Proline Codon Translational Fidelity in *Rhodopseudomonas palustris*: Characterization of Novel *Trans*-editing Factor ProXp-abu

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Jo Marie Bacusmo, B.Sc.

Graduate Program in Chemistry

The Ohio State University

2014

Dissertation Committee:

Karin Musier-Forsyth, Advisor

Michael Ibba

Thomas Magliery

ABSTRACT

Ribosomal protein synthesis is a core biological process essential to all living systems. Proteomic aberrations caused by translational inaccuracies lead to cellular dysfunction and in certain cases, cell death. Aminoacyl-tRNA synthetases (aaRSs) play a central role in protein synthesis. They are responsible for covalent attachment of the correct amino acid to their cognate tRNA substrates. However, this process is inherently error prone due to the isosteric nature of amino acids thus quality control mechanisms have evolved to maintain translational fidelity. In the case of bacterial prolyl-tRNA synthetase (ProRS), a triple-sieve editing mechanism is generally employed, which consists of the ProRS active site that discriminates amino acids based primarily on volume and size, a *cis*-editing domain (INS) that hydrolyzes Ala-tRNA^{Pro}, and a *trans*-editing factor, YbaK, which clears Cys-tRNA^{Pro}. Extensive genomic analysis revealed five distinct *trans*-editing domains homologous to the INS domain. While a subset of these factors correct ProRS-dependent errors, recent studies have revealed distinct substrate specificities and tRNA recognition capabilities that extend beyond Ala- and Cys-tRNA^{Pro}. All known editing mechanisms clear standard non-cognate amino acids; how non-protein amino acids are prevented from misincorporation is unclear. Non-protein amino acids are found in many foods and have the potential to adversely affect human health. The non-protein amino acid aminobutyrate (Abu) is a metabolite involved in various cellular processes. Due to Abu's similarity to Ala and Val, it is recognized and

misactivated by several aaRSs. The metabolically versatile bacterium *Rhodopseudomonas palustris* (*Rp*) encodes for a ProRS containing a catalytically inactive, truncated INS domain, in addition to two distinct INS homologs: YbaK, which deacylates Cys-tRNA^{Pro}, and ProXp-x of unknown function. Comparison of known crystal structures reveals that the catalytic pocket of ProXp-x is larger than that of INS, which suggests substrates larger than Ala are preferred. Indeed, ProXp-x weakly deacylates tRNAs charged with Ile and Val, but robustly edits Abu mischarged onto tRNA^{Pro}, tRNA^{Val}, and tRNA^{Ile}. Semi-promiscuous editing may offer advantages to cells and our data suggest ProXp-x may act as a general Abu-tRNA deacylase. *Rp* ProRS specificity for activation of Pro over Abu is only about 1,000:1, which strongly suggests that editing of Abu-tRNA^{Pro} is required *in vivo*. Taken together, these data suggest that Abu-tRNA editing by the *trans*-editing factor ProXp-x, now renamed ProXp-abu, is likely to be a critical checkpoint to ensure high fidelity in codon translation.

DEDICATION

Dedicated to my Dad, who has inspired and motivated me through my educational life… to my Mom, who taught me to embrace my faith… faith that held me steadfast through challenges in life… and finally to my brothers, who have always motivated me to be the big sister they can be truly proud of… I love you…

ACKNOWLEDGMENTS

I would like to express my sincerest gratitude to my advisor, Prof. Karin Musier Forsyth. I could not ask for a better mentor. I am so blessed to have her as my guide and motivator through such a monumental undertaking in my educational life. I could never thank her enough for all the perpetual support and patience through these years.

I would like to thank my committee members Drs. Thomas Magliery and Michael Ibba, for their valuable advice and helpful insights into my project and for their time and availability.

I would like to thank Dr. Birgit Alber for providing me with expert advice and guidance in the *in vivo* aspect of my project. I would also like to thank Michael Carter and Kelsey Baron who have been kind enough to show me around lab and answer all my endless questions.

I would like to thank Drs. Sandeep Kumar and Mom Das for their patience and perseverance in imparting lab skills necessary to succeed in the KMF lab. Drs. Marina Bhaktina and William Cantara for all their input and assistance with experiments and in helping shape my project into what it is now.

I would like to thank my labmates Oscar, Brianne, Tiffiny, Roopa, Ziwei, Meng and Chris, for all the good times in lab and for being my buddies be it for the gym, fun

night out at short north, or simply a quick lunch/coffee run. You guys are awesome! I would also like to thank the rest of the citizens of KMF Nation, for all your support and fun times together, interactions, bouncing ideas, random conversations, insane moments in lab. I love you guys!

I would like to thank my amazing friends Erik, Carlo, Lara, Minette, Sanjay, Carmille & Joseph for all the support and guidance for being there during the holidays for all the fun times together and for making this journey through grad school a wonderful and meaningful one. You are the best friends I could ever want.

I would like to thank my family, for all the love and support…for being my strength and motivation to overcome challenges in life.

Last but not the least, I would like to give thanks to God almighty, without whom none of this would be possible.

