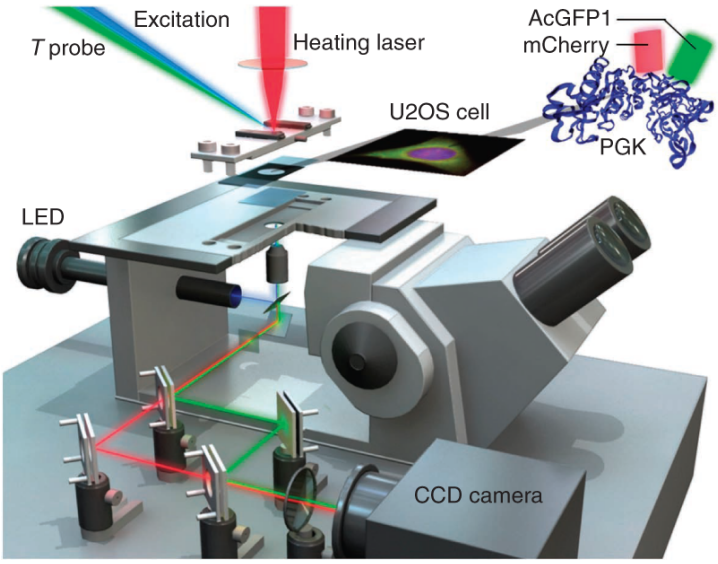


Figure 5.12: Pictorial representation of an FReI experimental setup.

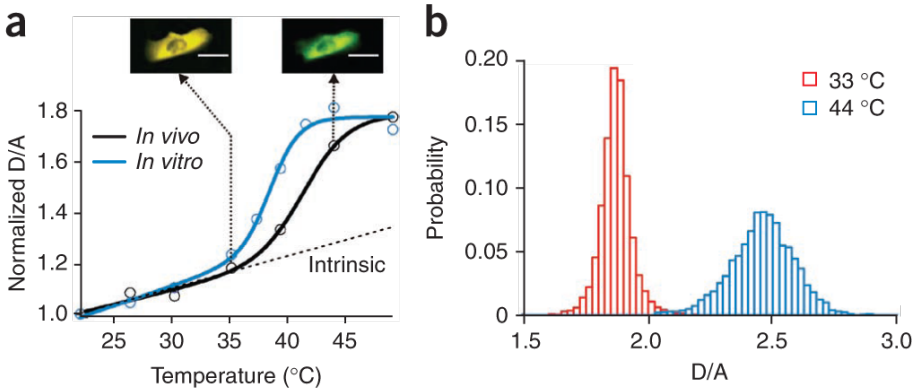


Figure 5.13: Data you would expect to see from an FReI experiment

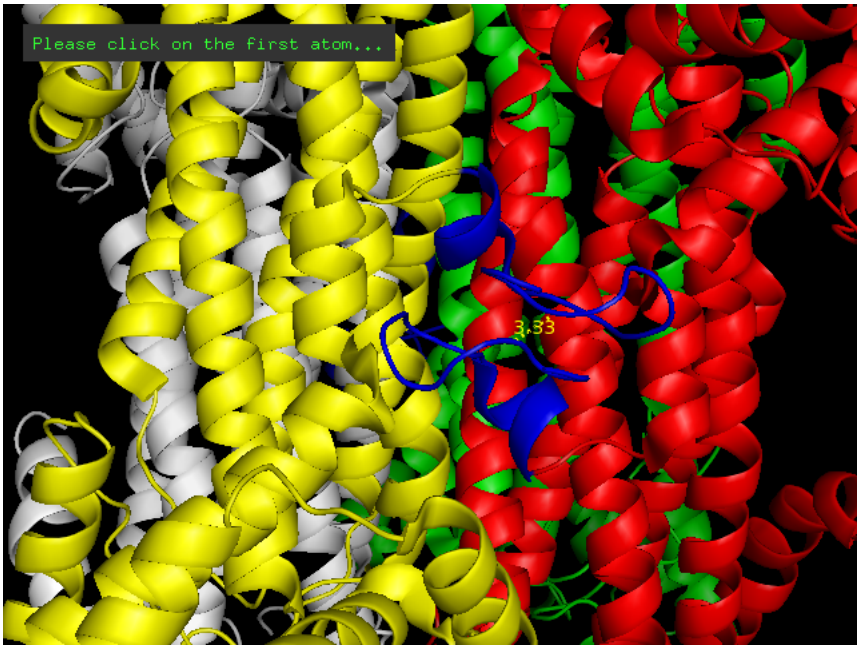


Figure 5.14: Approximate location for fluorescently labeled proteins. Tag location is depicted by blue chains. Area is approximate because the first 15 amino acids do not have a specified location

Bibliography

[1] M ALEXIOU and HJ LEESE. PURINE UTILIZATION, DE NOVO SYNTHESIS AND DEGRADATION IN MOUSE PREIMPLANTATION EMBRYOS. *DEVELOPMENT*, 114(1):185–192, JAN 1992.

[2] Songon An, Ravindra Kumar, Erin D. Sheets, and Stephen J. Benkovic. Reversible compartmentalization of de novo purine biosynthetic complexes in living cells. *SCIENCE*, 320(5872):103–106, APR 4 2008.

[3] Lushanti De Zoysa Ariyananda, Peychii , Christina Antonopoulos, and Roberta F. Colman. Biochemical and Biophysical Analysis of Five Disease-Associated Human Adenylosuccinate Lyase Mutants. *BIOCHEMISTRY*, 48(23):5291–5302, JUN 16 2009.

[4] Lushanti De Zoysa Ariyananda, Christina Antonopoulos, Jenna Currier, and Roberta F. Colman. In Vitro Hybridization and Separation of Hybrids of Human Adenylosuccinate Lyase from Wild-Type and Disease-Associated Mutant Enzymes. *BIOCHEMISTRY*, 50(8):1336–1346, MAR 1 2011.

[5] William E. Balch, Richard I. Morimoto, Andrew Dillin, and Jeffery W. Kelly. Adapting proteostasis for disease intervention. *SCIENCE*, 319(5865):916–919, FEB 15 2008.

[6] Jennifer K. Barry and Kathleen S. Matthews. Thermodynamic analysis of unfolding and dissociation in lactose repressor protein. *Biochemistry*, 38(20):6520–6528, 1999.

[7] B.A. Barshop, A.S. Albets, P.K. Laikin, and H.E. Gruber. Studies of mutant human adenylosuccinate lyase. *Adv. Exp. Med. Biol.*, 253A:23–30, 1989.

[8] Manuela Bartolini and Vincenza Andrisano. Strategies for the Inhibition of Protein Aggregation in Human Diseases. *CHEMBIOCHEM*, 11(8):1018–1035, MAY 17 2010.

[9] Katharina Bluemlein, Nana-Maria Gruening, Rene G. Feichtinger, Hans Lehrach, Barbara Kofler, and Markus Ralser. No evidence for a shift in pyruvate kinase PKM1 to PKM2 expression during tumorigenesis. *ONCOTARGET*, 2(5):393–400, MAY 2011.

[10] WA BRIDGER and LH COHEN. KINETICS OF ADENYLOSUCCINATE LYASE. *JOURNAL OF BIOLOGICAL CHEMISTRY*, 243(3):644–&, 1968.

[11] JL Brosius and RF Colman. Three subunits contribute amino acids to the active site of tetrameric adenylosuccinate lyase: Lys(268) and Glu(275) are required. *BIOCHEMISTRY*, 41(7):2217–2226, FEB 19 2002.

[12] JD BRYNGELSON, JN ONUCHIC, ND SOCCI, and PG WOLYNES. FUNNELS, PATHWAYS, AND THE ENERGY LANDSCAPE OF PROTEIN-FOLDING - A SYNTHESIS. *PROTEINS-STRUCTURE FUNCTION AND GENETICS*, 21(3):167–195, MAR 1995.

[13] JD BRYNGELSON and PG WOLYNES. SPIN-GLASSES AND THE STATISTICAL-MECHANICS OF PROTEIN FOLDING. *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA*, 84(21):7524–7528, NOV 1987.

[14] Kelsey Caetano-Anolles and Gustavo Caetano-Anolles. Structural Phylogenomics Reveals Gradual Evolutionary Replacement of Abiotic Chemistries by Protein Enzymes in Purine Metabolism. *PLOS ONE*, 8(3), MAR 13 2013.

[15] Jie Chen and Kathleen S. Matthews. Subunit dissociation affects dna binding in a dimeric lac repressor produced by c-terminal deletion. *Biochemistry*, 33(29):8728–8735, 1994.

[16] MS Cheung, D Klimov, and D Thirumalai. Molecular crowding enhances native state stability and refolding rates of globular proteins. *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA*, 102(13):4753–4758, MAR 29 2005.

[17] F Ciardo, C Salerno, and P Curatolo. Neurologic aspects of adenylosuccinate lyase deficiency. *JOURNAL OF CHILD NEUROLOGY*, 16(5):301–308, MAY 2001.

[18] FHC CRICK. THE PACKING OF ALPHA-HELICES - SIMPLE COILED-COILS. *ACTA CRYSTALLOGRAPHICA*, 6(8-9):689–697, 1953.

[19] KA DILL. THEORY FOR THE FOLDING AND STABILITY OF GLOBULAR-PROTEINS. *BIOCHEMISTRY*, 24(6):1501–1509, 1985.

[20] CM Dobson, A Sali, and M Karplus. Protein folding: A perspective from theory and experiment. *ANGEWANDTE CHEMIE-INTERNATIONAL EDITION*, 37(7):868–893, APR 20 1998.

[21] Nathan Duval, Kyleen Luhrs, Terry G. Wilkinson, II, Veronika Baresova, Vaclava Skopova, Stanislav Kmoch, Guido N. Vacano, Marie Zikanova, and David Patterson. Genetic and metabolomic analysis of AdeD and AdeI mutants of de novo purine biosynthesis: Cellular models of de novo purine biosynthesis deficiency disorders. *MOLECULAR GENETICS AND METABOLISM*, 108(3):178–189, MAR 2013.

[22] Simon Ebbinghaus, Apratim Dhar, Douglas McDonald, and Martin Gruebele. Protein folding stability and dynamics imaged in a living cell. *NATURE METHODS*, 7(4):319–323, APR 2010.

[23] RJ Ellis and AP Minton. Cell biology - Join the crowd. *NATURE*, 425(6953):27–28, SEP 4 2003.

[24] AR FERSHT, A MATOUSCHEK, and L SERRANO. THE FOLDING OF AN ENZYME .1. THEORY OF PROTEIN ENGINEERING ANALYSIS OF STABILITY AND PATHWAY OF PROTEIN FOLD-

ING. *JOURNAL OF MOLECULAR BIOLOGY*, 224(3):771–782, APR 5 1992.

[25] F Franks, CN Pace, K Wilson, RM Daniel, PJ Halling, DS Clark, and A Purkiss. Protein structure, stability and solubility in water and other solvents - Discussion. *PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY OF LONDON SERIES B-BIOLOGICAL SCIENCES*, 359(1448):1234–1235, AUG 29 2004.

[26] H. Frauenfelder, P. W. Fenimore, G. Chen, and B. H. McMahon. Protein folding is slaved to solvent motions. *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA*, 103(42):15469–15472, OCT 17 2006.

[27] Kingshuk Ghosh and Ken A. Dill. Computing protein stabilities from their chain lengths. *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA*, 106(26):10649–10654, JUN 30 2009.

[28] Kingshuk Ghosh, S. Banu Ozkan, and Ken A. Dill. The ultimate speed limit to protein folding is conformational searching. *JOURNAL OF THE AMERICAN CHEMICAL SOCIETY*, 129(39):11920–11927, OCT 3 2007.

[29] SC GILL and PH VONHIPPEL. CALCULATION OF PROTEIN EXTINCTION COEFFICIENTS FROM AMINO-ACID SEQUENCE DATA. *ANALYTICAL BIOCHEMISTRY*, 182(2):319–326, NOV 1 1989.

[30] Cyril Gitiaux, Irene Ceballos-Picot, Sandrine Marie, Vassili Valayannopoulos, Marlene Rio, Severine Verrieres, Jean Francois

Benoist, Marie Francoise Vincent, Isabelle Desguerre, and Nadia Bahi-Buisson. Misleading behavioural phenotype with adenylosuccinate lyase deficiency. *EUROPEAN JOURNAL OF HUMAN GENETICS*, 17(1):133–136, JAN 2009.

[31] R GRANTHAM. AMINO-ACID DIFFERENCE FORMULA TO HELP EXPLAIN PROTEIN EVOLUTION. *SCIENCE*, 185(4154):862–864, 1974.

[32] W Guo, JE Shea, and RS Berry. The physics of the interactions governing folding and association of proteins. In Lee, RC and Despa, F and Hamann, KJ, editor, *CELL INJURY: MECHANISMS, RESPONSES, AND REPAIR*, volume 1066 of *ANNALS OF THE NEW YORK ACADEMY OF SCIENCES*, pages 34–53, 2 EAST 63RD ST, NEW YORK, NY 10021 USA, 2005. NEW YORK ACAD SCIENCES. Conference on Cell Injury - Responses and Repair, Univ Chicago, Chicago, IL, MAR 31-JUN 02, 2004.

[33] Davin M. Henderson, Ann Lee, and James M. Ervasti. Disease-causing missense mutations in actin binding domain 1 of dystrophin induce thermodynamic instability and protein aggregation. *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA*, 107(21):9632–9637, MAY 25 2010.

[34] Hans C. Huerlimann, Benoit Laloo, Barbara Simon-Kayser, Christelle Saint-Marc, Fanny Coulpier, Sophie Lemoine, Bertrand Daignan-Fornier, and Benoit Pinson. Physiological and Toxic Effects of Purine Intermediate 5-Amino-4-imidazolecarboxamide Ribonucleotide (AICAR) in Yeast.

JOURNAL OF BIOLOGICAL CHEMISTRY, 286(35):30994–31002, SEP 2 2011.

[35] Zoya Ignatova, Beena Krishnan, Jeffrey P. Bombardier, Anna Marie C. Marcelino, Jiang Hong, and Lila M. Gierasch. From the test tube to the cell: Exploring the floding and aggregation of a beta-clam protein. *BIOPOLYMERS*, 88(2):157–163, 2007. 4th Peptide Engineering Meeting (PEM4), Yokohama, JAPAN, NOV 05-08, 2006.

[36] J JAEKEN and G VANDENBERGHE. AN INFANTILE AUTISTIC SYNDROME CHARACTERIZED BY THE PRESENCE OF SUCCINYL PURINES IN BODY-FLUIDS. *LANCET*, 2(8411):1058–1061, 1984.

[37] J JAEKEN and G VANDENBERGHE. Heterogeneity in adenylosuccinate deficiency. *Pediatr. Res*, 19:1075–1075, 1985.

[38] J JAEKEN, SK WADMAN, M DURAN, FJ VANSPRANG, FA BEEMER, RA HOLL, PM THEUNISSEN, P DECOCK, F VANDENBERGH, MF VINCENT, and G VANDENBERGHE. ADENYLOSUCCINASE DEFICIENCY - AN INBORN ERROR OF PURINE NUCLEOTIDE SYNTHESIS. *EUROPEAN JOURNAL OF PEDIATRICS*, 148(2):126–131, NOV 1988.

[39] JL Johnson and EA Craig. Protein folding in vivo: Unraveling complex pathways. *CELL*, 90(2):201–204, JUL 25 1997.

[40] A. Jurecka. Inborn errors of purine and pyrimidine metabolism. *JOURNAL OF INHERITED METABOLIC DISEASE*, 32(2):247–263, APR 2009.

[41] Agnieszka Jurecka, Elzbieta Jurkiewicz, and Anna Tylki-Szymanska. Magnetic resonance imaging of the brain in adenylosuccinate lyase deficiency: a report of seven cases and a review of the literature. *EUROPEAN JOURNAL OF PEDIATRICS*, 171(1):131–138, JAN 2012.

[42] Agnieszka Jurecka, Marie Zikanova, Anna Tylki-Szymanska, Jakub Krijt, Anna Bogdanska, Wanda Gradowska, Karolina Mullerova, Jolanta Sykut-Cegielska, Stanislav Kmoch, and Ewa Pronicka. Clinical, biochemical and molecular findings in seven Polish patients with adenylosuccinate lyase deficiency. *MOLECULAR GENETICS AND METABOLISM*, 94(4):435–442, AUG 2008.

[43] Kirstie E. Keller, Irene S. Tan, and Young-Sam Lee. SAICAR Stimulates Pyruvate Kinase Isoform M2 and Promotes Cancer Cell Survival in Glucose-Limited Conditions. *SCIENCE*, 338(6110):1069–1072, NOV 23 2012.

[44] S Kmoch, H Hartmannova, B Stiburkova, J Krijt, M Zikanova, and I Sebesta. Human adenylosuccinate lyase (ADSL), cloning and characterization of full-length cDNA and its isoform, gene structure and molecular basis for ADSL deficiency in six patients. *HUMAN MOLECULAR GENETICS*, 9(10):1501–1513, JUN 12 2000.

[45] Aaron J. Knox, Christine Graham, John Bleskan, Gary Brodsky, and David Patterson. Mutations in the Chinese hamster ovary cell GART gene of de novo purine synthesis. *GENE*, 429(1-2):23–30, JAN 15 2009.

[46] Paula Leandro and Claudio M. Gomes. Protein misfolding in conformational disorders: Rescue of folding defects and chemical chaperoning. *MINI-REVIEWS IN MEDICINAL CHEMISTRY*, 8(9):901–911, AUG 2008.

[47] Peychii Lee and Roberta F. Colman. Expression, purification, and characterization of stable, recombinant human adenylosuccinate lyase. *PROTEIN EXPRESSION AND PURIFICATION*, 51(2):227–234, FEB 2007.

[48] TT Lee, C Worby, ZQ Bao, JE Dixon, and RF Colman. Implication of His(68) in the substrate site of Bacillus subtilis adenylosuccinate lyase by mutagenesis and affinity labeling with 2-[(4-bromo-2,3-dioxobutyl)thio]adenosine 5 '-monophosphate. *BIOCHEMISTRY*, 37(23):8481–8489, JUN 9 1998.

[49] TT Lee, C Worby, ZQ Bao, JE Dixon, and RF Colman. His(68) and His(141) are critical contributors to the intersubunit catalytic site of adenylosuccinate lyase of Bacillus subtilis. *BIOCHEMISTRY*, 38(1):22–32, JAN 5 1999.

[50] TT Lee, C Worby, JE Dixon, and RF Colman. Identification of His(141) in the active site of Bacillus subtilis adenylosuccinate lyase by affinity labeling with 6-(4-bromo-2,3-dioxobutyl)thioadenosine 5'-monophosphate.

JOURNAL OF BIOLOGICAL CHEMISTRY, 272(1):458–465, JAN 3 1997.

[51] LEVINTHAL.C. ARE THERE PATHWAYS FOR PROTEIN FOLDING. *JOURNAL DE CHIMIE PHYSIQUE ET DE PHYSICO-CHIMIE BIOLOGIQUE*, 65(1):44–&, 1968.

[52] Leila M. Luheshi, Damian C. Crowther, and Christopher M. Dobson. Protein misfolding and disease: from the test tube to the organism. *CURRENT OPINION IN CHEMICAL BIOLOGY*, 12(1):25–31, FEB 2008.

[53] Claire T. Lundy, Heinz Jungbluth, Keith R. E. Pohl, Ata Siddiqui, Anthony M. Marinaki, Helen Mundy, and Michael P. Champion. Adenylosuccinate Lyase Deficiency in the United Kingdom Pediatric Population: First Three Cases. *PEDIATRIC NEUROLOGY*, 43(5):351–354, NOV 2010.

[54] S Marie, H Cuppens, M Heuterspreute, M Jaspers, EZ Tola, XX Gu, E Legius, MF Vincent, J Jaeken, JJ Cassiman, and G Van den Berghe. Mutation analysis in adenylosuccinate lyase deficiency: Eight novel mutations in the re-evaluated full ADSL coding sequence. *HUMAN MUTATION*, 13(3):197–202, 1999.

[55] Mirela Matecic, Daniel L. Smith, Jr., Xuewen Pan, Nazif Maqani, Stefan Bekiranov, Jef D. Boeke, and Jeffrey S. Smith. A Microarray-Based Genetic Screen for Yeast Chronological Aging Factors. *PLOS GENETICS*, 6(4), APR 2010.

[56] William H. Mather, Natalie A. Cookson, Jeff Hasty, Lev S. Tsimring, and Ruth J. Williams. Correlation Resonance Generated by Coupled Enzymatic Processing. *BIOPHYSICAL JOURNAL*, 99(10):3172–3181, NOV 17 2010.

[57] Hanna Mierzewska, Bogna Schmidt-Sidor, Elzbieta Jurkiewicz, Anna Bogdanska, Katarzyna Kusmierska, and Tomasz Stepien. Severe encephalopathy with brain atrophy and hypomyelination due to adenylosuccinate lyase deficiency - MRI, clinical, biochemical and neuropathological findings of Polish patients. *FOLIA NEUROPATHOLOGICA*, 47(4):314–320, 2009.

[58] Katharina Mouchegh, Marie Zikanova, Georg F. Hoffmann, Benno Kretzschmar, Thomas Kuehn, Eva Mildenberger, Gisela Stoltenburg-Didinger, Jakub Krijt, Lenka Dvorakova, Tomas Honzik, Jiri Zeman, Stanislav Kmoch, and Rainer Rossi. Lethal fetal and early neonatal presentation of adenylosuccinate lyase deficiency: Observation of 6 patients in 4 families. *JOURNAL OF PEDIATRICS*, 150(1):57–61, JAN 2007.

[59] Joydeep Mukherjee, Joanna J. Phillips, Shichun Zheng, John Wiencke, Sabrina M. Ronen, and Russell O. Pieper. Pyruvate Kinase M2 Expression, but Not Pyruvate Kinase Activity, Is Up-Regulated in a Grade-Specific Manner in Human Glioma. *PLOS ONE*, 8(2), FEB 25 2013.

[60] JK MYERS, CN PACE, and JM SCHOLTZ. DENATURANT M-VALUES AND HEAT-CAPACITY CHANGES - RELATION TO CHANGES IN ACCESSIBLE SURFACE-AREAS OF PROTEIN UNFOLDING. *PROTEIN SCIENCE*, 4(10):2138–2148, OCT 1995.

[61] CN PACE. CONFORMATIONAL STABILITY OF GLOBULAR-PROTEINS. *TRENDS IN BIOCHEMICAL SCIENCES*, 15(1):14–17, JAN 1990.

[62] JB Palenchar and RF Colman. Characterization of a mutant Bacillus subtilis adenylosuccinate lyase equivalent to a mutant enzyme found in human adenylosuccinate lyase deficiency: Asparagine 276 plays an important structural role. *BIOCHEMISTRY*, 42(7):1831–1841, FEB 25 2003.

[63] JB Palenchar, JM Crocco, and RF Colman. The characterization of mutant Bacillus subtilis adenylosuccinate lyases corresponding to severe human adenylosuccinate lyase deficiencies. *PROTEIN SCIENCE*, 12(8):1694–1705, AUG 2003.

[64] Benoit Pinson, Sabine Vaur, Isabelle Sagot, Fanny Coulpier, Sophie Lemoine, and Bertrand Daignan-Fornier. Metabolic intermediates selectively stimulate transcription factor interaction and modulate phosphate and purine pathways. *GENES & DEVELOPMENT*, 23(12):1399–1407, JUN 15 2009.

[65] ET Powers and DL Powers. A perspective on mechanisms of protein tetramer formation. *BIOPHYSICAL JOURNAL*, 85(6):3587–3599, DEC 2003.

[66] PL PRIVALOV and SJ GILL. STABILITY OF PROTEIN-STRUCTURE AND HYDROPHOBIC INTERACTION. *ADVANCES IN PROTEIN CHEMISTRY*, 39:191–234, 1988.

[67] V Race, S Marie, MF Vincent, and G Van den Berghe. Clinical, biochemical and molecular genetic correlations in adenylosuccinate lyase deficiency. *HUMAN MOLECULAR GENETICS*, 9(14):2159–2165, SEP 1 2000.

[68] Sheena E. Radford. Protein folding: progress made and promises ahead. *Trends in Biochemical Sciences*, 25(12):611 – 618, 2000.

[69] Roee Ramot, Krishna Kishore Inampudi, and Corey J. Wilson. Lactose repressor experimental folding landscape: Fundamental functional unit and tetramer folding mechanisms. *Biochemistry*, 51(38):7569–7579, 2012.

[70] Stephen P. Ray, Michelle K. Deaton, Glenn C. Capodagli, Lauren A. F. Calkins, Lucas Sawle, Kingshuk Ghosh, David Patterson, and Scott D. Pegan. Structural and Biochemical Characterization of Human Adenylosuccinate Lyase (ADSL) and the R303C ADSL Deficiency-Associated Mutation. *BIOCHEMISTRY*, 51(33):6701–6713, AUG 21 2012.

[71] K Rebora, C Desmoucelles, F Borne, B Pinson, and B Daignan-Fornier. Yeast AMP pathway genes respond to adenine through regulated synthesis of a metabolic intermediate. *YEAST*, 18(1):S217, AUG 2001.

[72] K Rebora, B Laloo, and B Daignan-Fornier. Revisiting purine-histidine cross-pathway regulation in Saccharomyces cerevisiae: A central role for a small molecule. *GENETICS*, 170(1):61–70, MAY 2005.

[73] AD Robertson and KP Murphy. Protein structure and the energetics of protein stability. *CHEMICAL REVIEWS*, 97(5):1251–1267, JUL-AUG 1997.

[74] Yannick Rondelez. Competition for Catalytic Resources Alters Biological Network Dynamics. *PHYSICAL REVIEW LETTERS*, 108(1), JAN 5 2012.

[75] ML Segall and RF Colman. Gln(212), Asn(270), and Arg(301) are critical for catalysis by adenylosuccinate lyase from Bacillus subtilis. *BIOCHEMISTRY*, 43(23):7391–7402, JUN 15 2004.

[76] S Sivendran, D Patterson, E Spiegel, I McGown, D Cowley, and RF Colman. Two novel mutant human adenylosuccinate Lyases (ASLs) associated with autism and characterization of the equivalent mutant Bacillus subtilis ASL. *JOURNAL OF BIOLOGICAL CHEMISTRY*, 279(51):53789–53797, DEC 17 2004.

[77] Sharmila Sivendran and Roberta F. Colman. Effect of a new non-cleavable substrate analog on wild-type and serine mutants in the signature sequence of adenylosuccinate lyase of Bacillus subtilis and Homo sapiens. *PROTEIN SCIENCE*, 17(7):1162–1174, JUL 2008.

[78] Erin K. Spiegel, Roberta F. Colman, and David Patterson. Adenylosuccinate lyase deficiency. *MOLECULAR GENETICS AND METABOLISM*, 3(1-2):19–31, SEP-OCT 2006.

[79] RL STONE, H ZALKIN, and JE DIXON. EXPRESSION, PURIFICATION, AND KINETIC CHARACTERIZATION OF RECOMBINANT HUMAN ADENYLOSUCCINATE LYASE. *JOURNAL OF BIOLOGICAL CHEMISTRY*, 268(26):19710–19716, SEP 15 1993.

[80] TW Stone, LA Roberts, BJ Morris, PA Jones, HA Ogilvy, WMH Behan, JA Duley, HA Simmonds, MF Vincent, and G van den Berghe. Succinylpurines induce neuronal damage in the rat brain. In Griesmacher, A and Chiba, P and Muller, MM, editor, *PURINE AND PYRIMIDINE METABOLISM IN MAN IX*, volume 431 of *ADVANCES IN EXPERIMENTAL MEDICINE AND BIOLOGY*, pages 185–189. PLENUM PRESS DIV PLENUM PUBLISHING CORP, 233 SPRING ST, NEW YORK, NY 10013 USA, 1998.

[81] M Tanaka, Y Machida, SY Niu, T Ikeda, NR Jana, H Doi, M Kurosawa, M Nekooki, and N Nukina. Trehalose alleviates polyglutamine-mediated pathology in a mouse model of Huntington disease. *NATURE MEDICINE*, 10(2):148–154, FEB 2004.

[82] EA Toth and TO Yeates. The structure of adenylosuccinate lyase, an enzyme with dual activity in the de novo purine biosynthetic pathway. *STRUCTURE WITH FOLDING & DESIGN*, 8(2):163–174, FEB 15 2000.

[83] May Tsai, Jason Koo, Patrick Yip, Roberta F. Colman, Mark L. Segal, and P. Lynne Howell. Substrate and product complexes of Escherichia coli adenylosuccinate lyase provide new insights into the enzymatic mechanism. *JOURNAL OF MOLECULAR BIOLOGY*, 370(3):541–554, JUL 13 2007.

[84] AS TU and D PATTERSON. BIOCHEMICAL GENETICS OF CHINESE-HAMSTER CELL MUTANTS WITH DEVIANT PURINE

METABOLISM .6. ENZYMATIC STUDIES OF 2 MUTANTS UNABLE TO CONVERT INOSINIC ACID TO ADENYLIC ACID. *BIOCHEMICAL GENETICS*, 15(1-2):195–210, 1977.

[85] AH Van Gennip. Defects in metabolism of purines and pyrimidines. *Ned tijdsh Klin Chem*, 24:171–175, 1999.

[86] F VANDENBERGH, MF VINCENT, J JAEKEN, and G VANDENBERGHE. FUNCTIONAL-STUDIES IN FIBROBLASTS OF ADENYLOSUCCINASE-DEFICIENT CHILDREN. *JOURNAL OF INHERITED METABOLIC DISEASE*, 16(2):425–434, 1993.

[87] F VANDENBERGH, MF VINCENT, J JAEKEN, and G VANDENBERGHE. RESIDUAL ADENYLOSUCCINASE ACTIVITIES IN FIBROBLASTS OF ADENYLOSUCCINASE-DEFICIENT CHILDREN - PARALLEL DEFICIENCY WITH ADENYLOSUCCINATE AND SUCCINYL-AICAR IN PROFOUNDLY RETARDED PATIENTS AND NONPARALLEL DEFICIENCY IN A MILDLY RETARDED GIRL. *JOURNAL OF INHERITED METABOLIC DISEASE*, 16(2):415–424, 1993. 4TH INTERNATIONAL SYMP ON THE NEURONAL CEROID-LIPOFUSCINOSIS (BATTEN DISEASE), HAMBURG, GERMANY, JUN 11-13, 1992.

[88] G. Vandenberghe and J. Jaeken. Adenylosuccinase deficiency. *Pediatr. Res*, 19:780–780, 1985.

[89] Florence Verrier, Songon An, Ann M. Ferrie, Haiyan Sun, Minjoung Kyoung, Huayun Deng, Ye Fang, and Stephen J. Benkovic. GPCRs regulate

the assembly of a multienzyme complex for purine biosynthesis. *NATURE CHEMICAL BIOLOGY*, 7(12):909–915, DEC 2011.

[90] Lydia K. Vliet, Terry G. Wilkinson II, Nathan Duval, Guido Vacano, Christine Graham, Marie Ziknov, Vaclava Skopova, Veronika Baresova, Ale Hnzda, Stanislav Kmoch, and David Patterson. Molecular characterization of the adei mutant of chinese hamster ovary cells: A cellular model of adenylosuccinate lyase deficiency. *Molecular Genetics and Metabolism*, 102(1):61 – 68, 2011.

[91] Richard W.E. Watts. Some regulatory and integrative aspects of purine nucleotide biosynthesis and its control: An overview. *Advances in Enzyme Regulation*, 21(0):33 – 51, 1983.

[92] Deborah K. Wilkins, Shaun B. Grimshaw, Vronique Receveur, Christopher M. Dobson, Jonathan A. Jones, and Lorna J. Smith. Hydrodynamic radii of native and denatured proteins measured by pulse field gradient nmr techniques. *Biochemistry*, 38(50):16424–16431, 1999.

[93] Hao Wu, Bin-Guang Ma, Ji-Tao Zhao, and Hong-Yu Zhang. How similar are amino acid mutations in human genetic diseases and evolution. *BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS*, 362(2):233–237, OCT 19 2007.

[94] XM Zhang, XT Wang, HW Yue, SW Leung, PH Thibodeau, PJ Thomas, and SE Guggino. Organic solutes rescue the functional defect in Delta F508 cystic fibrosis transmembrane conductance regulator. *JOURNAL OF BIOLOGICAL CHEMISTRY*, 278(51):51232–51242, DEC 19 2003.

[95] JH Zhao, HL Liu, HY Lin, CH Huang, HW Fang, SS Chen, Y Ho, WB Tsai, and WY Chen. Chemical chaperone and inhibitor discovery: potential treatments for protein conformational diseases. *Perspect Medicin Chem*, 11(1):39–48, Dec 2007.

[96] Huan-Xiang Zhou, German Rivas, and Allen P. Minton. Macromolecular crowding and confinement: Biochemical, biophysical, and potential physiological consequences. *ANNUAL REVIEW OF BIOPHYSICS*, 37:375–397, 2008.

[97] M Zikanova, J Krijt, H Hartmannova, and S Kmoch. Preparation of 5-amino-4-imidazole-N-succinocarboxamide ribotide, 5-amino-4-imidazole-N-succinocarboxamide riboside and succinyladenosine, compounds usable in diagnosis and research of adenylosuccinate lyase deficiency. *JOURNAL OF INHERITED METABOLIC DISEASE*, 28(4):493–499, AUG 2005.

[98] Marie Zikanova, Vaclava Skopova, Ales Hnizda, Jakub Krijt, and Stanislav Kmoch. Biochemical and Structural Analysis of 14 Mutant ADSL Enzyme Complexes and Correlation to Phenotypic Heterogeneity of Adenylosuccinate Lyase Deficiency. *HUMAN MUTATION*, 31(4):445–455, APR 2010.

Appendix A

Materials and Methods

A.1 Materials

Chemicals, biochemicals, buffers, and solvents were purchased from Sigma-Aldrich Chemical Co. (St. Louis, MO), Fisher Scientific Inc. (Pittsburgh, PA), Fluka Chemical Corp. (Milwaukee, WI), or EM Science (Cincinnati, OH). The Centricon and Ultrafree centrifugal filter devices were obtained from Millipore Co. (Billerica, MA). Nickel-nitrilotriacetic acid-agarose, a QIAspin kit, and high throughput crystal condition screens were purchased from QIAGEN. Additive HT Screen was purchased from Hampton Research. QuikChange site directed mutagenesis kit was purchased from Stratagene. SAICAR was prepared enzymatically from AICAR purchased from Sigma-Aldrich Chemical Co. as described by Zikanova, et al [97]. Enzymes and reagents used for molecular biology procedures were obtained from New England Biolabs, Inc. (Ipswich, MA).

A.2 Methods

A.2.1 Site-Directed Mutagenesis, Enzyme Expression, and Purification

The initial WT ADSL construct was obtained from Dr. Roberta F. Colman, Department of Chemistry and Biochemistry, University of Delaware, Newark, DE. The full description of the initial ADSL WT construct is described in Lee and Colman [47]. In short, the full length human ADSL gene (1-484 residues) was constructed in pET-14b vector containing a 5'-end NdeI restriction site and a 3'-end BlpI restriction site and a thrombin cleavable N-terminal histidine tag. In order to overexpress the human enzyme in E. coli, the vector was transformed into E. coli Rosetta 2(DE3)pLysS. Introduction of point mutations in the human ADSL plasmid was done by the QuikChange site-direction mutagenesis. The QIAspin kit was used for cDNA extraction and purification. DNA sequencing performed at CU Cancer Center DNA Sequencing and Analysis Core confirmed mutations.