VITA

2005…….. B.Sc. (Cumlaude) Chemistry,
Visayas State University
(former Leyte State University)

2007-2009……..Graduate Teaching Assistant,
The Ohio State University

2007-present……Graduate Research Associate,
The Ohio State University

FIELDS OF STUDY

Major Field: Chemistry

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS/ABBREVIATIONS

A	adenosine or L-alanine
Å	angstrom (unit)
aa	amino acid
aaRS	aminoacyl-tRNA synthetase (three letter amino acid code followed by suffix RS)
aa-tRNA	aminoacyl-tRNA
Abu	2-aminobutyrate
Ala	L-alanine
AMP	adenosine 5'-monophosphate
ATP	adenosine 5'-triphosphate
BSA	bovine serum albumin
C	cytidine or L-cysteine
°C	degree Celsius (centigrade)
CP1	connective polypeptide 1
cpm	counts per minute
CTP	cytidine 5'-triphosphate
Cys	L-cysteine

Da	dalton (unit)
DEPC	diethylpyrocarbonate
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DTD	D-tyrosyl-tRNATyr deacylase
DTT	dithiothreitol
E	enzyme or L-glutamic acid
EDTA	ethylenediaminetetraacetic acid
EF-Tu	elongation factor-Tu
F	L-phenylalanine
g	gram
G	guanosine or L-glycine
Gln	L-glutamine
Glu	L-glutamic acid
Gly	L-glycine
h	hour (unit)
H	L-histidine
Hcy	homocysteine
HEPES	N-[2-hydroxyethyl]piperazine-N'-[2-ethanesulfonic acid]
His	L-histidine
Hse	homoserine

I	L-isoleucine
Ile	L-isoleucine
INS	insertion
IPTG	isopropyl-β-D-thiogalactoside
K	L-lysine
KD	binding constant
kobs	observed rate
L	L-leucine
LB	Luria-Bertani
Leu	L-leucine
Lys	L-lysine
m	mili
M	molar or L-methionine
Met	L-methionine
min	minute
N	L-asparagine
Ntase	terminal tRNA nucleotidyl transferase
nLeu	L-norleucine
nVal	L-valine
p	pico
P	L-proline
PAGE	polyacrylamide gel electrophoresis

PCR	polymerase chain reaction
PEI	polyethyleneimine
Phe	L-phenylalanine
PMSF	phenylmethylsulfonyl fluoride
Ppi	inorganic pyrophosphate
Pro	L-proline
Q	L-glutamine
R	L-arginine
RNA	ribonucleic acid
S	substrate or L-serine
Ser	L-serine
T	L-threonine
Thr	L-threonine
TLC	thin-layer chromatography
Tris-HCl	tris-(hydroxylmethyl) aminomethane hydrochloride
tRNA	transfer RNA
Trp	L-tryptophan
Tyr	L-tyrosine
U	uridine
V	L-valine
Val	L-valine

W	water or L-tryptophan
WT	wild-type
Y	L-tyrosine
μ	micro

Chapter 1 : Introduction

1.1 Ribosomal Protein Synthesis

The process of decoding the genome from DNA to protein is one of the central biological processes in living systems. DNA is transcribed into mRNA and subsequently translated into proteins. Translating the nucleotide message into amino acids is a multi-step process involving several integral players.

The tRNA plays a pivotal role acting as an adapter molecule, which bridges the code between nucleotides and amino acids. The tRNA contains an anticodon sequence complementary to the nucleotide triad on the mRNA, and an acceptor stem onto which the cognate amino acid is covalently attached. Synthesis of the aminoacyl-tRNA (aa-tRNA) is achieved by aminoacyl-tRNA synthetases (aaRS). aaRS belongs to an ancient family of enzymes, which catalyze the covalent linkage of the correct amino acid onto the correct tRNA (1). The synthesized aminoacyl-tRNA is then recognized by elongation factors and escorted to the ribosome for incorporation into the growing peptide chain. Errors in translation may be fatal to the cell; thus, proofreading mechanisms have evolved to maintain translational fidelity. Synthesis of the correct aminoacyl-tRNA by aaRS and proper recognition of the aminoacylated-tRNA by the ribosome are crucial steps in maintaining translational fidelity (1–3).

1.2 Aminoacyl-tRNA Synthetases

Aminoacyl-tRNA synthetases belong to an ancient family of enyzymes that catalyze the step-wise addition of amino acids onto the 3' end of the tRNA (Figure 1.1) (1, 4). The first step involves activation of the amino acid with ATP forming an amino acid adenylate and release of inorganic pyrophosphate. The second step is the transfer of the activated amino acid adenylate onto the 3' end of the tRNA (5). There are two distinct classes of aaRS based on the architecture of their active site, function, and mode of tRNA binding (1, 6, 7). The class I aaRS active site is characterized by the Rossmann fold and two signature motifs, HIGH (His-Ile-Gly-His) and KMSKS (Lys-Met-Ser-Lys-Ser) (1, 8). Class I aaRS generally interact with the tRNA via the minor grove and aminoacylation occurs on the 2'OH (2, 9). On the other hand, the class II active site is comprised of 3 conserved motifs folded into an anti-parallel β-sheet. Motif 1 is integral to self-dimerization and motifs 2 and 3 are essential for amino acid and ATP binding. Binding to the tRNA occurs on the major groove and the amino acid adenylate is commonly transferred onto the 3' OH position (1, 2, 8). In terms of their catalytic function, product release is rate-limiting for class I but not for class II (10, 11). Another distinctive feature between classes is their oligomeric state. Class I aaRS are generally found as functional monomers whereas class II aaRS are typically in a multimeric form (2, 3, 5). In higher taxonomic domains, multiple aaRS interact with each other forming a large complex called multi-synthetase complex (MSC) (12).