ADSL protein was then grown in LB broth at 37°C until cell density (OD_{600}) reached 0.4-0.6. The cell culture was then cooled to 25°C and induced with 0.4 mM IPTG overnight. Cells were pelleted by centrifugation at x2500 g. The cell pellet was resuspended in cell lysis buffer (50 mM potassium phosphate buffer, pH 8.0, 300 mM KCl, and 10% glycerol) based on 10 mL lysis buffer per gram of cell pellet. DNase (1 mg/mL) and RNase (25 mg/mL) were added to the lysis buffer at 0.001 (w/v) and 0.01 (w/v) for DNase and RNase respectively. The pellet was freeze-thawed three times to

rupture the bacterial cell membrane. The soluble protein fraction (crude cell lysate) was separated from cell debris by centrifugation at x32,000 g. The crude cell lysate was then loaded onto a Qiagen Ni-NTA column equilibrated with lysis buffer. After loading the crude cell lysate, the Ni-NTA column was washed with 100 mL lysis buffer, followed by 100 mL lysis buffer with 20 mM imidazole to remove any loosely bound protein. The histidine tagged ADSL was eluted by a column gradient or 150 mL of lysis buffer containing 20 mM imidazole and 250 mM imidazole. ADSL was collected in 4 mL fractions with 400 μL of enzyme storage buffer (50 mM potassium phosphate, pH. 7.0, 150 mM KCL, 1 mM DTT, 1 mM EDTA, 10% glycerol). Fractions were tested for ADSL concentration by the Bio-Rad assay. Fractions with high ADSL concentration were pooled and concentrated using Amicon Ultra-15 (MWCO 10,000). The concentrated enzyme was dialyzed against enzyme storage buffer overnight. Purity was assessed by SDS-PAGE Gel Electrophoresis (data not shown). Protein Concentration was calculated by absorbance at 280 nm using ${E_{280}}^{1\%} = 7.7$ [29]. After purification, protein was stored at -80°C.

A.2.2 Population Distribution by Static Light Scattering

Size Exclusion Chromatography-Multi-Angle Light Scattering measurements were performed on a Wyatt miniDAWN TREOS connected to a Shimadzu UFLC and a Wyatt WTC-030S5 size exclusion column. Samples were run in duplicate of 350 μg of His-tagged ADSL diluted in Enzyme Storage buffer. PBS was used as the mobile phase and flowed at .5 mL/min. Calibration

was checked using 2 mg/mL bovine serum albumin. Data were analyzed with ASTRA software. Values of .185 and 770 were used as the $\frac{dn}{dc}$ value and extinction coefficient ($\frac{mL}{g*cm}$) respectively. The extinction coefficient was calculated by previously established methods [29].

A.2.3 Population Distribution by Analytical Ultracentrifugation

Sedimentation velocity (SV) experiments of the WT and mutant enzymes were conducted using a Proteomelab XL-I analytical ultracentrifuges (Beckman Coulter, Fullerton, CA) analytical ultracentrifuge equipped with an An-60Ti analytical rotor. Prior to each SV experiment, frozen samples were incubated for ∼2 hours at 25°C before measurements were taken. ADSL samples (∼0.35 mg/mL) were centrifuged at 40,000 rpm at 25°C. The number of scans collected in each run was adequate to capture the entire sedimentation process. Data were analyzed using the SEDPHAT program. The density of the buffer at 25°C was calculated using the SEDNTERP program. The partial specific volume of human ASL at 25°C is 0.7366.

A.2.4 Secondary Structure by Circular Dichroism

The secondary structure of the WT and mutant enzymes was assessed using CD spectroscopy. Ellipticity was measured on a Jasco - J815 spectropolarimeter from 200 to 250 nm, in 0.2 nm increments using a 0.1 cm quartz cuvette. The samples were scanned three times and averaged, and the background from the buffer (50 mM KPO4 buffer, pH 7, containing 150 mM KCl, 1 mM EDTA,

1 mM DTT, and 10% glycerol) was subtracted. The mean molar residue ellipticity $[\theta]$ (deg cm^2 dmol^{-1}) was calculated from the equation $[\theta]=\frac{\theta}{10*n*C*\ell}$, where θ is the measured ellipticity in millidegrees, C is the molar concentration of enzyme subunits, ℓ is the path length in centimeters, and n is the number of residues per subunit (503 residues per monomer, including the His6 tag and thrombin cleavage site). The WT and mutant enzyme samples were incubated for ∼2 hours at 25°C before measurements were taken. The ellipticity of the samples was measured at 25°.

A.2.5 Enzyme Kinetics by UV

Enzyme kinetics was performed on a UV-Vis Spectrophotometer Evolution 3000 from Thermo Scientific using 1 mL quartz cuvettes at 25°C. ADSL with the His-tag intact was used for enzyme assays, as it has been shown that the His-tag does not affect enzyme activity [47]. Experiments were run with concentrations of 0.11 mg/mL and 0.19 mg/mL for WT and R303C respectively. Frozen samples were incubated for ∼2 hours at 25°C before measurements were taken. SAMP enzyme assays of ADSL were measured at 25°C in 40 mM Tris-HCl (pH 7.4) with varying concentrations of SAMP (1 - 60 μM). Specific activity was measured from the decrease in absorbance of SAMP at 282 nm as it was converted to AMP and fumarate. The assay was monitored over 30 seconds in a 1 mL reaction volume. The difference extinction coefficient of 10000 M^{-1} cm^{-1} between SAMP and AMP was used to calculate the specific activity. SAICAR enzyme assays of ADSL were measured at 25°C in 40 mM Tris-HCl (pH 7.4) with varying concentrations of SAICAR (1-60 μM). Specific activity was measured from the decrease in absorbance of SAICAR at 269 nm

as it was converted to AICAR and fumarate. The assay was monitored over 30 seconds in a 1 mL reaction volume. The difference extinction coefficient of 700 M^{-1} cm^{-1} between SAICAR and AICAR was used to calculate the specific activity. The initial velocities (ν) and kinetic constants (V_{max}, $K_{0.5}$, and $\frac{k_{cat}}{K_{0.5}}$) were calculated by custom Mathematica templates. The initial velocity was calculated by importing the absorbance vs. time graphs into Mathematica where the linear portion of the graph was fit with a linear line. The change in absorbance (ΔA) was calculated for one minute from the linear fit. The initial velocity was calculated by Beer's Law, $\Delta C = \frac{\Delta A}{\varepsilon \ell}$, where ΔC is the change in concentration over time, ΔA is the change in absorbance over time, ε is the difference extinction coefficient, and ℓ is the pathlength. The initial velocity was put in terms of specific activity (μmol min^{-1} mg^{-1}) by dividing ν by the total enzyme injected into the reaction volume. To determine the kinetic constants, the initial velocity data were fitted to the Hill equation, $\nu = \frac{V_{max} * [S]^n}{K_{0.5}^n + [S]^n}$. The k_{cat} value was calculated from V_{max} and the enzyme concentration [E] via the equation $k_{cat} = \frac{V_{max}}{[E]}$.

A.2.6 Enzyme Kinetics by HPLC-EC

Enzyme activity was measured by HPLC-EC analysis of AMP and AICAR formed from SAMP and SAICAR respectively. Assays were performed at 25°C in 40 mM Tris-HCl (pH 7.4) with varying concentrations of SAMP and SAICAR (0-500 μM). Experiments were run in duplicate with approximately 250-500 μg ADSL injected into 1 mL reaction volume at 100% SAMP, 100% SAICAR, and different ratios of SAMP to SAICAR concentration: 3:1, 1:1, and 1:3. Every minute, 100 μL of the reaction mixture was extracted into

25 μL of 2.5 M perchloric acid on ice to stop the reaction. The solution was neutralized by adding 25 μL of 2.5 M potassium hydroxide. Samples were clarified by centrifugation in a Spectrafuge 16M at 14,000 rpm at -5°C.

Separation of SAICAR, AICAR, SAMP, and AMP was achieved by HPLC-EC analysis similar to our previously described method [21]. Briefly, separation was obtained by a reverse phase HPLC-EC with a TSKgel ODS-80Tm C-18 column (250 mm x 4.6 mm ID, 5 μM) protected by Tosoh Bioscience TSKgel guard cartridge. A column temperature of 35°C was maintained throughout the analysis. A mobile phase consisting of 50 mM lithium acetate, 2% acetonitrile, and 5 mM tetrabutyl ammonium phosphate (TBAP) at pH 4.8 was delivered isocratically at a flow rate of 0.8 mL/min. Sample extracts and standards were kept at 10°C until a 30 μL aliquot of each sample was injected using an ESA autosampler (model 542). After injection and separation, analytes were detected using a CoulArray HPLC system (model 5600A, ESA) with three electrochemical detector modules. Each module contains four flow-through coulometric detectors in series. Nine EC channels were set to a range of potentials from 0 - 900 mV in 100 mV increments. Two channels were used for a high potential boron-doped diamond (BDD) amperometric detector set to 1700 mV and UV detector set at 265 nm.

This HPLC-EC software allows for quantitative measurement of the amounts of both substrates and products. Peak quantitation was done using the CoulArray software package with manual curation to assure appropriate peak calling by the software. Standard curves of concentration versus peak area in units

of coulombs were generated for AICAR, and AMP. For these experiments the standard curves consisted of a concentration range of 2-150 μM for the products and fit with a linear trendline. Time points 0-5 minutes for each mixing ratio were analyzed. AMP and AICAR concentrations were calculated from the standard curve, plotted, and fit with a linear line. The activity was given by the slope of the line in units of Molarity*min^{-1} and changed into units of μmol*min^{-1}*mg^{-1}.

A.2.7 Resource Sharing Kinetics Model

We carried out an enzyme mixing kinetics analysis to predict the rate of product formation when two substrates catalyzed by the same enzyme are present in solution at different concentrations. The analysis is based on the coupling of two independent enzymatic reactions via resource sharing. By independent reactions, we mean two separate reactions for each substrate in the absence of the other. However, resource sharing induces a coupling due to the fact that if one enzyme molecule is occupied by substrate one, it is not available for substrate two. Although the enzyme has multiple binding sites, we assume on a given enzyme each site is occupied by substrates of the same type thus avoiding any other possible sources of coupling that may alter the equilibrium constant (K_m). Thus competitive binding has been enforced in the simplest form by conserving the total number of resources between the two substrates. This is the only source of interaction between the two reactions in our resource sharing model. Based on work by Ray, Deaton et al, cooperativity may play a small role in substrate binding and catalysis [70]. Therefore, for simplicity, we disregard cooperativity in the present model.

From these assumptions, the rate of product formation for each substrate in mixing conditions can be predicted. The analysis starts with two independent kinetics reactions:

$$[E] + [S_1] \underset{\mathrm{k}_{-1}}{\overset{\mathrm{k}_1}{\rightleftharpoons}} [E \cdot S_1] \xrightarrow{k_{cat,1}} [E] + [P_1] \tag{A.2.1}$$

$$[E] + [S_2] \underset{\mathrm{k}_{-2}}{\overset{\mathrm{k}_2}{\rightleftharpoons}} [E \cdot S_2] \xrightarrow{k_{cat,2}} [E] + [P_2] \tag{A.2.2}$$

where $[E]$ = ADSL, $[S_1]$ = SAICAR, $[P_1]$ = AICAR, $[S_2]$ = SAMP, and $[P_2]$= AMP. The two independent enzyme reactions indicate that a single enzyme molecule binds $[S_1]$ or $[S_2]$. Using our assumption no cooperativity, the rate of product formation,$\frac{d[P]}{dt}$, in mixing conditions can be predicted in the following manner:

Applying the Law of Mass Action to $[P_1]$ and the $[E \cdot S_1]$ complex in equation A.2.1

$$\frac{d[P_1]}{dt} = k_{cat,1} \cdot [E \cdot S_1] \tag{A.2.3}$$

$$\frac{d[E \cdot S_1]}{dt} = k_1 \cdot [E][S_1] - k_{-1} \cdot [E \cdot S_1] - k_{cat,1} \cdot [E \cdot S_1] \tag{A.2.4}$$

Assuming a steady-state approximation, $\frac{d[E \cdot S_1]}{dt} = 0$, leads to

$$\frac{[E][S_1]}{[E \cdot S_1]} = \frac{k_{-1} + k_{cat,1}}{k_1} = K_{m,1} \tag{A.2.5}$$

Applying the same to equation A.2.2

$$\frac{d[P_2]}{dt} = k_{cat,2} \cdot [E \cdot S_2] \tag{A.2.6}$$

$$\frac{[E][S_2]}{[E \cdot S_2]} = \frac{k_{-2} + k_{cat,2}}{k_2} = K_{m,2} \tag{A.2.7}$$

If both substrates are present, the total enzyme concentration, $[E_{total}]$, can be written as

$$[E_{total}] = [E] + [E \cdot S_1] + [E \cdot S_2] \tag{A.2.8}$$

Solving for $[E]$ from equation A.2.5 and equation A.2.7

$$[E] = \frac{[E \cdot S_1]K_{m,1}}{[S_1]} = \frac{[E \cdot S_2]K_{m,2}}{[S_2]} \tag{A.2.9}$$

Utilizing equation A.2.9 to write equation A.2.8 in terms of $[E \cdot S_1]$

$$[E_{total}] = \frac{[E \cdot S_1]K_{m,1}}{[S_1]} + [E \cdot S_1] + \frac{[E \cdot S_1] \cdot K_{m,1} \cdot [S_2]}{[S_1] \cdot K_{m,2}} \tag{A.2.10}$$

Solving for $[E \cdot S_1]$ from equation A.2.10

$$[E \cdot S_1] = \frac{[E_{total}]}{\frac{K_{m,1}}{[S_1]} + \frac{K_{m,1}[S_2]}{K_{m,2}[S_1]} + 1} \tag{A.2.11}$$

Substituting equation A.2.11 into equation A.2.3

$$\frac{d[P_1]}{dt} = \frac{k_{cat,1}[E_{total}]}{\frac{K_{m,1}}{[S_1]} + \frac{K_{m,1}[S_2]}{K_{m,2}[S_1]} + 1} = \frac{V_{max,1}}{\frac{K_{m,1}}{[S_1]} + \frac{K_{m,1}[S_2]}{K_{m,2}[S_1]} + 1} \tag{A.2.12}$$

Since our experiment was always in saturating conditions, $K_{m,1} << S_1$

$$\frac{d[P_1]}{dt} = \frac{V_{max,1}}{\frac{K_{m,1}[S_2]}{K_{m,2}[S_1]} + 1} \tag{A.2.13}$$

Setting x = $\frac{[S_2]}{[S_1]}$

$$\frac{d[P_1]}{dt} = \frac{V_{max,1}}{\frac{K_{m,1}}{K_{m,2}} \cdot x + 1} \tag{A.2.14}$$

Utilizing equation A.2.9 to write equation A.2.8 in terms of $[E \cdot S_2]$, , solving for $[E \cdot S_2]$, and substituting into equation A.2.6 results in the rate of formation for $[P_2]$

$$\frac{d[P_2]}{dt} = \frac{V_{max,2}}{\frac{K_{m,2}}{K_{m,1}} \cdot \frac{1}{x} + 1} \tag{A.2.15}$$

Equations A.2.14 and A.2.15 are predictions for the rate of product formation, d$[P]$/dt, as a function of substrate ratio, x. V_{max} and K_m values are obtained experimentally by enzyme kinetic assays on each substrate independently.

A.2.8 Two Site, Two Substrate Model

I did not have an opportunity to really delve into this model, and it seems pretty moot at this point to continue. But it was fun to think about and was good practice.

(*Initial concentrations for reaction*)

i0=4; (*ADSL*)

J0=50;(*SAMP*)

P0=0;(*AMP*)

s0=50;(*SAICAR*)

A0=0;(*AICAR*)

(*complexes*)

iJ0=0;

is0=0;

Ji0=0;

si0=0;

JiJ0=0;

sis0=0; siJ0=0;

Jis0=0;

(*rate constants*)

k_1=1;

k_{-1}=0;

k_2=1;

k_{-2}=0;

k_3=1;

k_{-3}=0;

k_4=1;

k_{-4}=0;

(*catalytic constants*)

k_{stoA}=7.4*0.43;

k_{JtoP}=13.5*0.052;

(*coop and catalytic rate constants*)

α=0.97; (*coop for after J binds for another J to bind*)

β=1; (*changes in catalytic rate with two active sites filled with J*)

γ=0.94; (*coop for after s binds for another s to bind*)

ϕ=1;(*changes in catalytic rate with two active sites filled with s*)

θ=1; (*coop for after s binds for another J to bind*)

ψ=1;(*changes in catalytic rate with two active sites filled with s and J*)

η=1; (*coop for after J binds for another s to bind*)

ϵ=1;(*changes in catalytic rate with two active sites filled with J and s*)

(*Time constants*)

NDSolveTime=2000;(*How far out to solve the differential equations*)

solution=NDSolve[i'[t]==-k_1 * i[t] * J[t] + k_{-1} * iJ[t] + k_{JtoP} * iJ[t] - k_3 * i[t] * s[t] + k_{-3} * is[t] + k_{stoA} * is[t] - k_2 * i[t] * J[t] + k_{-2} * Ji[t] - k_4 * i[t] * s[t] + k_{-4} * si[t], (* + k_{stoA} * si[t] + k_{JtoP} * Ji[t], *)

J'[t]==-k_1 * i[t] * J[t] + k_{-1}iJ[t] - α * k_1 * Ji[t] * J[t] + α * k_{-1}JiJ[t] - θ * k_1 * si[t] * J[t] + θ * k_{-1}siJ[t] - k_2 * i[t] * J[t] + k_{-2}Ji[t] - α * k_2 * iJ[t] * J[t] + α * k_{-2}JiJ[t] - θ * k_2 * is[t] * J[t] + θ * k_{-2} * Jis[t],

iJ'[t]==k_1 * J[t] * i[t] - k_{-1} * iJ[t] - k_{JtoP} * iJ[t] - α * k_2 * iJ[t] * J[t] + α * k_{-2} * JiJ[t] - η * k_4 * iJ[t] * s[t] + η * k_{-4} * siJ[t], (* + β * k_{JtoP} * JiJ[t] + ϵ * k_{stoA} * siJ[t], *)

P'[t]==k_{JtoP} * iJ[t] + β * k_{JtoP} * JiJ[t] + ψ * k_{JtoP} * siJ[t], (* + k_{JtoP} * Ji[t] + β * k_{JtoP} * JiJ[t] + ψ * k_{JtoP} * Jis[t] *)

s'[t]==-k_3 * i[t] * s[t] + $k_{-3}$is[t] - γ * k_3 * si[t] * s[t] + γ * k_{-3}sis[t] - η * k_3 * Ji[t] * s[t] + η * k_{-3} * Jis[t] - k_4 * i[t] * s[t] + k_{-4}si[t] - γ * k_4 * is[t] * s[t] + γ * k_{-4}sis[t] - η * k_4 * iJ[t] * s[t] + η * k_{-4}siJ[t],

is'[t]==k_3 * s[t] * i[t] - k_{-3} * is[t] - k_{stoA} * is[t] + k_4 * s[t] * i[t] - k_{-4} * si[t] - γ * k_4 * is[t] * s[t] + γ * k_{-4} * sis[t] - θ * k_2 * is[t] * J[t] + θ * k_{-2} * Jis[t], (* + ϕ * k_{stoA} * sis[t] - k_{stoA} * si[t] + ψ * k_{JtoP} * Jis[t], *)

A'[t]==k_{stoA} * is[t] + ϕ * k_{stoA} * sis[t] + ϵ * k_{stoA} * Jis[t], (* + k_{stoA} * si[t] + ϕ * k_{stoA} * sis[t] + ϵ * k_{stoA} * siJ[t] *)

Ji'[t]==-α * k_1 * Ji[t] * J[t] + α * k_{-1} * JiJ[t] + β * k_{JtoP} * JiJ[t] - η * k_3 * Ji[t] * s[t] + η * k_{-3} * Jis[t] + ϵ * k_{stoA} * Jis[t] + k_2 * J[t] * i[t] - k_{-2} * Ji[t], (* - k_{JtoP} * Ji[t], *)

JiJ'[t]==α * k_1 * J[t] * Ji[t] - α * k_{-1} * JiJ[t] - β * k_{JtoP} * JiJ[t] + α * k_2 * J[t] * iJ[t] - α * k_{-2} * JiJ[t], (* - β * k_{JtoP} * JiJ[t], *)

si'[t]==-γ * k_3 * si[t] * s[t] + γ * k_{-3} * sis[t] + ϕ * k_{stoA} * sis[t] - θ * k_1 * si[t] * J[t] + θ * k_{-1} * siJ[t] + ψ * k_{JtoP} * siJ[t],

sis'[t]==γ * k_3 * s[t] * si[t] - γ * k_{-3} * sis[t] - ϕ * k_{stoA} * sis[t] + γ * k_4 * s[t] * is[t] - γ * k_{-4} * sis[t], (* - ϕ * k_{stoA} * sis[t], *)

siJ'[t]==θ * k_1 * J[t] * si[t] - θ * k_{-1} * siJ[t] - ψ * k_{JtoP} * siJ[t] + η * k_4 * s[t] * iJ[t] - η * k_{-4} * siJ[t], (* - ϵ * k_{stoA} * siJ[t], *)

Jis'[t]==η * k_3 * s[t] * Ji[t] - η * k_{-3} * Jis[t] - ϵ * k_{stoA} * Jis[t] + θ * k_2 * J[t] * is[t] - θ * k_{-2} * Jis[t], (* - ψ * k_{JtoP} * Jis[t], *)

i[0]==i0,

J[0]==J0,

iJ[0]==iJ0,

P[0]==P0,

s[0]==s0,

is[0]==is0,

A[0]==A0,

Ji[0]==Ji0,

JiJ[0]==JiJ0,

si[0]==si0,

sis[0]==sis0,

siJ[0]==siJ0,

Jis[0]==Jis0,

i[t],J[t],iJ[t],P[t],s[t],is[t],A[t],Ji[t],JiJ[t],si[t],sis[t],siJ[t],Jis[t], t,0,1200]

PlotTime=25;(*How far out to plot your solution*)

Plot[Evaluate[s[t],A[t],J[t],P[t]/.solution],t,0,PlotTime,

PlotRange → All,

(*PlotRange → 0,50,0,10,*)

Frame →True,

FrameStyle →Thick,

(*FrameTicks → 0,1,2,3,4,5,6,0,0.2,0.4,0.6,0.8,1.0,,,*)

FrameLabel → "Concentration"," ","Time","R303C Kinetic Parameters",

LabelStyle → FontSize → 32,

PlotStyle → RGBColor[0,0,0], Thickness[0.005], RGBColor[1,0,0],

Thickness[0.005], RGBColor[0,1,0], Thickness[0.005], RGBColor[0,0,1], Thickness[0.005]]

A.2.9 Purine Biosynthesis Pathway

This is another model I did not have time to develop further than just starting. This model predicts the intermediate accumulation in the *de novo* purine biosynthesis pathway using the principle of mass action.

(*Package for Legend*)

needs[PlotLegends‘]

(*Initial concentrations for enzymes*)

T0=0.1;(*PPAT*)

G0=0.1;(*GART*)

B0=0.1;(*FGAMS*)

d0=0.1;(*PAICS*)

i0=0.1; (*ADSL*)

F0=0.1; (*ATIC*)

H0=0.1; (*ADSS2*)

(*Initial concentrations for metabolites*)

p0=10;(*PRPP*)

w0=0;(*PRA*)

y0=0;(*GAR*)

z0=0;(*FGAR*)

g0=0;(*FGAM*)

a0=0;(*AIR*)

c0=0;(*CAIR*)

s0=0; (*SAICAR*)

A0=0; (*AICAR*)

r0=0;(*FAICAR*)

M0=0;(*IMP*)

J0=0;(*SAMP*)

P0=0;(*AMP*)

(*Initial concentrations for complexes*)

Tp0=0;

Gw0=0;

Gy0=0;

Bz0=0;

Gg0=0;

da0=0;

dc0=0;

is0=0;

FA0=0;

Fr0=0;

HM0=0;

iJ0=0;

(*rate constants*)

k_{12}=1;

k_{-12}=0;

k_{11}=1;

k_{-11}=0;

k_{10}=1;

k_{-10}=0;

k_{9}=1;

k_{-9}=0;

k_{8}=1;

k_{-8}=0;

k_{7}=1;

k_{-7}=0;

k_{6}=1;

k_{-6}=0;

k_1=1;

k_{-1}=0;

k_2=1;

k_{-2}=0;

k_3=1;

k_{-3}=0;

k_4=1;

k_{-4}=0;

k_5=1;

k_{-5}=0;

(*catalytic rate constants*)

k_{ptow}=10;

k_{wtoy}=10;

k_{ytoz}=10;

k_{ztog}=10;

k_{gtoa}=10;

k_{atoc}=10;

k_{ctos}=10;

k_{StoA}=7.4*0.43;

k_{Ator}=10;

k_{rtoM}=10;

k_{MtoJ}=10;

k_{JtoP}=13.4*0.052;

(*Time constants*)

NDSolveTime=2000;(*How far out to solve the differential equations*)

PlotTime=300;(*How far out to plot your solution*)

solution=NDSolve[

T'[t]== - k_{12} * T[t] * p[t] + k_{-12} * Tp[t] + k_{ptow} * Tp[t],

p'[t]== - k_{12} * T[t] * p[t] + k_{-12}Tp[t],

Tp'[t]==k_{12} * p[t] * T[t] - k_{-12} * Tp[t] - k_{ptow} * Tp[t],

G'[t]== - k_8 * G[t] * g[t] + k_{-8} * Gg[t] + k_{gtoa} * Gg[t] - k_{10} * G[t] * y[t] + k_{-10} * Gy[t] + k_{ytoz} * Gy[t] - k_{11} * G[t] * w[t] + k_{-11} * Gw[t] + k_{wtoy} * Gw[t],

w'[t]== - k_{11} * G[t] * w[t] + k_{-11}Gw[t] + k_{ptow} * Tp[t],

Gw'[t]==k_{11} * w[t] * G[t] - k_{-11} * Gw[t] - k_{wtoy} * Gw[t],

y'[t]== - k_{10} * G[t] * y[t] + k_{-10}Gy[t] + k_{wtoy} * Gw[t],

Gy'[t]==k_{10} * y[t] * G[t] - k_{-10} * Gy[t] - k_{ytoz} * Gy[t],

B'[t]== - k_9 * B[t] * z[t] + k_{-9} * Bz[t] + k_{ztog} * Bz[t],

z'[t]== - k_9 * B[t] * z[t] + k_{-9}Bz[t] + k_{ytoz} * Gy[t],

Bz'[t]==k_9 * z[t] * B[t] - k_{-9} * Bz[t] - k_{ztog} * Bz[t],

g'[t]== - k_8 * G[t] * g[t] + k_{-8}Gg[t] + k_{ztog} * Bz[t],

Gg'[t]==k_8 * g[t] * G[t] - k_{-8} * Gg[t] - k_{gtoa} * Gg[t],

d'[t]== - k_6 * d[t] * c[t] + k_{-6} * dc[t] + k_{ctos} * dc[t] - k_7 * d[t] * a[t] + k_{-7} * da[t] + k_{atoc} * da[t],

a'[t]== - k_7 * d[t] * a[t] + k_{-7}da[t] + k_{gtoa} * Gg[t],

da'[t]==k_7 * a[t] * d[t] - k_{-7} * da[t] - k_{atoc} * da[t],

c'[t]== - k_6 * d[t] * c[t] + k_{-6}dc[t] + k_{atoc} * da[t],

dc'[t]==k_6 * c[t] * d[t] - k_{-6} * dc[t] - k_{ctos} * dc[t],

s'[t]== - k_1 * i[t] * s[t] + $k_{-1}$is[t] + k_{ctos} * dc[t],

i'[t]== - k_1 * i[t] * s[t] + k_{-1} * is[t] + k_{StoA} * is[t] - k_5 * i[t] * J[t] + k_{-5} * iJ[t] + k_{JtoP} * iJ[t],

is'[t]==k_1 * s[t] * i[t] - k_{-1} * is[t] - k_{StoA} * is[t],

A'[t]==k_{StoA} * is[t] - k_2 * F[t] * A[t] + k_{-2}FA[t],

F'[t]== - k_2 * F[t] * A[t] + k_{-2} * FA[t] + k_{Ator} * FA[t] - k_3 * F[t] * r[t] + k_{-3} * Fr[t] + k_{rtoM} * Fr[t],

FA'[t]==k_2 * A[t] * F[t] - k_{-2} * FA[t] - k_{Ator} * FA[t],

r'[t]==k_{Ator} * FA[t] - k_3 * F[t] * r[t] + k_{-3}Fr[t],

Fr'[t]==k_3 * r[t] * F[t] - k_{-3} * Fr[t] - k_{rtoM} * Fr[t],

M'[t]==k_{rtoM} * Fr[t] - k_4 * H[t] * M[t] + k_{-4}HM[t],

H'[t]== - k_4 * H[t] * M[t] + k_{-4} * HM[t] + k_{MtoJ} * HM[t],

HM'[t]==k_4 * M[t] * H[t] - k_{-4} * HM[t] - k_{MtoJ} * HM[t],

J'[t]== - k_5 * i[t] * J[t] + k_{-5}iJ[t] + k_{MtoJ} * HM[t], iJ'[t]==k_5 * J[t] * i[t] - k_{-5} * iJ[t] - k_{JtoP} * iJ[t],

P'[t]==k_{JtoP} * iJ[t],

T[0]==T0,

p[0]==p0,

Tp[0]==Tp0,

G[0]==G0,

w[0]==w0,

Gw[0]==Gw0,

y[0]==y0,

Gy[0]==Gy0,

B[0]==B0,

z[0]==z0,

Bz[0]==Bz0,

g[0]==g0,

Gg[0]==Gg0,

d[0]==d0,

a[0]==a0,

da[0]==da0,

c[0]==c0,

dc[0]==dc0,

s[0]==s0,

i[0]==i0,

is[0]==is0,

A[0]==A0,

F[0]==F0,

FA[0]==FA0,

r[0]==r0,

Fr[0]==Fr0,

M[0]==M0,

H[0]==H0,

HM[0]==HM0,

J[0]==J0,

iJ[0]==iJ0,

P[0]==P0,

T[t], p[t], Tp[t], G[t], w[t], Gw[t], y[t], Gy[t], B[t], z[t], Bz[t], g[t], Gg[t], d[t], a[t], da[t], c[t], dc[t], i[t], s[t], is[t], A[t], F[t], FA[t], r[t], Fr[t], M[t], H[t], HM[t], J[t], iJ[t], P[t], t,0,NDSolveTime]

Plot[Evaluate[s[t],A[t],J[t],P[t]/.solution],t,0,PlotTime,

PlotRange → All,

(*PlotRange → 0,50,0,10,*)

Frame → True,

FrameStyle → Thick,

FrameLabel → "Concentration (AU)"," ","Time (AU)","R303C ADSL Metabolite Accumulation",

PlotStyle → RGBColor[0,0,0],,Thickness[0.005], RGBColor[1,0,0], Thickness [0.005], RGBColor[0,1,0], Thickness[0.005],RGBColor[0,0,1], Thickness[0.005], LabelStyle → FontSize → 32]

A.2.10 Modeling ADSL Tetramer Formation

This is as far as I went with my code before I found the Powers paper on tetramer formation [65].

m0=10;

d0=0;

r0=0;

e0=0;

```
k1=1;
k-1=0;
k2=1;
k-2=0;
k3=1;
k-3=0;
k4=1;
k-4=0;
k5=0;
k-5=0;
k6=0;
k-6=0;

solution=NDSolve[
m'[t]==-k1*m[t]*m[t]-k2d[t]*m[t]-k3r[t]*m[t]+k-1d[t]*m[t]+
k-2r[t]*m[t]-k6*m[t]*m[t]*m[t]*m[t],
d'[t]==k1*m[t]*m[t]-k-1d[t]*m[t]-k2d[t]*m[t]-k4*d[t]*d[t],
r'[t]==k2d[t]*m[t]-k3r[t]*m[t]-k-2r[t]*m[t],
e'[t]==k3r[t]*m[t]+k4*d[t]*d[t]+k6*m[t]*m[t]*m[t]*m[t],
m[0]==m0,
d[0]==d0,
r[0]==r0,
e[0]==e0,
m[t],d[t],r[t],e[t],
t,0,100];
```

Plot[Evaluate[m[t],d[t],r[t],e[t]/.solution],t,0,3,

PlotRange→ All,

AxesLabel → "Time (sec)","Concentration (M)",

PlotStyle → RGBColor[0,0,0], Thickness[0.005], RGBColor[1,0,0],

Thickness[0.005], RGBColor[0,1,0], Thickness[0.005], RGBColor[0,0,1], Thickness[0.005],

LabelStyle → FontSize → 22]

A.2.11 Thermal Denaturation

Fraction Folded by Circular Dichroism

Thermal denaturation was first tested on a Jasco-J815 spectropolarimeter by moniroting the absorbance at 222 nm as temperature increased from 10 to 80 °C at a ramping speed of 1 °C/min. Samples were measured at 0.2 mg/mL in a 0.1 cm pathlength quartz cuvette. The mean molar residue ellipticity $[\theta]$ (deg cm^2 dmol^{-1}) was calculated from the equation $[\theta]=\frac{\theta}{10*n*C*\ell}$, where θ is the measured ellipticity in millidegrees, C is the molar concentration of enzyme subunits, ℓ is the path length in centimeters, and n is the number of residues per subunit (503 residues per monomer, including the His6 tag and thrombin cleavage site). The WT and mutant enzyme samples were incubated for ∼2 hours at 25°C before measurements were taken. The unfolding data obtained from CD measurements were treated by linear extrapolation of the pre- post transitions states and converted to the form of fraction folded using the equation

$$F_f = \frac{I_{exp} - I_u}{I_f - I_u} \tag{A.2.16}$$

where I_u and I_f are the baseline intensities of the unfolded and folded signal respectively and I_{exp} is the signal from the protein.

Determining ΔG, ΔH, and ΔC_p by Differential Scanning Calorimetry

Differential Scanning Calorimetry (DSC) scans for measuring the change in excess heat capacity as a function of temperature were performed on a VP-DSC by MicroCal located at the Biophysics Core facility at Anschutz Medical Campus. DSC scans were carried out with 1 mg/mL protein, 1 mL total volume at a scan rate of 1°C/min. The temperature range for thermal denaturation was 15-90°C. Data were analyzed with a two state, one peak model using the ORIGIN DSC software provided by MicroCal. Testing for reversibility was not performed.

A collaboration was attempted with Colette Quinn, PhD, located at TA Instruments. Samples were prepared and shipped to her to measure the excess heat capacity on their Nano DSC. The capillary cell design on the Nano DSC was thought to provide better baseline, more accurate data, and better potential for reversibility. One eppendorf tube of 0.5 mg/ mL adenylosuccinate lyase (ADSL) and one tube of 15 mL of buffer, were received on ice an immediately stored at -20 °C. The buffer and protein were thawed previous to their use at 10 C in a controlled block heat and cooler. Initially, 600 μL of

buffer was loaded into the sample and reference cell. The buffer was scanned several times, with the second scan serving as the background as the first scan is a considered a conditioning scan, which is standard practice. The buffer was removed from the sample cell and 600 μL of ADSL was loaded into the cell.

All of the data was fit with the NanoAnalyze Software package. The concentration used in the analysis of was 0.5 mg/mL. The enthalpy and scaling factor (Aw) are dependent on both the molecular weight and the concentration of the samples. The scaling factor is an integer value that is used to better fit the data when the concentration or molecular weight might be slightly inaccurate. It is also necessary when fitting multiple unfolding events simultaneously as in this case.