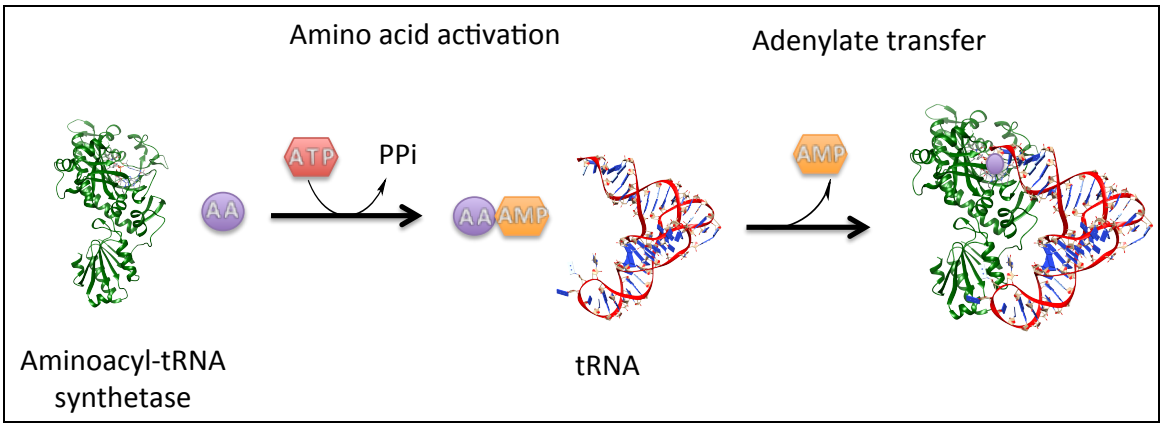


Figure 1.1. The two-step aminoacylation reaction. The first step involves activation of the amino acid with ATP forming an amino acid adenylate complex with release of inorganic pyrophosphate. The second step is the esterification of the amino acid adenylate onto the 3'-end of the tRNA. Here shown is Rp ProRS (unpublished crystal structure from S. Cusack).

1.3 Extra Domain Structures of Prolyl-tRNA Synthetases

Throughout evolution, tRNAs and aaRS co-evolved to adapt to requirements of the cell. Although all Proly-tRNA synthetases (ProRS) maintain a class II aminoacylation core and a class IIa anticodon binding domain, various extra domains have been appended to ProRS in each of the different taxonomic domains (Figure 1.2) (13). ProRS are divided into 2 types. The Eukarya and Archaeon-like ProRS and the Prokaryotic-like ProRS. These 2 types are architecturally distinct from each other (13–15).

Eukaryotic and Archaeon ProRS are typically about 480 residues in size and encode for an ~80 residue C-terminal extension beyond the anticodon binding domain (13, 16). The C-terminal extension folds into a long α-helix terminated by an anti-parallel

β-sheet structure similar to the aminoacylation core (13). The C-terminal extension contains 4 Cysteines, conserved throughout most eukaryotic-like ProRSs that fold into a tetrahedral zinc-binding site. In *Thermus thermophilus* (Tt), the C-terminal extension contains 5 Cys residues forming the zinc-binding site. The 5th Cys is free but still demonstrated to be able to bind ions (13). Interestingly, in lower eukarya such as in *Sacaromyces cerivisiae* (Sc), a 180 residue N-terminal extension is displayed in addition to the C-terminal extension (16, 17). Moreover, this N-terminal extension is homologous to an insertion domain found only in Prokaryotic-like ProRS (17).

Prokaryotic ProRS commonly encode for a ~180 residue insertion (INS) between motifs 2 and 3 (16, 18). This insertion folds into an anti-parallel β-sheet structure and forms a separate domain linked to the aminoacylation core by 2 long unstructured loops. Unlike the C-terminal extension, INS possesses catalytic activity and has been shown to edit mischarged Ala-tRNA^{Pro} but not Cys-tRNA^{Pro} (16, 19). Substrate discrimination is achieved via a size exclusion principle, where smaller Ala is accepted but not larger Cys and Pro (20). In *Escherichia coli* (Ec), a highly conserved critical Lys in the 279 position in interacts with the backbone phosphate of A76 on the tRNA and is proposed to be essential for proper orientation of the substrate in the catalytic pocket (16, 21). Mutating K279 into Ala significantly decreases editing function (16).

In a smaller subset of bacterial ProRS, the INS domain is lacking and a truncated mini-INS domain is presents, comprised of ~25 residues, which fold into 2 α-helices joined by a loop. Sequence analysis reveals mini-INS aligns to various fragments within the INS (22). Unlike INS, which folds into an independent domain, the mini-INS is fused

to the aminoacylation core by folding over the posterior side of the synthetic site, thereby projecting into the dimerization interface. Mini-INS appears to lack any catalytic activity but its conservation and impact on overall ProRS function has not been carefully investigated (15).

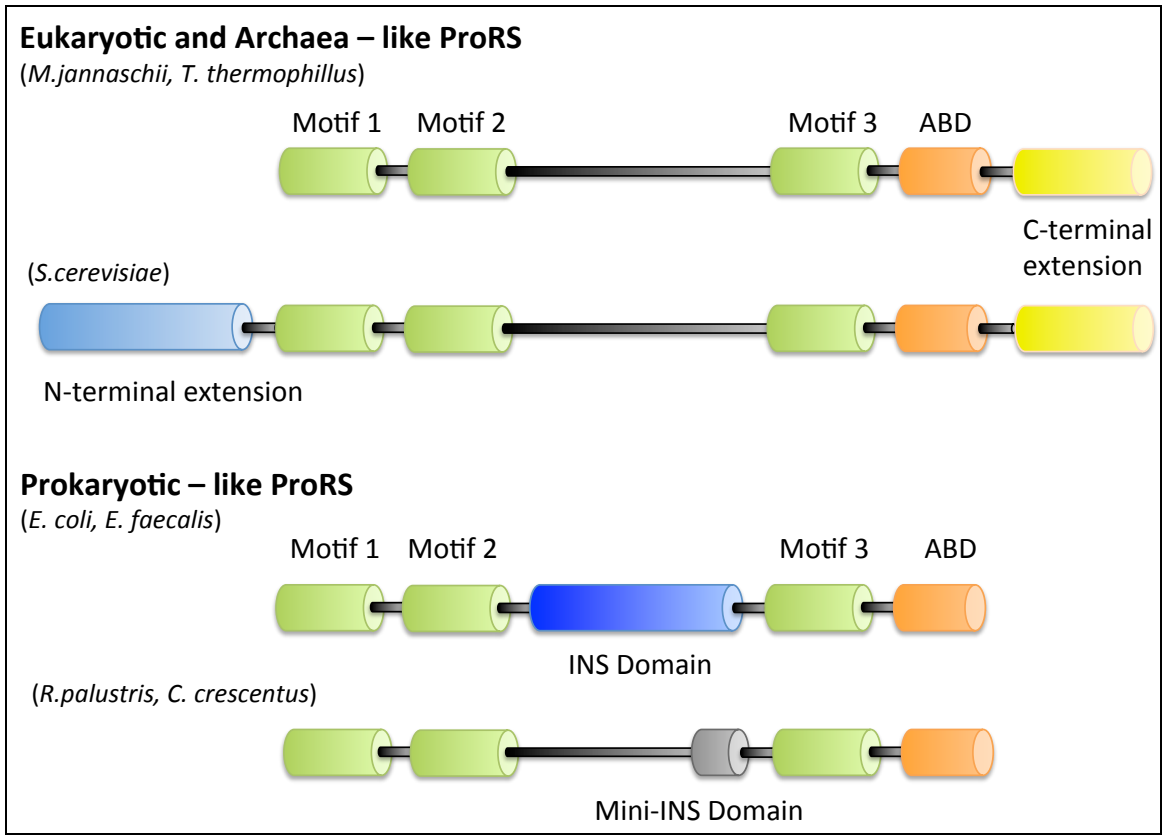


Figure 1.2. Domain Representation of Structures on ProRS. Conserved motifs 1-3 are shown in green together with the anticodon binding domain (ABD), shown in orange. The C-terminal extension (yellow), N-terminal extension (light blue), insertion domain (INS) (blue), and mini-INS (gray) are also shown.

1.4 Editing by Aminoacyl-tRNA Synthetases

Maintaining translational fidelity is essential as aberrations in the proteome may lead to cellular dysfunction and death. Error-free translation of a nucleotide triad into an amino acid is dependent on the attachment of the correct amino acid onto the correct tRNA by the aaRS (1, 5, 23). The error rate of this synthesis step is about 1 in 10^4, primarily due to the inherent similarities in volume and size of amino acids (4, 23). Although amino acids are diverse in their display of various side-chain functional groups, their overall structural relatedness poses a challenge for many aaRS (1, 3). In general, enzymes cannot differentiate between two substrates that differ by only one methyl group (3, 24). For example, isoleucyl-tRNA synthetase (IleRS) cannot distinguish between cognate Ile and non-cognate Val. Hence, Val is recognized and misactivated by IleRS (25). This limitation in substrate specificity does not only apply to proteinogenic amino acids. Non-proteinogenic amino acids are often misincorporated by aaRS (26–28). The striking similarity of 2-aminobutyrate (Abu) to branched amino acids, not only in terms of size and volume but also because of its side-chain properties, causes it to be misactivated by ValRS, and IleRS (29–31). Moreover, it also shares properties with Ala; thus, aaRS that recognize Ala may also accept Abu (e.g., ProRS and AlaRS) (32). To overcome this inevitable propensity for error, aaRS have evolved several editing mechanisms to correct translational errors (Figure 1.3). “Pre-transfer” editing occurs in the first step of aminoacylation, where hydrolysis of the non-cognate aminoacyl adenylate is initiated by the aaRS to yield free amino acid and AMP. This occurs in either a tRNA-dependent or -independent manner (33). Conversely, “post-transfer” editing

occurs after the second step, where hydrolysis of the misacylated tRNA is performed by an independent editing domain (Figure 1.3) (25). These editing domains are either fused to the synthetic core as a *cis*-editing site, or function in *trans* as autonomous enzymes (14, 16, 34). In some cases, a single editing mechanism is not sufficient for correcting errors in translation; thus, several strategies and combinations of editing mechanisms may be employed to ensure synthesis of the correct aminoacyl-tRNA (Table 1.1).