A.3 Thermal Denaturation Modeling

The math of thermal denaturation is as follows:

$$\Delta G(T) = \Delta H(T_R) + \int_{T_R}^{T} \Delta C_p dT - T\Delta S(T_R) - T\int_{T_R}^{T} \Delta C_p dln\, T \qquad \text{(A.3.1)}$$

If ΔC_p is temperature independent;

$$\Delta G(T) = \Delta H(T_R) + \Delta C_p(T - T_R) - T\Delta S(T_R) - T\Delta C_p ln\frac{T}{T_R} \qquad \text{(A.3.2)}$$

Simplifying and setting T_R is equal to the melting temperature, T_m;

$$\Delta G(T) = \Delta H(T_m) - T\Delta S(T_m) + \Delta C_p(T - T_m - Tln[\frac{T}{T_m}]) \qquad \text{(A.3.3)}$$

When $T = T_m$, $\Delta G = 0$. This allows $\Delta S(T_m)$ to be solved in terms of $\Delta H(T_m)$ and T_m

$$0 = \Delta H(T_m) - T_m\Delta S(T_m) + \Delta C_p(T_m - T_m - Tln[\frac{T_m}{T_m}]) \qquad \text{(A.3.4)}$$

$$\Delta S(T_m) = \frac{\Delta H(T_m)}{T_m} \qquad \text{(A.3.5)}$$

Substituting back into Equation A.3.3

$$\Delta G(T) = \Delta H(1 - \frac{T}{T_m}) + \Delta C_p(T - T_m - Tln[\frac{T}{T_m}]) \qquad \text{(A.3.6)}$$

Theoretical Values Based on Chain Length

These equations use a different reference temperature so it is not an "apples to apples" comparison. But it is interesting to see how close they may be and can give us a general idea of what to expect for thermodynamic parameters.

Table A.1: Thermodynamic Parameters from DSC and Predictions

Thermodynamic Parameter	Experimental Value	Tetramer Prediction	Monomer Prediction
ΔC_p (kJ/K - mol)	30	99	25
ΔH_1 (kJ/mol)	535	6500	1710
ΔH_2 (kJ/mol)	1157	6500	1710
ΔH_{total} (kJ/mol)	1692	6500	1710
ΔS (J/K - mol)	-	21	5.6

Comparison of experimentally measured thermodynamic parameters from Differential Scanning Calorimetry and predicted thermodynamic values based on chain length

Theoretical predictions of ΔH, ΔS, and ΔC_p based on chain length:

$$\Delta H(373.5) = (3.3N + 112)kJ/mol \tag{A.3.7a}$$

$$\Delta S(385) = (10.9N + 291)J/K - mol \tag{A.3.7b}$$

$$\Delta C_p = (0.051N + 0.26)kJ/K - mol \tag{A.3.7c}$$

A.3.1 Chemical Denaturation

Unfolding and Refolding of ADSL by Guanidine Hydrochloride

For the unfolding of ADSL, a stock solution was prepared of 1 mg/mL protein in ESB with 0 M denaturant. The stock solution was then diluted with ESB with denaturant concentrations, ranging from 0 to 6 M in 0.25 M increments, for a final concentration of 0.2 mg/mL (0.9 μM). For the refolding of ADSL, a stock solution was prepared of 1 mg/mL protein that was dialyzed in 1 L of ESB with 6 M denaturant. The stock solution was then diluted to

0.2 mg/mL with ESB with lower concentrations of denaturant. Samples were allowed to equilibrate overnight before measurements were taken.

Fraction Folded by Circular Dichroism

To monitor the fraction folded, circular dichroism (CD) was used to monitor α-helical secondary structure of ADSL. CD measurements were performed on a Jasco J-815 CD spectrometer equipped with a Peltier temperature control device that maintained the sample at 25°C. Changes in conformational secondary structure of ADSL were monitored in the region between 210 and 230 nm of a 200 μL of the sample loaded in a 1 mm pathlength quartz cuvette. The scanning speed, bandwidth, and data pitch were set to default values. Six scans were taken and averaged for complete spectra. Samples containing no ADSL protein were scanned identically and subtracted. Absorbance at 222 nm was recorded in mdeg and changed to ME by equation $[\theta]=\frac{\theta}{10*n*C*\ell}$for each concentration of denaturant. Baselines for completely folded and unfolded were characterized by a linear trendline. Fraction folded curves were then produced by

$$F_f = \frac{I_{exp} - I_u}{I_f - I_u} \tag{A.3.8}$$

where I_u and I_f are the baseline intensities of the unfolded and folded signal respectively and I_{exp} is the signal from the protein.

Determining Protein State by Sedimentation Velocity

Analytical ultracentrifugation sedimentation velocity (AUC-SV) experiments were conducted in a Beckman Coulter XL-I equipped with an An-Ti 50 8-place titanium rotor. Cells were filled with 440 μL of prepared samples were placed in the rotor and equilibrated at 25 °C for approximately 30 min before data acquisition. Data were recorded at 20 krpm and 25 °C using absorbance optics at 280 nm scanning was performed between the radial positions of 5.8 and 7.2 cm until the sample was completely sedimented. AUC-SV runs were performed singly.

The data were analyzed using the c(s) method of SEDFIT. For the analysis, the meniscus and bottom were allowed to float with a fixed frictional ratio. The meniscus was allow to float in a location in the vicinity of the maximum absorbance spike. Then, the bottom and meniscus were fixed with the frictional ratio allowed to float. The sedimentation coefficient range was 1-15 or 1-30 S so no partial peaks were present at the edge. A resolution was chosen to correspond to a step size of 0.1 S. Initial fixed frictional ratios ranged from 1.2 to 2. A regularization of 0.68 was used. The buffer density and viscosity changed with increasing denaturant concentration and ranged from 1.01648 to 1.13316 g/ml and 0.0091562 to 0.012564 cP respectively. The partial specific volume was changed to 0.7366 cm^3/g. Buffer density and viscosity as well as the partial specific volume were calculated using SEDNTERP. Goodness of fit was determined by the rmsd value, as well as no visible diagonal lines were present on the residuals bitmap. To estimate the relative amount of different states present in the samples, integration of the c(s) distributions was per-

formed. The percentage of ADSL monomer, tetramer, and aggregation were calculated by dividing the corresponding peak area by the sum of all peak area

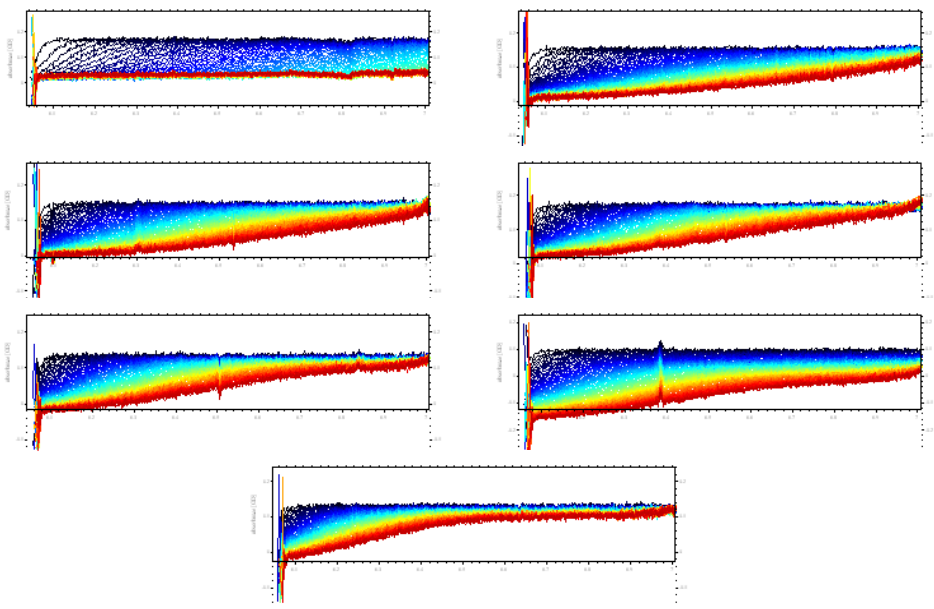


Figure A.1: Sedimentation Velocity graphs from analytical ultracentrifugation

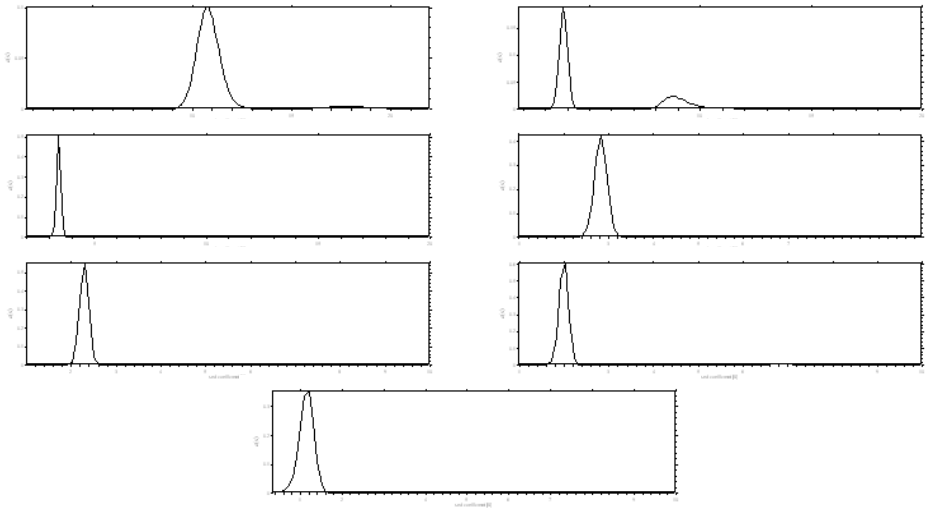


Figure A.2: Sedimentation Velocity graphs from analytical ultracentrifugation

A.3.2 Stability Modeling

Stability models were built in Mathematica.

Two State Model - No Dissociation

```
Clear[ChemMelt, model, bestfit, G, m, t];
SetDirectory[ "/Users/Stephen/Documents/research/DU/
ADSLProject/Stability/ChemicalDenature"];
ChemMelt = Take[Import["WT25ChemDenature.csv"]];
t = 298.15; (*Temperature in kelvin*)
model = (1/(1 + e^(G−m∗x)));
bestfit = NonlinearModelFit[ChemMelt, model, G, m, x];
bestfit["ParameterTable", "RSquared"]


G = 3.4;
m = 1.6;
fit = (1/(1 + e^(G−m∗x)));
Show[ Plot[fit, x, 0, 10, PlotStyle → Black, Thickness[0.005]],
ListPlot[ChemMelt, PlotStyle → PointSize[0.012], Red],
Frame → True,
FrameStyle → Thick,
FrameTicks → 0, 1, 2, 3, 4, 5, 6, 0, 0.2, 0.4, 0.6, 0.8,1.0, , ,
FrameLabel → "Fraction Folded", " ", "GdnHCl (M)", "WT M4 → 4U",
LabelStyle → FontSize → 32,
PlotRange → 0, 6, 0, 1,
AxesOrigin → Automatic]
```

Two State Model - Folded Tetramer to Unfolded Monomer

For models that require an iterative process, I was unable to figure out how to use an iterative process without equal step size in the data, which is the case for the raw data. Therefore, the data file WTRaw.csv, has extrapolated points to fill in the spots that do not have data at certain concentrations.

```
SetDirectory[ "/Users/Stephen/Documents/research/DU/
ADSLProject/Stability/ChemicalDenature"];

(*import the raw data*)
ChemMelt = Take[Import["WTRaw.csv"]];

(*min and max values for ΔF and m value fit parameters*)

F1min = 0;
F1max = 100;

M1min = 0;
M1max = 10;

(*step size between the min and max values*)
F1incriment = 1;

M1incriment = 1;
```

```
(*Creates an array of range of the fit parameters*)

F1 = Range[F1min, F1max, F1incriment];

M1 = Range[M1min, M1max, M1incriment];

(*calculates the number of iterations in the For loops*)

F1stop = (F1max - F1min)/F1incriment;

M1stop = (M1max - M1min)/M1incriment;

(*monomer concentration calculation*)
Subunits = 4;(*number of
monomers in the native structure*)
MonomerMolecularWeight = 228000/Subunits;(*native molecular weight in
Daltons*)
mgmL = 0.2;(*concentration of protein in mg/ml*)
Concentration = mgmL/MonomerMolecularWeight;(*sample concentration in
M
(*Concentration=1.3*10^-6;(*If protein concentration was already in Molarity,
comment out the previous and just use this*)*)
(*min, max, and step size for denaturant concentration*)
dmin = 0;
dmax = 5.2;
```

dincriment = 0.2;

dstop = (dmax - dmin)/dincriment + 1;

(*starting value for if statement to keep track of variance values*)

o = 0;
Γ = 20;

For[i = 0, i ≤ F1stop,
For[j = 0, j ≤ M1stop,

Do[If[o ≤ dstop - 1, o++, o = 1],
t = NSolve[Concentration == 4*mon^4*E$^{(F1[[i]]-M1[[j]]*d)}$ + mon && mon > 0,
mon, Reals],
p = (4*(mon /. t)^4*E$^{(F1[[i]]-M1[[j]]*d)}$)/(Concentration),
Subscript[v, o] = (ChemMelt[[o, 2]] - p)^2 ,
d, dmin, dmax, dincriment],
χ = Sum[Subscript[v, q], q, 1, dstop],
If[χ[[1]] ¡ Γ, Γ = χ[[1]], Print[Γ, F1[[i]], M1[[j]], "best"], Γ]
, j++],
i++]
TimeUsed[]

Three State Model - Folded Tetramer to Folded Monomer to Unfolded Monomer

```
SetDirectory["/Users/Stephen/Documents/research/DU/
ADSLProject/Stability/ChemicalDenature"];

(*import the raw data*)
ChemMelt = Take[Import["WTRaw.csv"]];

(*min and max values for ΔF and m value fit parameters*)

F1min = 45;
F1max = 60;

M1min = 7;
M1max = 17;

F2min = 3.1;
F2max = 3.1;

M2min = 1.5;
M2max = 1.5;

(*step size between the min and max values*)
F1incriment = 0.1;
```

```
M1incriment = 0.1;

F2incriment = 0.1;

M2incriment = 0.1;

(*Creates an array of range of the fit parameters*)

F1 = Range[F1min, F1max, F1incriment];

M1 = Range[M1min, M1max, M1incriment];

F2 = Range[F2min, F2max, F2incriment];

M2 = Range[M2min, M2max, M2incriment];

(*calculates the number of iterations in the For loops*)

F1stop = (F1max - F1min)/F1incriment;

M1stop = (M1max - M1min)/M1incriment;

F2stop = (F2max - F2min)/F2incriment;
```

M2stop = (M2max - M2min)/M2incriment;

(*monomer concentration calculation*)
Subunits = 4;(*number of monomers in the native structure*)
MonomerMolecularWeight =228000/Subunits;(*native molecular weight in Daltons*)
mgmL = 0.2;(*concentration of protein in mg/ml*)
Concentration = mgmL/MonomerMolecularWeight;(*sample concentration in M*)

(*Concentration=1.3*10^{-6};(*If protein concentration was already in Molarity, comment out the previous and just use this*)*)

(*min, max, and step size for denaturant concentration*)
dmin = 0;
dmax = 5.2;
dincriment = 0.2;

dstop = (dmax - dmin)/dincriment + 1;

(*starting value for if statement to keep track of variance values*)

o = 0;
Γ = 20;

For[i = 0, i ≤ F1stop,

For[j = 0, j ≤ M1stop,

For[k = 0, k ≤ F2stop,

For[l = 0, l ≤ M2stop,

Do[If[o ≤ dstop - 1, o++, o = 1],

t = NSolve[Concentration == (4*mon^4*E$^{(F1[[i]]-M1[[j]]*d)}$) + mon

+ (mon*E$^{-(F2[[k]]-M2[[l]]*d)}$) && mon > 0, mon, Reals],

p = (4*(mon /. t)^4*E$^{(F1[[i]]-M1[[j]]*d)}$ + (mon /. t))/(Concentration),

Subscript[v, o] = (ChemMelt[[o, 2]] - p)^2,

d, dmin, dmax, dincriment],(*end Do loop*)

χ =Sum[Subscript[v, q], q, 1, dstop],

If[χ[[1]] ¡ Γ, Γ = χ[[1]],

Print[Γ, F1[[i]], M1[[j]], F2[[k]],

M2[[l]]], Γ]

, l++](*end M2 (l) For loop*),

k++](*end F2 (k) For loop*)

, j++],(*end M1 (j) For loop*)

i++](*end F1 (i) For loop*)

TimeUsed[]

Three State Best Fit

After running the brute force code, I would use this one to see what the fit looked like using the actual data file.

```
SetDirectory["/Users/Stephen/Documents/research/DU/
ADSLProject/Stability/ChemicalDenature"];

ChemMelt = Take[Import["WTRaw1.csv"]];

(*monomer concentration calculation*)

Subunits = 4;(*number of monomers in the native structure*)

MonomerMolecularWeight = 228000/ Subunits;(*native molecular weight in
Daltons*)

mgmL = 0.2;(*concentration of protein in mg/ml*)

Concentration = mgmL/MonomerMolecularWeight;(*sample concentration in
M*)

F1 = 65;
M1 = 23;
F2 = 3.1;
M2 = 1.5;

dmin = 0;
dmax = 6;
```

dincriment = 0.05;

dstop = (dmax - dmin)/dincriment + 1;

o = 0;

Do[If[o ≤ dstop - 1, o++, o = 1],
t = NSolve[Concentration == 4*mon^{4}*E$^{(F1-M1*d)}$ + mon + mon*E$^{-(F2-M2*d)}$
&& mon > 0, mon, Reals],
p = (4*(mon /. t)^{4}*E$^{(F1-M1*d)}$ + (mon /. t))/(Concentration),
Subscript[a, o] = Flatten[d, p],
d, dmin, dmax, dincriment]