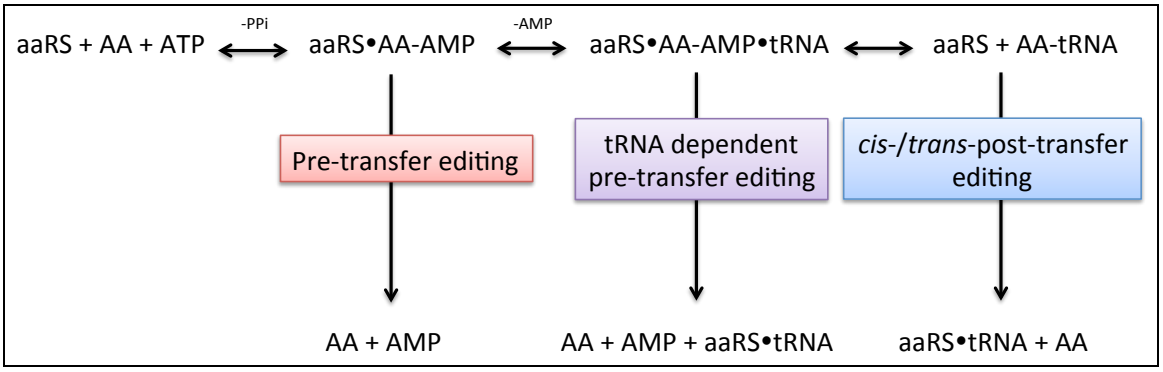


Figure 1.3. Schematic representation of the various editing pathways of aaRS (3).

1.4.1 Pre-transfer Editing

Several models of pre-transfer editing have been proposed; this include translocation to an independent editing site, selective release of the amino-acid adenylate, and active hydrolysis in the aminoacylation core (2–4).

Baldwin and Berg first discovered hydrolytic editing in the IleRS system, where they observed a tRNA dependent pre-transfer editing of Val-AMP (35). Active

translocation is the supported model of pre-transfer editing for IleRS (36–38). Several biophysical and biochemical studies have shown that upon tRNA binding, the CP1 domain rotates 47°, thereby creating a channel through which non-cognate Val-AMP is actively shuttled into the CP1 domain for hydrolysis (36, 37, 39). Furthermore, crystal structures of Val-AMP analog bound in the CP1 domain active site weakly support this as the catalytic site for pre-transfer editing (40).

In the selective release model, the affinity for the noncognate amino acid adenylate is significantly lower compared to cognate, thereby mediating its release into the solution where hydrolysis of the unstable moiety rapidly occurs (31, 41). In certain cases, kinetic proofreading occurs when the rate of adenylate release is faster than the rate of transfer onto the tRNA (42).

Active hydrolysis is a tRNA-independent pathway where editing of the miactivated noncognate amino acid occurs within the aaRS active site. Studies on bacterial LeuRS containing an editing defective CP1 domain (caused by deleterious mutations or complete deletion of CP1 domain) suggest pre-transfer editing also occurs in the aminoacylation site (40, 43, 44). Similarly, ProRS uses a combination of active hydrolysis and selective release to clear misactivated Ala-AMP (45). ProRS has been shown to robustly edit aa-AMP in the absence of tRNA. Although ProRS encodes for an editing site, pre-transfer editing solely occurs in the aminoacylation site (45). This observation is consistent with studies of the editing defective Ec ProRS K279A, and of several ProRS that do not encode for an editing domain such as *Methanococcus jannaschii* and *Caulobacter crescentus* (22, 45).

1.4.2 Post-transfer Editing

In the event of a misactivated noncognate amino acid adenylate successfully esterified onto the tRNA, post-transfer editing is required. This highly specialized proofreading mechanism occurs in a separate editing site independent of the synthetic core. The principle of post-transfer editing is analogous to the use of multiple specialized sieves, where utilization of several highly specialized catalytic sites is significantly advantageous compared to use of a single catalytic site for multiple functions (34, 46).

The aaRS active site is considered as the coarse sieve, which screens against much larger and smaller amino acids, while amino acids similar in size to the cognate are recognized and misactivated. Most aaRS hydrolyze and clear noncognate amino acid adenylates. However, this editing mechanism may not be sufficient when the rate of transfer is faster than the rate of adenylate release; thus, the misactivated amino acids are transferred onto the tRNA resulting in the synthesis of a mismatched aa-tRNA pair. In this case, post-transfer editing acts as a fine sieve, where misacylated aa-tRNAs are specifically recognized and cleared, releasing free amino acid and tRNA (31, 34). Post-transfer editing occurs either in *cis*, where the editing domain is fused to the aminoacylation core, or in *trans*, where free-standing editing domains act on mismatched aa-tRNAs Cis independent of the aaRS (14, 16, 34).

1.4.2.1 *Cis*-editing

Hydrolysis of aminoacyl-tRNA was first demonstrated in the ValRS system where formation of Thr-tRNA^{Val} followed by rapid hydrolysis was observed (47). ValRS,

as well as most class I aaRS, encode for an insertion domain (highly conserved connective polypeptide 1, CP1) demonstrated to possess editing function (39, 48–50). The CP1 domain is present as a non-catalytic ~150 residue insertion in MetRS, and ~250 residue insertion domain in ValRS, IleRS, and LeuRS (51). In 1977, Fersht initially proposed the concept of a double-sieve editing model based on the IleRS system (25). However, it wasn't until 2 decades later that direct evidence for post-transfer editing was shown. Cross-linking studies showed a tagged Val-tRNA^{Ile} can be crosslinked to both the IleRS active site and CP1 domain, while the tagged cognate Ile-tRNA^{Ile} only crosslinked to the active site (52). This observation supports translocation of the misacylated Val from the synthetic site and into the CP1 editing site. In 1998, crystal structures elucidating the double-sieve editing model were solved (48). Val-IleRS complexes indicated a primary site (coarse sieve), big enough to accommodate both Ile and Val, and a secondary smaller site (fine sieve) where only Val can be accommodated. Deletion of 47 residues around the second site completely abolishes hydrolytic editing and accumulation of Val-tRNA^{Ile} is observed (34, 48). Furthermore, expression of CP1 as an independent domain does not affect its hydrolytic activity (53, 54).