Show[ListPlot[ChemMelt, PlotStyle → PointSize[0.012], Red],
ListLinePlot[Table[Subscript[a, z], z, dstop],
PlotStyle → Black, Thickness[0.005]],
Frame → True,
FrameStyle → Thick,
FrameTicks → 0, 1, 2, 3, 4, 5, 6, 0, 0.2, 0.4, 0.6, 0.8, 1.0, , ,
FrameLabel → "Fraction Folded", " ", "GdnHCl (M)", "WT M_4 → 4M → 4U",
LabelStyle → FontSize → 32,
PlotRange → 0, 6, 0, 1,
AxesOrigin → Automatic]

Prediction Fraction of Tetramer and Monomer for Three State Model

(*monomer concentration calculation*)

Subunits = 4;(*number of monomers in the native structure*)

MonomerMolecularWeight = 228000/Subunits;(*native molecular weight in Daltons*)

mgmL = 0.2;(*concentration of protein in mg/ml*)

Concentration = mgmL/MonomerMolecularWeight;(*sample concentration in M*)

F1 = 50.8;
M1 = 12;
F2 = 3.1;
M2 = 1.5;
d = 0;

t = NSolve[Concentration == 4*mon^{4}*E$^{(F1-M1*d)}$ + mon + mon*E$^{-(F2-M2*d)}$ && mon > 0, mon, Reals];
tetfrac = (4*(mon /. t)^{4}*E$^{(F1-M1*d)}$)/(Concentration) monfrac = ((mon /. t) + ((mon /. t)*E$^{-(F2-M2*d)}$))/(Concentration)

Hydrodynamic Radius Prediction for Three State Model

r = 2.8; (*Wilkins hyd. radius for folded*)

R = 6.02; (*Experimental hyd. radius for unfolded*)

SetDirectory["/Users/Stephen/Documents/research/DU/
ADSLProject/Stability/ChemicalDenature"];

Radius = Take[Import["HydRadiusAUC.csv"]];

(*monomer concentration calculation*)

Subunits = 4;(*number of monomers in the native structure*)

MonomerMolecularWeight = 228000/ Subunits;(*native molecular weight in Daltons*)

mgmL = 0.2;(*concentration of protein in mg/ml*)

Concentration = mgmL/MonomerMolecularWeight;(*sample concentration in M*)

F1 = 50.8;
M1 = 12;
F2 = 3.1;

M2 = 1.5;

dmin = 0.9;

dmax = 6;

dincriment = 0.05;

dstop = (dmax - dmin)/dincriment + 1;

o = 0;

Do[If[o ≤ dstop - 1, o++, o = 1],

t = NSolve[Concentration == 4*mon^{4}*E$^{(F1-M1*d)}$ + mon + mon*E$^{-(F2-M2*d)}$

&& mon > 0, mon, Reals],

p = ((mon /. t))/((mon /. t) + (mon /. t)*E$^{-(F2-M2*d)}$),

P = ((mon /. t)* E$^{-(F2-M2*d)}$)/((mon /. t) + (mon /. t)*E$^{-(F2-M2*d)}$),

Subscript[R, avg] = ((r*p) + (R*P))/(P + p),

Subscript[a, o] = Flatten[d, Subscript[R, avg]] ,

d, dmin, dmax, dincriment]

Show[

ListLinePlot[Table[Subscript[a, z], z, dstop],

PlotStyle → Black, Thickness[0.005]],

ListPlot[Radius, PlotStyle → PointSize[0.02], Red],

Frame → True,

FrameStyle → Thick,

FrameTicks → 0, 1, 2, 3, 4, 5, 6, 0, 2, 3, 4, 5, 6, 7, , ,

FrameLabel → "R_h (nm)", " ", "GdnHCl (M)", "WT M_4 → 4M → 4U",

LabelStyle → FontSize → 32,

PlotRange → 1, 5, 2, 6.5,

AxesOrigin → Automatic]

A.3.3 Isothermal Titration Calorimetry

ITC experiments were conducted using a NANO-ITC system (TA Instruments, Utah, USA). Thrombin cleaved WT ADSL was dialyzed overnight against a solution of 25 mM HEPES pH 7.0, 150 mM KCl, and 2 mM DTT. AMP was brought to a concentration of 2.5 mM using the buffer in which WT ADSL was dialyzed. ITC runs of WT ADSL with AMP were performed in duplicate and comprised of one 1 μl injection followed by 24, 2 μl injections for a total of 25 injections of 2.5 mM AMP into 0.250 mM WT ADSL. Each injection was spaced 250 seconds apart. ITC experiments were performed in an identical method, injecting AICAR instead of AMP into 0.225 mM WT ADSL. In addition, AMP and AICAR were titrated in an equal manner into 0.280 mM R303C ADSL. Data sets were analyzed with NanoAnalyze software and fit to an independent model concurrently with a blank constant model to adjust for heat of dilution.

Appendix B

Detailed Lab Protocol

B.1 Fabricating and Purifying Proteins

B.1.1 Day to Day Details

Day 1 - Estimated Time = 4 hours

1) Checks amount of buffer solution. If not enough see buffer page on how to make more. Will need 1 L Lysis Buffer, ∼2 L Enzyme Storage Buffer, 1 M Imidazole stock.

2) Make 1.5 L of LB per mutation to grow cultures in, see solutions page for recipe.

3) Autoclave LB to sterilize using L8 cycle. Make sure to cover flasks with Al foil and mark with autoclave tape. Autoclave takes 1.5 hours.

4) Once cycle is finished, load in incubator at 37^{o}C. If incubator is being used, cool LB first so you will not to kill cultures in incubator. Make 5 - 10 mL fresh ^{34}Cm and ^{100}Ap during autoclave or take some from the - 50 C freezer

to thaw. Best to make fresh.

5) Make 2 overnight cultures in 13 mL falcon tube.

Label with date, WT or mutation, initials.

Fill test tube with 5 mL LB and 5 μL of ^{34}Cm and ^{100}Ap. Aloquot and freeze ^{34}Cm and ^{100}Ap to use tomorrow

Find desired ADSL clone in -80^{o}C freezer. See chart below.

Use yellow pipet tip to scoop out ADSL and place in test tube. It is also optional to streak colonies of the mutant a day or two in advance.

Put into shaker @ 37^{o}C for overnight.

6) Check Ni purifying column. If it does not look ok or if purifying a new mutant, proceed to step 7.

7) To make a new column...

Cut off the top of 25 mL pipet, and put glass wool in the pipet to support the resin (just add enough to where the resin won't leak through). Ni-NTA His bind resin from deli fridge. Suspend Ni resin gently and add $\sim$24 mL to the pipet. Let it settle. Run about 50 mL Lysis buffer through column, label with date and mutant it will be used for, and store in deli fridge. To store the column, plug tip and leave 3 mL buffer on top of the resin.

ADSL Type	Location
WT	A-9
R426H	B-8
E80D	C-8
D87E	C-9
A291V	B-7
R303C	B-6
M225T	Not Made

Day 2 - Estimated Time = 6 hours

1) Thaw ^{34}Cm and ^{100}Ap, make sure you have enough for .5 mL per flask.

Check temp of both incubators, they should be 37^{o}C.

2) Label each flask as Control, A, B, C

Add .5 mL ^{34}Cm and ^{100}Ap, then 2 mL of overnight culture for 1:250 dilution to each flask.

3) Put into shaker @ 37^{o}C, 250 rpm, and record time or start timer

4) Measure OD @ $\lambda = 600$ using UV-VIS with 1 mL samples from each flask (this measures the bacteria growth rate).

Use OD print out sheet to record data

5) Make 5 mL fresh IPTG.

6) Once OD is between .4 and .6 (aim for .5), induce 2 of the 3 flasks with 2 mL 100 mM IPTG.

Immediately turn off the shaker incubator and grow overnight

Best to take a reading at 2 and 3 hours to get an idea of how fast the protein is growing.

Day 3 - Estimated Time = 5 hours

1) Get water bath of 37 - 50 ^{o}C ready for Freeze-Thaw cycle. Get out DNase and RNase from freezer to thaw.

2) Take 1 mL sample of each ADSL type and put in labeled eppie.

Will need to eventually spin down in mini-centrifuge in deli fridge, discard supe, and freeze.

3) Get out 2 nalgene bottles per flask to pour remaining culture in.

Use wet pellet table to record empty bottle weights, then pour about 250 mL in each bottle. Make sure the weights are within one gram of each other for centrifugation.

4) Spin remaining culture down @ 4000 rpm in GSA rotor in the Sorvall RC-5 for 10-15 min. to pellet cells.

Colman reported a wet cell pellet of $\sim$ 3.7 g per 2 L cell culture

Decontaminate flasks with 10% bleach.

5) Save 200 μL supe in eppie, store in -50 C freezer and discard rest

Record weight of bottle and pellet to get pellet weight

6) Add 10 times the weight of the pellet of Lysis buffer (mL). Add 1 μg DNase and 10 μg RNase per mL Lysis buffer added. If needed add 45 - 60 KU lysozyme per gram wet pellet. Easy way to do this is to add the total amount of Lysis Buffer to one bottle, resuspend pellet, and pour solution into the next bottle of the same type of ADSL until done. DO NOT MIX DIFFERENT

MUTANTS OR WT if purifying more than one ADSL.

7) Begin Freeze-Thaw cycle of 1 hour in -80^{o}C freezer folowed by 10 - 15 min in 37 - 50^{o}C water bath.

Repeat cycle 3 times.

Leave in -80^{o}C freezer over night. This is a potentially good break point if get busy and can't continue the next few days.

8) During Freeze-Thaw, make SDS-PAGE gels for diagnotic tests.

Expect ADSL monomer band at 55 kDa

See SDS-PAGE gel protocol for making gels.

9) Prep the fraction collector by loading with glass tubes. Number the tubes with permanent marker and put 400 μL (10% total volume of fraction) of ADSL Storage buffer with 10 mM DTT in each test tube, cover with seran wrap and leave on counter overnight.

Day 4 - Estimated Time = 14 - 16 hours. Bring Reading Material

1) Set water bath to 37^{o}C if it is not already there.

Thaw lysate in -80^{o}C freezer in 37^{o}C water bath once it is up to temp.

2) Start washing column with Lysis Buffer

Want to flow $\sim$ 100 mL through before loading ADSL.

3) Once Lysate is thawed, spin for 30 min @ 12,000 (23,500 x g) rpm in GSA rotor in Sorvall RC-5. Make sure to use the "hard" plastic nalgene bottles because the other ones cave in from the high rpms.

4) Make 20 and 250 mM Imidazole solutions of 400 and 160 mL total volume respectively.

5) Retrieve Lysate, pour supe into 50 mL conical tube(s). Weigh the bottles containing cell debris and calculate percent lysis.

May want to filter ADSL supe using 4 x 4 gauze by PAGE Gel solutions.

6) Start loading column with ADSL

Once you start loading the column, IT CANNOT RUN DRY!

Save at least 200 μL of Flow through in 50 mL conical. Probably want to do this sometime after half of ADSL has been loaded on the column.

7) Flow 200 mL Lysis Buffer though the column.

Catch at least 200 μL of flow through in 50 mL conical tube labeled Wash Buffer 1. Catching flow through is optional, but it is always good practice just in case purification does not work so you can trace back to where you "lost" the protein.

8) Flow 200 mL Lysis Buffer with 20 mM Imidazole though the column.

Catch at least 200 μL of flow through in 50 mL conical tube labeled Wash

Buffer 2.

9) Once the last 20 mL is put into the pipet, prepare column gradient.

• stirring plate

• smallest stirring rod

• large column gradient

• Lysis Buffer, Lysis Buffer with 20 mM Imidazole, Lysis Buffer with 250 mM Imidazole

• fraction collector

Make sure the out-take tube is pinched shut and the valve connecting the inner and outer cylinders is open.

Place the smallest stir bar in the inner cylinder and pour 150 mL 20 mM Imidazole into the inner cylinder and 150 mL 250 mM Imidazole into the outer cylinder.

Once there is enough room to start the column gradient, open the out-put tube slowly. Adjust the input into the pipet so it is equals the output. Put on the connecting tube from the pipet to the fraction collector.

10) Once the column gradient is started, put the start the fraction collector and start collecting 4 mL fractions. You will need to time out 4 mL for the fraction collector timing.

11) Speed of output usually increases with time. May need to check on it every half hour to hour. Once fraction 45 or 50 has been collected, equilibrate the flow and you should be ok to leave because unless you started at 6 am, it is going to be late.

Day 5 - Estimated Time = 6 hours

1) Run bulk Bio-Rad without the standard to find peak of ADSL (see concentration protocol).

2) Get out resin column and run 200 mL Lysis buffer with 250 mM Imidazole to elute any proteins left on the Ni.

Then run 200 mL of plain Lysis buffer through the resin and store in deli fridge.

3) Run SDS-Page Gel.

See SDS-PAGE Gel protocol.

4) Stain SDS-Page Gel overnight with Commassie Blue Staining solution.

Day 6 - Estimated Time = 4 hours

1) Destain gel and take photo.

2) If data looks good, take the fractions and start filter/buffer swap sequence

All spins are for 15 min at 4,000 rpm at room temp.

Get 2 out 10,000 kDa Amni-con filters per ADSL type and spin with 3 mL Storage Buffer

Load ADSL fractions 10 mL at a time. This may require more than one spin. Keep retentate, dump filtrate.

Add 10 mL of Storage Buffer and spin. Do this 3 times.

Extract ADSL via pipet into 15 mL conical tube. Rinse membrane with 3 mL Storage Buffer and place in respective conical tube.

If protein is not quite soluble, add about 1 mL Storage Buffer

3) Soak ∼6 in of 6-8 kDa dialysing membrane per conical tube in dH_2O for 30 min.

4) After soaking, load ADSL in membrane and dialyze in 1 L of Storage Buffer overnight.

Day 7 - Estimated Time = 5 hours

1) Put dialyzed ADSL into 15 mL conical tubes. Put 1 mL into eppie to check concentration.

2) Check concentration with E 1% (see concentration protocol).

3) Check linearity (see kinetics protocol).

4) Protein is good for ∼3 weeks on the counter. Best to do all experiments as soon as possible. Freeze leftover for experiments down the line.

5) ALL DONE WITH PROCESS!.

B.1.2 Buffers for Fabricating and Purifying Proteins

ADSL Lysis Buffer

50 mM Potassium Phosphate buffer @ pH 8.0 containing 300 mM KCl and 10% glycerol

Sterile Filter when pouring into bottle

Volume (mL)	1 M K_2HPO_4 (mL)	1 M KH_2PO_4 (mL)	KCL (g)	Glycerol (mL)
500	23.5	1.5	11.2	50
1000	47	3	22.4	100
1500	70.5	4.5	33.6	150
2000	94	6	44.8	200

ADSL Enzyme Storage Buffer

50 mM potassium phospate buffer @ pH7.0 containing 150 mM KCl, 1 mM DTT, 1 mM EDTA, and 10% glycerol

Sterile Filter when pouring into bottle

Volume (mL)	1 M K_2HPO_4 (mL)	1 M KH_2PO_4 (mL)	KCL (g)	Glycerol (mL)	1 M DTT	.5 M EDTA
500	15.4	9.6	5.6	50	.5	1
1000	30.8	19.2	11.2	100	1	2
1500	46.2	28.8	16.8	150	1.5	3
2000	61.6	38.4	22.4	200	2	4

1 M Imidazole in ADSL Lysis Buffer

50 mM Potassium Phosphate buffer containing 300 mM KCl, 20 mM Imidazole and 10% glycerol

4 x stock of 250 mM, 50 x stock of 20 mM.

Sterile Filter when pouring into bottle

Volume (mL)	1 M K_2HPO_4 (mL)	1 M KH_2PO_4 (mL)	KCL (g)	Glycerol (mL)	Imidazole (g)
500	23.5	1.5	11.2	50	34.04
1000	47	3	22.4	100	68.08

20 and 250 mM Imidazole in ADSL buffer from 1 M stock

After dilution pH to 8.0. The total volume is optimal for one purification run.

pH with Phosphoric Acid.

Final Concentration (mM)	Volume (mL)	1 M Imidazole (mL)	ADSL Lysis Buffer (mL)
20	400	8	392
250	160	40	120

20 mM Imidazole in ADSL Lysis Buffer

50 mM Potassium Phosphate buffer @ pH 8.0 containing 300 mM KCl, 20 mM Imidazole and 10% glycerol

Sterile Filter when pouring into bottle

pH with Phosphoric Acid.

Volume (mL)	1 M K_2HPO_4 (mL)	1 M KH_2PO_4 (mL)	KCL (g)	Glycerol (mL)	Imidazole (g)
500	23.5	1.5	11.2	50	.68
1000	47	3	22.4	100	1.36
1500	70.5	4.5	33.6	150	2.04
2000	94	6	44.8	200	2.72

250 mM Imidazole in ADSL Lysis Buffer

50 mM Potassium Phosphate buffer @ pH 8.0 containing 300 mM KCl, 250 mM Imidazole and 10% glycerol

Sterile Filter when pouring into bottle

pH with Phosphoric Acid.

Volume (mL)	1 M K_2HPO_4 (mL)	1 M KH_2PO_4 (mL)	KCL (g)	Glycerol (mL)	Imidazole (g)
500	23.5	1.5	11.2	50	8.51
1000	47	3	22.4	100	17.02
1500	70.5	4.5	33.6	150	25.53
2000	94	6	44.8	200	34.04

B.1.3 Culture Solutions

LB

pH to 7.0

Volume (mL)	yeast (g)	peptone (g)	NaCl (g)
500	2.5	5	5
1000	5	10	10
1500	7.5	15	15

100 mM IPTG - Isopropl β-D-1-thiogalactopyranoside

Stuff that begins the over-expression of protein

IPTG (mg)	dH_2O (mL)
238	10

Antibiotics

^{34}Cm - Chloramphenicol

34 mg/mL solution in 70% ethanol

Store in eppies in -50 C freezer

Cm (mg)	70% Ethanol (mL)
170	5
340	10

[100]Amp - Ampicillin Sodium Salt

100 mg/mL solution in dH_2O

Store in eppies in -50 C freezer

Ap (g)	dH_2O (mL)
.5	5
1	10

B.1.4 Extra Solutions

DNase I

Dilute to 1 mg/mL with ADSL lysis buffer in bottle it came in. Store in -50 C freezer.