Post-transfer editing has also been observed for several class II aaRS including AlaRS, ThrRS, PheRS, and ProRS. In contrast to the class I CP1 domain, the class II editing sites are distinct from each other (3). ThrRS encodes for a ~158 residue domain (*E. coli*) called N2 domain (55, 56). This has been shown to possess editing activity against Ser-tRNA^{Thr} via water mediated hydrolysis mechanism. The N2 domain is

located 39 Å away from the aminoacylation domain thus translocation of the tRNA 3'-end from the synthetic site to the editing site is supported in this model (56).

In AlaRS, the editing site for Gly- and Ser-$tRNA^{Ala}$ consists of a highly conserved specialized internal domain located ~37 Å away, posterior to the synthetic site (57). This suggests translocation of the 3'-end of the tRNA is unlikely to occur (58, 59). Characterization of the editing site reveals two RNA complex formations; tRNA bound to the synthetic site, and tRNA bound to the editing site. This suggests AlaRS uses a different model of editing, where the misacylated tRNA released from the synthetic site to allow reassociation with the editing site in *trans* (59).

PheRS encodes for its synthetic site and editing site in separate subunits, α and β respectively, and exists in a $(\alpha\beta)_2$ heterotetramer oligomeric structure in solution (60, 61). All PheRS are active in the heterotetrameric form except for mitochondrial PheRS, which displays activity in its monomeric form (62). The B3/B4 editing site is specific for Tyr-$tRNA^{Phe}$ and is located ~20 Å away from the synthetic site, which readily supports the translocation editing model (60). In *E. coli*, EF-Tu is unable to distinguish between Tyr-$tRNA^{Phe}$ and Phe-$tRNA^{Phe}$ and shuttles both moieties into the ribosome for protein synthesis. This highlights the significance of PheRS post-transfer editing as the primary quality control checkpoint in maintaining fidelity in Phe codon translation (63).

ProRS encodes for a 180 residue insertion domain (INS) shown to possess robust editing activity against Ala-$tRNA^{Pro}$ but not Cys-$tRNA^{Pro}$ (16, 64). Substrate discrimination is based on a size exclusion principle where smaller Ala is accepted in the editing site but not larger Pro and Cys (20, 65). The INS encodes for a universally

conserved Lys on position 279 in *E. coli*, which is proposed to interact with the phosphate backbone on A76 on the tRNA and is proposed to aid in proper orientation of the substrate (16, 20). Mutating this residue to Ala significantly reduces editing activity (16). In some lower eukarya, a catalytically inactive N-terminal extension moderately homologous to INS is displayed (17). Eukaryotic and Archaeon ProRS lack the INS but encode for a C-terminal extension instead, shown to lack any catalytic activity (14). In a small subset of bacterial ProRS, the INS domain is severely truncated and has been shown to be catalytically inactive (22). This suggests these aaRS may have evolved alternate proofreading mechanisms, or they may have evolved highly specialized synthetic cores to discriminate against noncognate amino acids (14, 22). Crystal structure of *Ef* ProRS reveals INS is about 35 Å away from the synthetic site suggesting a translocation mechanism similar to the N2 domain in ThrRS may be employed (15). Similar to the CP1 domain present in class I aaRS, isolation of INS and B3/B4 editing domains from their respective full-length enzymes show they retain their hydrolytic activity (66, 67). This observed trend of independent editing activity of isolated domains suggests the existence of functional autonomous editing factors.

1.4.2.2 *Trans*-editing

Autonomous editing enzymes that specifically recognize D-Tyr-tRNA^{Tyr} (DTD) were the first evidence of *trans*-editing (68). Sequence and structural analysis suggests DTD shares ancestry with the N2 domain on *Pyrococcus abyssi* ThrRS and may have evolved to gain substrate enantioselectivity (69–71). Phylogenetic analysis with aaRS *cis*-

editing domains has lead to the discovery of several other *trans*-editing factors. The most common of which are ones associated with editing $tRNA^{Ala}$, $tRNA^{Thr}$, and $tRNA^{Pro}$.

Most archaeal ThrRS are devoid of an N2 domain shown to be important for editing Ser-$tRNA^{Thr}$ (72). Genomic analysis of crenarchaeal organisms reveal the catalytic and editing domains are encoded as separate genes referred to as ThrRS-cat and ThrRS-ed respectively. ThrRS-cat has been shown to synthesize Ser-$tRNA^{Thr}$ in addition to cognate Thr-$tRNA^{Thr}$ and lacks hydrolytic editing activity. The autonomous ThrRS-ed maintains translational fidelity by clearing Ser-$tRNA^{Thr}$ in *trans* (72).