RNase A

Dilute to 10 mg/mL with ADSL lysis buffer in bottle it came in. Store in -50 C freezer.

0.5 M EDTA

Tricky to make. Dissolve given amount of EDTA with dH_2O (half total volume) first. Then add about 20 g NaOH per liter solution. Top off with dH_2O and adjust pH, aim for 8.0.

Volume (mL)	EDTA (g)	NaOH (g)
500	93.05	10
1000	186.1	20

1 M DTT - Dithiothreitol

DTT (g)	dH_2O (mL)
1.54	10

B.2 Fabricating and Running SDS-Page Gels

B.2.1 Step by Step Instructions

Wear gloves while making and pouring gels and will need a camera to take a photo of the resulting gel.

1) Compose gel molding device

Need

- large glass slide (check for previous gel residues)
- two spacers
- white divider
- white comb
- two large paper clamps

Make sure the mold is flush with the table!

2) Seal bottom of mold with agarose by pouring a straight line of agarose and

placing the mold on top. Continue to seal with more agarose. Make sure the white divider is on the right side when placing the mold on agarose.

3) Get out 10% Ammonium Persulfate to thaw or make fresh.

Mark glass with permanent marker at the same height as the top of the black paper clamp.

4) Start mixing 12% resolving gel.

Using a transfer pipet, carefully run the resolving gel down one side of the mold. Continue adding until the gel has reached the mark at the top of the paper clamp. You may need to refill once or twice up to the mark.

Once completed, top off with .1% SDS of dH_2O to prevent oxidation. Let gel polymerize for 30 min. Save excess gel in designated flask.

Note: The max amount you want to make at one time is 15 mL for two gels. Otherwise the gel polymerizes before you can pour a third.

5) As the resolving gel cures, start making mixing the 5% stacking gel without TEMED.

6) After 30 min. gently rinse out the gel mold with dH_2O. Will need to use filter paper to soak up the excess water.

7) Place comb partially in one corner of the mold, leaving one side to pour stacking gel.

8) Mix in TEMED, swirl around the solution and top off the gel mold. Smoothly place comb into gel. If you do this carelessly, bubbles will form in the lanes. The stacking gel takes 20 minutes to cure.

9) If do not need until gels until later: Wrap each mold in paper towels soaked in 1 x Tris Glycine, then in seran wrap, and store in deli fridge overnight. On day you need the gels, bring up to room temp before moving on.

10) Get out SDS-Page Gel contraption and partially fill with 1 x Tris Glycine Electrophoresis Buffer.

11) Place gels in the contraption securing it with the red clamps and mark the center of the wells. It is easiest to do this with the comb still in. Then gently take the comb out. Fill in wells with 1 x Tris Glycine Electrophoresis Buffer.

12) Start heating up the incubator and take out protein standards to thaw. Fill eppies with 25 μL of SDS Buffer and 5 μL of the sample. Boil in water bath for 3 min. Afterwards spin them down for 20 sec. May need to add 1 M DTT (about 50 μL to SDS buffer.

13) Load outer wells with 10 μL of denatured standard. Load rest of lanes with 10 μL of sample. Be sure to label in note book what sample goes in what lane.

14) Let gel sit for $\sim$5 min after loading lanes. Then put on cover and attached to voltage supply. Set voltage to 45 Volts. It takes $\sim$50 min for protein to run through the stacking gel.

15) When protein has reached resolving gel, top off buffer in inner cavity if needed. Crank voltage up to 80 Volts. It takes $\sim$1.5 hours to run through the resolving gel. Make sure the protein runs all the way off the gel.

16) Carefully dismantle the gel apparatus, pry open the gel plates with a razor, discard the stacking gel, cut the corner off of the top left and/or bottom left of the resolving gel, remove the gel and place it in commassie blue stain solution. Cover with seran wrap and place on teter-table set @ 3-4 overnight.

Note: Commassie blue staining solution is reusable up to 5 times.

Note: If do not want to wait overnight, stain for minimum of 2 hours.

17) Save what stain you can by pouring it into the beaker and mark that it

has been used. Use a kimwipe to soak up extra stain. Set kimwipes in corners to soak up destainer, put in enough destainer to cover the gel(s), recover with seran wrap and put back on teter-table. Process takes ∼1.5 hours. You do not want to destain to go too far, otherwise, you will not see any bands and will have to restain. Check every 30 min.

18) Set on projector and take a picture.

19) Wrap in seran wrap, label with tape, and store in the deli fridge.

B.2.2 Buffers & Gel Solutions

6% Resolving Gel

	Component Volumes (mL)			
Solution Components	5 mL	10 mL	15 mL	20 mL
dH_2O	2.6	5.3	7.9	10.6
30% Acrylamide Mix	1.0	2.0	3.0	4.0
1.5 M Tris (pH 8.8)	1.3	2.5	3.8	5.0
10% SDS	0.05	0.1	0.15	0.2
10% Ammonium Persulfate	0.05	0.1	0.15	0.2
TEMED	0.004	0.008	0.012	0.016

8% Resolving Gel

	Component Volumes (mL)			
Solution Components	5 mL	10 mL	15 mL	20 mL
dH_2O	2.3	4.6	6.9	9.3
30% Acrylamide Mix	1.3	2.7	4.0	5.3
1.5 M Tris (pH 8.8)	1.3	2.5	3.8	5.0
10% SDS	0.05	0.1	0.15	0.2
10% Ammonium Persulfate	0.05	0.1	0.15	0.2
TEMED	0.003	0.006	0.009	0.012

10% Resolving Gel

	Component Volumes (mL)			
Solution Components	5 mL	10 mL	15 mL	20 mL
dH_2O	1.9	4.0	5.9	7.9
30% Acrylamide Mix	1.7	3.3	5.0	6.7
1.5 M Tris (pH 8.8)	1.3	2.5	3.8	5.0
10% SDS	0.05	0.1	0.15	0.2
10% Ammonium Persulfate	0.05	0.1	0.15	0.2
TEMED	0.002	0.004	0.006	0.008

12% Resolving Gel - used for hADSL Gels

	Component Volumes (mL)			
Solution Components	5 mL	10 mL	15 mL	20 mL
dH_2O	1.6	3.3	4.9	6.6
30% Acrylamide Mix	2.0	4.0	6.0	8.0
1.5 M Tris (pH 8.8)	1.3	2.5	3.8	5.0
10% SDS	0.05	0.1	0.15	0.2
10% Ammonium Persulfate	0.05	0.1	0.15	0.2
TEMED	0.002	0.004	0.006	0.008

15% Resolving Gel

	Component Volumes (mL)			
Solution Components	5 mL	10 mL	15 mL	20 mL
dH_2O	1.1	2.3	3.4	4.6
30% Acrylamide Mix	2.5	5.0	7.5	10.0
1.5 M Tris (pH 8.8)	1.3	2.5	3.8	5.0
10% SDS	0.05	0.1	0.15	0.2
10% Ammonium Persulfate	0.05	0.1	0.15	0.2
TEMED	0.002	0.004	0.006	0.008

5% Stacking Gel

	Component Volumes (mL)			
Solution Components	1 mL	2 mL	3 mL	4 mL
dH_2O	0.68	1.4	2.1	2.7
30% Acrylamide Mix	0.17	0.33	0.5	0.67
1.0 M Tris (pH 6.8)	0.13	0.25	0.38	0.5
10% SDS	0.01	.02	0.03	0.04
10% Ammonium Persulfate	0.01	.02	0.03	0.04
TEMED	0.001	.002	0.003	0.004

5 x Tris Glycine Electrophoresis Buffer

25 mM Tris, 250 mM Glycine, .1% SDS, pH 6.8

Volume (mL)	Tris (g)	mM Glycine (pH 8.3) (g)	10% (w/v) SDS (mL)
500	7.6	47	25
1000	15.1	94	50
1500	22.7	141	75

1 x SDS Gel Loading Buffer

50 mM Tris Cl, 100 mM DTT, 2% SDS, .1% Bromophenol Blue, 10% Glycerol

dH_2O (mL)	10% SDS (mL)	1 M Tris Gly	Glycerol (mL)	Bromophenol Blue
5.5	2	.5	1	1

Sodium Dodecyl Sulfate (SDS)

Mass/Volume	SDS (g)	dH_2O (mL)
10%	1	10
1%	.1	10

Coomassie Brilliant Blue R250

Careful with this, is will permanently stain clothes and turn hands blue for extended period of time

Coomassie Brilliant Blue R250 (g)	MetOH:dH_2O (1:1 v/v) (mL)	Glacial Acetic Acid (mL)
0.25	90	10
0.50	180	20

Destaining Solution

MetOH:dH_2O (1:1 v/v) (mL)	Glacial Acetic Acid (mL)
90	10
180	20

10% Ammonium Persulfate

dH_2O (mL)	Ammonium Persulfate (g)
10	1

Agarose Bed (2% Gel)

Bed for making SDS page gels

dH_2O	Agarose (g)
50	1
100	2

B.3 Enzyme Kinetics

B.3.1 Linearity

Need

- 200 μL and 10 μL pipets and tips

- 200 μL quartz cuvettes
- Range of .1, .2, .3, .4, .5 μg /mL ADSL
- 60 μM S-AMP, best to make fresh solution from 600 μM stock.
- 50 mM HEPES
- notebook and writing utensil

To check the linearity of the protein, run kinetics for 2 to 3 samples of each ADSL dilution (see protocol below). Record change in absorbance over 1 min period and in your lab notebook. Then transcribe the numbers into the ADSL Characterization excel spread sheet for that mutant under the tab "Linearity". You should end up with a linear line. Print out graph and place in binder.

1) Make sure you have enough 60 μM S-AMP for at least 10 runs. Gather all the needed materials and go to room 254 (Margittai and Pegan's Lab or Dr. Verl's lab to use the Thermo UV Spec) and use the Varian UV/Vis Spec.
2) Turn on the Spec.
3) Open up software "kinetics" located in Cary folder on desktop.
4) Open up the method in Patterson folder for bulk runs.
5) Wash cuvettes 20 times prior to using.
6) Blank the spec with 50 mM HEPES. Keep for quality check run at beginning and end of experiment.
8) Start run, save data in Patterson Lab folder, name data set (date, type, linearity, ie: 042510 WT Linearity), press start until 2 min countdown starts.
9) Fill cuvette with 200 μL substrate.
10) Inject 2 μL of ADSL into the cuvette, pump once or twice, shake cuvette

3 - 5 times to mix and place in the UV spec (cell #1 if using the cell changer), make sure you press the cuvette all the way down. Close the chamber and press enter for the reading to start.

11) While the sample is reading, wash any previously used cuvette.

12) Linear regression data should show up. Record the slope in notebook.

13) Remove sample (may have to use cuvette remover if cell changer platform is in spec).

14) The next sample run will pop up on the screen.

15) Repeat steps 9 - 13 until finished.

B.3.2 Running Enzyme Kinetics

- 200 μL and 10 μL pipets and tips
- 200 μL quartz cuvettes
- $\sim$250 μg /mL ADSL
- 1 mL of 1, 2, 3, 4, 6, 7.5, and 10 μM S-AMP and/or SAICAR in 50 mM HEPES
- 50 mM HEPES
- parafilm
- notebook and writing utensil
- USB

To check the kinetic behavior (cooperativity, $K_{0.5}$) of ADSL toward either S-AMP or SAICAR, run 3 samples of each S-AMP or SAICAR dilution (see protocol below). Because the graph will not be linear over one minute, the analysis of V_o is more complicated. Raw data will need to be saved individually

as a .csv file and uploaded on Mathematica "Hill Plot Template". Instructions for uploading data onto the template are written in the mathematica file.

1) Make sure you have enough substrate for at least 3 runs each. Gather all the needed materials and go to room 254 (Margittai and Pegan's Lab or Dr. Verl's lab to use the Thermo UV Spec) and use the Varian UV/Vis Spec.
2) Turn on the Spec. If doing thermal stability, make sure cell changer is in place and turn on temp control.
3) Open up software "kinetics" located in Cary folder on Desktop.
4) Open up the method in Patterson folder for single runs. This automatically saves data in .csv form.
5) Wash cuvettes 20 times prior to running an experiment.
6) Blank the spec with 50 mM HEPES.
8) Start run, save data in Patterson Lab folder, name sample (date, type, substrate, concentration, run ie:122510WTSAMP1i), press start until 2 min countdown starts.
9) Fill cuvette with 200 μL substrate.
10) Inject 2 μL of ADSL into the cuvette, pump once or twice, shake cuvette 3 - 5 times to mix and place in the UV spec (cell #1 if using the cell changer), make sure you press the cuvette all the way down. Close the chamber and press enter for the reading to start.
11) While the sample is reading, wash any previously used cuvette.
12) Remove sample (may have to use cuvette remover if cell changer platform is in spec).
13) Repeat steps 8 - 12 until finished.

14) Save data on USB drive and then start transferring data to mathematica template.

B.3.3 Mathematica Templates

Calculating V_o

(*User input, sets the directory*)

SetDirectory["/Users/Stephen/Documents/research/DU/ADSLProject/ADSLkineticsdata/012512/SAMP"];

(*User input*)

concentration = 0.15;(*concentration of enzyme used in experiment in mg/mL.*)

injvol = .01;(*Volume of injected enzyme for .2 and 1 mL substrate volumes.*)

mg = concentration*injvol;(*Amount of milligrams enzyme used during experiment for .2 and 1 mL substrate volumes.*)

mL = 1;(*Substrate volume used for 60 micromolar concentraion to preserve sample. This produces data identical to 1 mL samples*)

[Epsilon] = 10000;(*Extinction coefficient (ϵ) in M^{-1} cm^{-1}*)

start = 9;(*the cell in the csv file that the data starts in*)

freq = 460; (*number of data points taken in 1 minute - 60 divided by time interval in seconds*)

n = 3; (*Number of trials per concentration*)

1 μM SAMP

Trial 1

Clear[data,datum,a,b,ΔA,lastpoint,V1,V2,V3]

data=Take[Import["012512RHSAMP1i.csv"],start,.5*freq];
ListPlot[data]
lastpoint=.1;

datum= Take[data,1,lastpoint*freq];

a=LinearModelFit[datum,x,x];
a["ParameterTable","RSquared"]
b=Plot[a[x],x,0,lastpoint];

Show[ListPlot[datum],b]
ΔA=(a[0]-a[lastpoint])*(1/lastpoint);

V1=(ΔA*mL)/(ϵ*1000*mg)

Trial 2

data=Take[Import[”012512RHSAMP1ii.csv”],start,.5*freq];

ListPlot[data]

lastpoint=.1;

datum= Take[data,1,lastpoint*freq];

a=LinearModelFit[datum,x,x];

a[”ParameterTable”,”RSquared”]

b=Plot[a[x],x,0,lastpoint];

Show[ListPlot[datum],b]

ΔA=(a[0]-a[lastpoint])*(1/lastpoint);

V2=(ΔA*mL)/(ϵ*1000*mg)

Trial 3

data=Take[Import[”012512RHSAMP1iii.csv”],start,.5*freq];

ListPlot[data]

lastpoint=.1;

datum= Take[data,1,lastpoint*freq];

a=LinearModelFit[datum,x,x];

a[”ParameterTable”,”RSquared”]

b=Plot[a[x],x,0,lastpoint];

Show[ListPlot[datum],b]

ΔA=(a[0]-a[lastpoint])*(1/lastpoint);

V3=(ΔA*mL)/(ϵ*1000*mg)

Average Subscript[V, 0]

Mean[V1,V2,V3]

StandardDeviation[V1,V2,V3]

SE= %/$\sqrt{n}$

B.3.4 Creating the Hill Plot

Clear[activity,model,bestfit,Hill,V,k,n];

(*Add on package that the fit needs*)

Needs["PlotLegends`"]

Needs["ErrorBarPlots`"]

(*User input; put in data from KineticsVoTemplate.nb*)

activity=1,1.49,2,4.26,3,4.73,4,5.60,6,6.23,7.5,6.49,10,5.61,60,6.52;

(*Model to fit Hill Plot*)

model=((V*x^n)/(k^n+x^n));

(*User input; The fit is sensitive to to the max hill coefficient. If it is too high, it won't fit properly. So you will need to adjust the number accordingly*)

Hill=8;

(*Creates a best fit*)

bestfit=NonlinearModelFit[activity,model,V≥0,k≥0,Hill≥n≥0,V,k,n,x];

(*Creates a table of V, k, and n with standard error*)

bestfit["ParameterTable","RSquared"]

(*User input; put in numbers from the above tables, V is max activity in μmol/min/mg, k is suppose to be $K_{0.5}$ in units of micromols, but it doesn't work out that way so have to solve for that seperately, n is the Hill Coefficient. *)

V_{max}=6.31158;

k=1.59736;

n=2.28658;

(*Defines the Hill equation with the above numbers*)

fit=((V_{max}*S^n)/(k^n+S^n));

(*Solves for $K_{0.5}$, concentration where half the V_{max} is reached, which is defined as S in the above equation b/c $K_{0.5}$ wouldn't work as a parameter*)

sol=Solve[fit==V_{max}/2,S];

(*Renames the answer to previous equation, the 10^{-6} explicitly puts answer in micromols*)

$K_{0.5}$=(S/.sol[[1]])*10^{-6};

(*Gives all the parameters from the Hill Equation*)