In addition to the catalytic domain and editing domains, the AlaRS system also includes a free-standing domain called AlaXp, which is homologous to the AlaRS editing domain. AlaXp is ubiquitous among all Domains of life and possesses robust editing activity against Ser-$tRNA^{Ala}$ (14, 73). Unlike the AlaRS editing domain, which recognizes both Gly-and Ser-$tRNA^{Ala}$, most AlaXps only recognize Ser-$tRNA^{Ala}$. Interestingly, a small subset of AlaXps have been shown to recognize Gly-$tRNA^{Ala}$ (74). AlaXps are divided into 3 types based on their sizes (75). AlaXp-Ia (AlaXp-S) is the smallest of the three and contains only the minimal requirements for hydrolytic function (74, 76). AlaXp-Ib (AlaXp-M) contains a short N-terminal extension is addition to the catalytic core (77, 78). Finally, AlaXp-II (AlaXp-L) is similar in size to the AlaRS editing domain and encodes for a C-terminal extension (C-Ala) (58, 59, 73).

Taxonomic sequence analysis has revealed a family of 5 autonomous enzymes, homologous to the ProRS INS domain, comprised of YbaK, ProXp-ala (PrdX), ProXp-x (ProX), ProXp-y (YeaK), and ProXp-z (PA2301) (14, 15, 21, 22, 79, 80). These domains

share structural and sequence similarities with INS but are clearly distinct from one another (22). Of the 5 members of the INS superfamily, only YbaK and ProXp-ala have been extensively characterized (14, 22, 79). YbaK is shown to posseses robust editing activity against misacylated Cys-tRNA^{Pro}. Unlike INS and the synthetic core which use size exclusion to confer specificity, Ybak uses chemical properties to specifically recognize Cys and reject other similar sized amino acids such as Pro (21). Furthermore, YbaK utilizes a unique sulfhydryl-cyclization mechanism to clear Cys. Although YbaK demonstrates stringent amino acid substrate specificity, it does not show any tRNA specificity. YbaK acts as a general deacylase and has been shown to deacylate cognate Cys-tRNA^{Cys} (79). Biophysical studies demonstrate YbaK gains specificity though interaction with ProRS in a complex formation (79). The discovery of YbaK answered the Cys paradox that plagued the ProRS system. ProRS catalyzes the synthesis of both Ala-and Cys-tRNA^{Pro} but only Ala-tRNA^{Pro} is recognized in the *cis*-editing site (16). This suggests a separate machinery must be utilized to clear misacylated Cys-tRNA^{Pro}. This lead to the triple-sieve editing model where the aminoacylation site acts as the coarse sieve by rejecting larger amino acids, the *cis*-editing domain INS acts as a fine sieve by clearing Ala-tRNA^{Pro}, and the highly specialized YbaK clears Cys-tRNA^{Pro} in *trans* as a chemical sieve (79, 81).

ProXp-ala has been shown to efficiently edit Ala-tRNA^{Pro} (14, 22). Unlike YbaK, ProXp-ala shows high tRNA specificity and is able to discriminate between Ala-tRNA^{Pro} and Ala-tRNA^{Ala} (22). Interstingly, ProXp-ala is present in ProRS systems lacking a full-length INS domain such as in *C. crescentus*. In these organisms, an alternative triple

sieve editing mechanism is employed where translational fidelity is maintained through the use of two distinct *trans* editing factors (YbaK and ProXp-ala) (22). The existence of three other distinct freestanding homologs ProXp-x, ProXp-z, and ProXp-y, whose functions have yet to be characterized, suggests different editing models may be employed by other systems. Moreover, these *trans*-editing factors are widely distributed among all Domains of life in various species-specific combinations (22). Co-evolution of these autonomous editing domains highlights the diversity of editing strategies used by various systems to adapt to challenges in codon translation.

1.5 Non-protein Amino Acids and Translation

Amino acids are biologically essential compounds composed of an amino group and a carboxylic group attached to a carbon atom (α-carbon) displaying a diverse array of functional side-chains. Only ~2% (20 protein amino acids including pyrolysine and selenocysteine) of the total number of amino acids found in nature are commonly utilized by organisms in protein synthesis (82). Many non-protein amino acids are synthesized and selectively utilized by plants and animals for ecological advantage over pathogens, predators, and competing organisms (83, 84). Fescue grasses secrete phytotoxins primarily composed of *meta*-tyrosine (m-Tyr) to compromise root development of competing plants (85). *Meta*-tyrosine is a non-protein amino acid misacylated by PheRS onto $tRNA^{Phe}$ resulting in misincorporation of m-Tyr instead of Phe in the proteome (85, 86). The highly poisonous compound azetidine-2-carboxylic acid (Aze), synthesized by *Convallaria majalis* (Lilly of the Valley) and present in trace amounts in *Beta vulgaris*

(beets), has been shown to replace Pro in protein sequences resulting in lethality in organisms that do not synthesize this non-protein amino acid (87, 88). *Canavalia ensiformis* (jack bean) synthesizes the arginine-mimetic L-2-amino-4-(guanidinooxy)butyric acid (canavanine) found to cause lethality in the larvae of its predators by misincorporation in Arg codons resulting to aberrant proteins (89–91). Non-protein amino acids have been exploited by plants as armaments for ecolocogical preservation and survival (26). Consequently, these toxic compounds are found in common produce such as legumes, fruits, and nuts; thus, they are generally present for consumption in human and animal diet (26).

Non-protein amino acids pose a threat to some living systems, as they are able to act as antimetabolites by interfering with metabolic processes that utilize protein amino acids, and as proteomimetics through misincorporation into the proteome (27, 84, 92). Misincorporation of non-protein amino acids has been implicated in a wide variety of diseases and physiological disorders (26, 27).