V_{max},$K_{0.5}$,n

(*Standard Deviations of average Subscript[v, 0]*)

error1=ErrorBar[0.07];

error2=ErrorBar[0.13];

error3=ErrorBar[0.04];

error4=ErrorBar[0.25];

error6=ErrorBar[0.11];

error8=ErrorBar[0.09];

error10=ErrorBar[0.05];

error60=ErrorBar[0.05];

(*Creates a list plot from the activity data. Will have to input Subscript[V, 0] again*)

data=ErrorListPlot[1,1.49,error1,

2,4.26,error2,

3,4.73,error3,

4,5.60,error4,

6,6.23,error6,

7.5,6.49,error8,

10,5.61,error10,

60,6.52,error60,

PlotStyle-¿Thickness[0.005],PointSize[0.012],Orange];

(*Plots the activity data and the best Hill fit together*)

Show[

Plot[fit,S,0,10,PlotStyle→Black,Thickness[0.005]],

Plot[fit,S,0,60,PlotStyle→Black,Thickness[0.005]],

data,

ListPlot[activity,PlotStyle→PointSize[0.012],Red],

Frame→True,

FrameStyle→Thick,

FrameTicks→0,10,20,30,40,50,60,0,1,3,5,7,,,

FrameLabel→"Specific Activity (μmol/min/mg)"," ","SAICAR (μM)"," ",

LabelStyle→FontSize-¿32,

PlotRange→0,60,0,7,

AxesOrigin→Automatic]

Show[

Plot[fit,S,0,10,PlotStyle→Black,Thickness[0.005]],

data,

ListPlot[activity,PlotStyle→PointSize[0.012],Red],

Frame→True,

FrameStyle→Thick,

FrameTicks→0,10,20,30,40,50,60,0,1,3,5,7,,,

FrameLabel→"Specific Activity (μmol/min/mg)"," ","SAICAR (μM)"," ",

LabelStyle→FontSize→32,

PlotRange→0,10,0,7,

AxesOrigin→Automatic]

Calculating k_{cat} and enzyme efficiency

(*User input; concentration of enzyme used in mg/mL*)

conc=.15;

B.3.5 Calculating k_{cat}

(*User input; volume of enzyme used in the reaction in liters*)

L=2*10^{-6};

(*Tetramer weight in Daltons*)

Da=228000;

(*Seconds in a minute*)

sec=60;

(*Amount of enzyme used in mg*)

mg=(conc*L)/1000;

(*Calculates k_{cat} in sec^{-1}*)

k_{cat}=(V_{max}*Da*mg)/(conc*L*sec)

(*Calculates the Efficiency in M^{-1} sec^{-1}. 10^8-10^9 are the diffusion limit.*)

Efficiency=$k_{cat}/K_{0.5}$

B.3.6 Solutions

50 mM HEPES

Solution will initially have pH ∼4.3. Adjust to pH 7.4 with NaOH

HEPES (g)	dH_2O (mL)
0.30	25
0.60	50
0.84	70
1.2	100

S-AMP

Initially Dilute 25 mg S-AMP in bottle with 9 mL 50 mM HEPES for 6 mM S-AMP.

Dilute 1 mL 6 mM S-AMP in 9 mL 50 mM HEPES for final concentration of 600 μM S-AMP

Store any extra substrate by aliquotting 1.05 mL in eppies and freeze unused aliquots of S-AMP

For 60 μM S-AMP for Linearity

600 μM S-AMP (mL)	50 mM HEPES (mL)
1	9
5	45

Dilutions for Kinetics

Final S-AMP (μM)	S-AMP for Dilution (μM)	S-AMP (mL)	50 mM HEPES (mL)
10	60	0.3	1.5
7.5	60	0.1	0.7
6	60	0.1	0.9
4	10	0.4	0.6
3	10	0.3	0.7
2	10	0.2	0.8
1	10	0.1	0.9

B.4 Fabricating and Purifying SAICAR

B.4.1 Day to Day Details

Day 1

1) Day prior to starting process, turn on water bath to 37 C.

2) Remove purified ADSL from -80 C freezer and allow to warm up to room temp ($\sim$30 min)

3) Turn on Hot Plate, set temp to 100 C.

4) Set up reaction mixture by dilution AICAR 1:5 in SAICAR Synthesis Buffer

5) Place unmixed aliquots of ADSL enzyme and reagents in 37 C incubator for 30 min prior to addition of ADSL to reaction mixture.

6) Add 12 μg of purified ADSL to reaction mixture and invert tube 3-4 times to mix.

7) Immediately take a sample for the analysis by Bratton-Marshall reaction and place tube in 37 C water bath. Start Time.

8) Get out TLC plates and pre-develop them with dH_2O. You will see a yellowish line close to the solvent front. When they're done, Store standing up on the counter over night.

9) At 4 hr, Pre-run Amicon Ultra with 10,000 MWCO and spin for 30 min at 3,000 g with 10 mM Tris.

10) At 5 hr, take a sample for analysis by Bratton-Marshall reaction.

11) Add reaction mixture to Amicon Ultra and spin for 10-20 min at 3,000 g to remove enzyme (keep the filtrate, discard the retentate). Note the volume of retentate and filtrate as these will influence the final yield. Use either our

centrifuge or the one in room 254 (Margittai and Pegan Lab).

12) Spin filtrate in speed vac with heater "off" to reduce total volume to about 400 μL. This takes a fair amount of time.

13) Once total volume is reduced, store in -80 C freezer for tomorrow.

Day 2

1) Set up spotting station with TLC plates, hairdryer, seran wrap, tape, and 25 μL capillary tubes.

2) Prior to spotting, draw a straight line with a pencil $\sim$3 cm from the bottom of a plate to mark the start of the chromatographic separation for calculation of RF values.

3) Spot 200 μL of reaction onto the TLC plates.

4) Fill developing tank with 200 mL 1 M ammonium acetate or 0.8 M lithium chloride.

5) Load plates into developing tank and develop until they are 1 in from the previous developing front.

6) Remove plates from tank and allow them to dry vertically.

7) Visualize spots onTLC plates using hand-held UV lamp located in same drawer as tube holders. Mark location of the solvent front, SAICAR, AICAR, and fumarate with a small tick mark on the side of the TLC plates.

8) Using the distance from the starting line to the locations, calculate RF values using the ratio of the distance traveled by the substrates and the solvent front.

9) Using a clean razor blade, scrape SAICAR from the TLC plates into OakRidge tubes (located on top of the incubator).

10) Elute SAICAR from the cellulose with a total volume of 10 mL 1 N NH_4OH.

11) Remove cellulose by centrifugation in SS34 rotor in Sorvall for 30 min at 16,000 rpm (30,000 g) at 4 C.

12) Remove supernatant from cellulose pellet and discard pellet. Prepare 1 mL aliquots in 1.7 ml eppies.

13) Reduce volume by 5 - 10 fold in speed vac with heater "off". Combine remaining liquid in 1 or 2 eppies.

14) Remove samples for HPLC or Bratton-Marshall and store SAICAR at -80 C.

B.5 Determining Protein Concentration

B.5.1 E 1% = 7.7

This is what Roberta Colman uses for her reserach

1) Take .5 mL from stock protein and put in eppie.

2) Make serial dilutions up to 1:16 in enzyme storage buffer.

3) Find ABS at $\lambda = 280$ nm.

4) A solution of 1 mg/mL has an ABS of 0.77. Use ratio $\frac{1(mg/mL)}{0.77} = \frac{unknownconcentration(mg/mL)}{measuredABS*dilutionfactor}$ to find concentration of protein.

B.5.2 Bio-Rad Assay

1) Make needed amount of 1 x Bio-Rad Dye. Will need 1 mL per sample plus the blank.

2) Check if have 0.2 to 1.0 mg/mL dilutions of BSA to use as standards. If not, see below for recipe.

3) Dilute ADSL to be in the range 0.2 to 1.0 mg/mL. If concentration is unknown, make serial dilutions from 1:2 to 1:8.

4) Add 980 μL of 1 x Dye to acrylic cuvettes.

5) Add 20 μL of standard or unknown ADSL, invert cuvette 3 times and incubate at room temp for 5 - 10 min.

6) Record Abs at λ= 595 nm for each sample. Blank with 1 mL 1 x Dye.

7) Record the data in Bio-Rad template in excel to make the standard curve and calculate ADSL concentration.

BSA Standards

Dilute 100 mg BSA in 10 mL dH_2O

Final Concentration BSA (mg/mL)	BSA @ 10 mg/mL (μL)	dH_2O (μL)
0.2	20	980
0.4	40	960
0.6	60	940
0.8	80	920
1.0	100	900

B.5.3 260-280 nm Concentration Formula

1) Fill cuvette with 2 mL Enzyme Storage buffer and place in position 1 of UV-Vis to use as blank.

2) Make dilution of ADSL to ∼.5 mg/mL in cuvette. If unsure of concentration, or just purified ADSL, dilutions of 1:20 to 1:40 will work. Reocrd Dilution Factor (D.F.) in printed out template.

3) Fill Positions 2 - 6 with ADSL protein as needed.

4) Blank UV-Vis at λ= 280 nm and record ADSL Abs in template.

5) Repeat for λ= 260 nm.

6) Using formula, calculate protein concentration.

Formula

ADSL [ ] (mg/mL) = $(1.55 \times \text{Abs}^{280} - 0.76 \times \text{Abs}^{260}) \times \text{Dilution}$

B.6 How to Use the Static Light Scattering Instrument

B.6.1 Setting Up and Running Experiments

What you need to know about your sample and mobile phase

Sample

1) Concentration in g mL^{-1}.

2) UV Extinction Coefficient in mol g^{-1} cm^{-1}.

3) dn/dc value. For proteins, the default value is .185.

Mobile Phase

1) Refractive Index. If you don't know this, there is a way to experimentally find it. See section below that describes how to do it.

Setting up Instrument for SEC-MALS

1) Turn on the HPLC and Static Light Scattering (SLS) instruments

2) Open the HPLC and SLS software (EZStart 7.4 SP3 and ASTRA)

3) In the HPLC software, open the method "PattersonLabSECMALS" by File $\rightarrow$ Method $\rightarrow$ Open $\rightarrow$ PattersonLabSECMALS.

4) Change the flow rate to 0 mL/min and push "download method". Check to make sure the HPLC now reads 0 mL/min.

5) Connect the size injector loop you will use. Typically the 100 μL loop is used, but smaller loops can be used. Push the injector knob down.

6) Connect the SLS to the HPLC by the two tubes located on the right side of the HPLC. The tube with the blue stripe goes to the "Out" and the other tube goes to the "In". The "In" and "Out" ports of the SLS are located on the left side of the instrument behind a door. You will need to remove two screws before attaching the tubes.

7) Attach the column you will use to the HPLC. There are two ways to do this: i) If you do not need to use the oven, attached the column on the left side of the HPLC using valves 1 ii) if you do need to use the oven, attach the column inside the oven using valves 3. The default for the method is valves 1, if you use ii) you will have to change the valves on the software under the "Oven" tab. When attaching the tubes to the column, make sure the mobile phase flows in the correct direction. The left main valve is where the flow originates.

8) Equilibrate the column by following these steps...

• Note, when changing the flowrate on the column, always to it **gradually** by increasing the flowrate through the column in increments of 0.25 mL/min using the method "outlined" in 4). This is to prevent disturbing the resin at the top of the column by suddenly increasing the flowrate.

• First wash the column with 5 column volumes of dH_2O at 1 mL/min.

• Before switching mobile phase, check for leaks.

• Equilibrate the column with the mobile phase for the experiment. It is best to use 15 - 20 column volumes to equilibrate the column.

• Things to keep in mind when equilibrating the column.

i) Flowrate and pressure limitation of the column.

ii) If it is ok to use your mobile phase on the column.

iii) Molecular weight range of the column.

iv) Since it takes a long time to equilibrate the column, it is best to do it overnight.

v) Make sure you have enough buffer to equilibrate and run experiments. If the column runs dry, you will be held responsible.

• After the column in equilibrated, you are ready to run a sample.

9) If you are using the oven, set the temperature in the HPLC software. Do this by going to the "Oven" tab, clicking the ☐ next to "Enable Oven", set the oven temperature, and push "Download Method". It is best and easiest to equilibrate the column with mobile phase and temperature at the same time.

Running a Sample

Setting up the Software

EZStart

1) Under "Pumps", set the flow rate to your desired value.

2) Under "Time Program", change the ending controller time to the length you need for the experimental run.

3) Under "Column Info", the Wyatt Tech. Co. column info is the default. Change values you need to change.

4) Hit $\triangleright$, name the run, hit start. The orange bar should eventually turn purple, which means the software is waiting the auto-inject signal.

ASTRA

• Whenever you want to close a window in ASTRA, click the "OK" button at the bottom of the window. This action should be taken every time you open something and change it. For example, see step 5)

1) File $\rightarrow$ New $\rightarrow$ Experiment from Template.

2) This should take you to the default folder "My Templates". Choose the "PattersonLabSECMALS" template.

3) A new experiment will appear on the left hand side of the software.

4) Click on the + next to "Configuration". Then click on the + next to "Generic Pump" and "Injector".

5) Double Click (DC) Generic Pump and change the flowrate to match the HPLC flowrate for the experiment.

6) DC Solvent and enter your values. If you have a well known buffer (ie: PBS), you can load its values by clicking on "..." and choose the solution you

have. If you don't know the refractive index and your buffer isn't included in the list provided by ASTRA, you will have to experimentally find it by using BSA. To change the Refractive Index, press the + next to "Refractive Index Model" and put in your value.

7) DC "Injector" and change the injector volume to the volume of solution you inject.

8) DC "Sample" and change the name of sample, dn/dc (if you're not using a protein b/c the dn/dc value of proteins is .185), UV Extinction Coefficient, and concentration.

9) Click the + next to "Procedures". DC Basic Collection and change duration if needed. Hit ▷ to start a run. You should see "Waiting for Auto-Inject" on the graph.

Loading a Sample

1) Make sure the injector knob is in the load position.

2) Load the appropriate amount into the injector. A good general rule would be to inject 150% of the injector loop volume that way you will saturate the loop with sample.

3) Press the injector knob down to inject the loop and start the run.

• Things to think about when running a sample.

i) It is always a good idea for the 1^{st} run to be a blank of the mobile phase, and the 2^{nd} run to be a well known standard like BSA. That way you will know if the injected mobile phase causes any UV or SLS signal and you will know if the setup is working properly if the standard come out correct.

ii) You will need at least 10 min. of baseline after the last peak.

Analyzing Data

1) See section 7 of the ASTRA manual.

2) Do not forget to save your experiment.

Running another Sample

1) Start on 4) in EZStart. You may have to hit the stop button to stop the current run.

2) Start on 1) in ASTRA.

3) Start on 1) in Loading a Sample.

4) Analyze your data.

5) Go through steps 1- 4 above until you are all done with your samples.

Dismantling SEC-MALS

Cleaning the Column

1) Bring the flow to a stop.

2) Reverse the column direction so the flow is now backwards.

3) Clean the column with .2 M KCl @ pH $\sim$3. running @ 0.25 mL/min

4) Cleaning will take 10 to 15 column volumes.

5)While cleaning the column, go to the SLS and, on the interface of the instrument, press the "Tab" button until the laser is selected, press enter. This will turn off the laser. Press "Tab" until the comet is selected, press enter. This will turn on a vibrating mechanism to sonicate the flowcell.

6) After 10 to 15 column volumes, turn off the comet.

7) Bring the flow to a stop.

8) Orient the column so the flow is in the correct direction.

9) Flow 5 column volumes of dH_2O .

10) Flow 5 column volumes of 20% Ethanol.

11) Stop flow and take off column. Place the plugs back and put away in its box.

12) Reconnect the two tubes with a union.

Tidy Up

1) Disconnect the SLS and reconnect the two tubes with a union.

2) Turn off the SLS and HPLC.

3) Shut down the EZStart software.

4) Once you are down analyzing and saving your data, turn off the ASTRA software.

5) Clean up any mess you created while using the SLS.

Find RI value for your running buffer

1) Equilibrate your column with your mobile phase.

2) Run a 100 μL sample of 2 mg/mL BSA.

3) When analyzing your data, fit all the peaks and look at the M_w. Adjust the RI value under "Solvent" until the M_w value is close to 66 kDa. This is your approximate RI value for your running buffer.

Advantages of SLS

1) Relatively quick to run samples - 40 to 90 minutes.

2) Detailed molecular weight and radius values.

3) Can easily look at species distribution (or stoichiometry).

4) Can use a very small amount of sample (100 μL) as long as the concentration

is enough to generate a light scattering signal.

)The other main alternatives are Analytical Ultracentrifugation (AUC) or SDS Gels. AUC is expensive and takes at least 5 hours to run the sample plus 3 hours of sample cell preperation. The analysis of data in non-trivial and can give different results depending on the model you are fitting the data to. SDS works well, but doesn't provide near the amount of detailed information SLS can.

Limitations of SLS

1) Particle radius must 10 nm particles to be detected.

2) Molecular weight for all peaks are averages. Therefore the better SEC you have, the better your results.

B.7 Computer Lab Tips

For running the code to predict the folding pathway of ADSL, I used the computer lab located in the physics building room 118. I found certain peculiarities with using these computers, namely that they "reset" the hard drive every night so if the code needed to run longer than 24 hours, I could not use the computers. The computers did not have a scheduled "reset" on the hard drive, therefore I thought unplugging the ethernet cord would prevent them from "resetting". This gave me the intended result. However, since I used Mathematica, the license requires the computer to be connected to the DU network in order for Mathematica to work. So, it is possible to start the Mathematica kernel while the ethernet cord is connected, disconnect it late at

night so the computer does not "reset" causing you to lose all of the computations, and come back in the morning to reconnect the ethernet. Now, all of this being said, I think it would be much simpler to use MatLab for multiple reasons. It does not need to be connected to the DU network, therefore you can run it overnight with the network disconnected. Mathematica is also a memory hog for larger computations. It seems to store past solutions. So, all this being said, I would just use MatLab or learn how to use the cluster.