A complex neurological disorder where individuals slowly develop symptoms of amytrophic lateral sclerosis (ALS), Parkinson's disease (PD) and dementia, collectively known as ALS-PDC, has been linked to consumption of cycad flour containing the non-protein amino acid β-methylamino-L-alanine (BMAA) (93, 94). BMAA was identified in kernels of *Cycas circinalis* (queen sago) and demonstrated to cause a degenerative motor-system disease when administered to *Macaca fascicularis* (cynomolgus monkeys) (95–97). It is also synthesized by cyanobacteria (*Nostoc sp.*) found in cycad roots (98). Recently, BMAA has been shown to infiltrate Ser codons resulting in synthesis of

aberrant proteins prone to misfolding and aggregation (99). The Tyr mimetic L-3,4 dihydroxyphenylalanine (L-DOPA) has been shown to generate aggregate-prone and protease-resistant proteins in human cells *in vitro* (99, 100). L-DOPA misincorporated into proteins has been shown to possess cytotoxic properties *in vitro* and potentially cause progressive chronic toxicity *in vivo* (101). *Macuna* species indigenous to Central America naturally synthesize L-DOPA, constituting 6-9% of their seeds (86, 100).

The Pro mimetic Aze has been implicated in neurodegeneration, autoimmune diseases, and multiple sclerosis (MS) (92). Rubenstein established a tight link between the geographical distribution of beet agriculture and the worldwide prevalence of MS (88, 102). Furthermore, Aze has been shown to hijack Pro codon translation and is proposed to alter myelin basic protein structure and function (103).

In order to overcome translational errors caused by misincorporation of non-protein amino acids, aaRS need to evolve better and highly specialized proofreading strategies. The predator *Caryedes brasiliensis* (bruchid beetle) has evolved an arginyl-tRNA synthetase (ArgRS) with high specificity for Arg, effectively discriminating between cognate Arg and canavanine (104). Moreover, bruchid beetles utilize canavanine as a nitrogen source for amino acid biosynthesis (105). Jack bean, a canavanine producer, has been shown to evolve an ArgRS that can discriminate against canavanine and prevent autotoxicity (106).

1.6 Significance of Editing *In Vivo*

Studies of cellular systems with editing-defective aaRS systems indicate that editing is essential for cell survival. In mice a single point mutation in the editing domain of AlaRS causes Ala to Ser mistranslation, which has been linked to protein misfolding and neurodegeneration (107). In bacterial systems, mistranslations have been shown to exhibit lethality. An editing defective ValRS, bearing the T222P mutation in its CP1 domain, is fine under normal conditions but is extremely sensitive when exposed to noncognate Abu and Thr (29, 30). In the case of IleRS, loss of editing activity results in an enhanced mutation rate in aging bacteria (108).

Organisms vary in threshold for proteomic mistranslations. For example, *E. coli* can tolerate misincorporations to about 10% and still retain regular cellular function. In some cases, mistranslation is required to restore shallow growth (109). Under severe growth conditions such as starvation and oxidative stress, mistranslations may be beneficial to the cell. Expression of both WT and editing-defective *E. coli* IleRS under starvation condition supplemented only with Val and norvaline (Nval) showed the editing defective IleRS exhibited a better phenotypic response compared to wild-type (110, 111). Some parasites exhibit phenotypic plasticity to evade host defense mechanisms. aaRS natural harboring point mutations and/or deletions in their editing domains is common in *Mycoplasma* parasites. The stastical misincorporations allow for the organism to develop antigen diversity thereby escaping its host's defenses (112).

CLASS	aaRS	Post-transfer editing site	*Trans*-editing factor homolog	Amino acids activated but not edited	Amino acids edited		
					Pre	Post-*cis*	Post-*trans*
Class I	IleRS	CP1			Val, Hcy, Cys, Thr	Val, Abu	
	ValRS	CP1		Ile	Hcy, Cys, Ala	Thr, Cys, Abu	
	LeuRS	CP1			Val, γ-hLeu, NLeu, Nva	Ile, Met	
Class II	ThrRS	N2	ThrRS-ed	β-hNva		Ser	
	ProRS	INS	YbaK, ProXps	Aze, 4-hPro	Ala	Ala	Ala (ProXp-ala), Cys (YbaK), Abu (ProXp-x)
	AlaRS	N2-like	AlaXps			Gly, Ser	Gly, Ser (ProXp-y)
	SerRS				Thr, Ser, Cys		
	PherRS	B3/B4			Tyr	Ile, Tyr, m-Tyr	
	LysRS-II			Arg, Ala, Thr, Met, Leu, Cys, Ser	Orn, Hcy, Hse		

Table 1.1. Summary of editing domains on aaRS and noncognate amino acids that are edited. This table is adapted from (2, 3). Abbreviations used: Abu, 2-aminobutyrate; Aze, azetidine-2-carboxylic acid; Hcy, homocysteine; Hse, homoserine; Nva, norvaline; β-hNva, β- hydroxynorvaline; Nleu, norleucine; γ-hLeu, γ-hydroxyleucine; 4-hPro, 4-hydroxyproline; m-Tyr, *meta*-tyrosine; Orn, ornithine

1.7 Purpose of this Study

Synthetases use various strategies in maintaining fidelity in codon translation. In this study, we investigate an alternative ProRS system distinct from the triple-sieve editing mechanisms elucidated in *E. coli* and *C. crescentus*. *R. palustris* (*Rp*) encodes for a ProRS with a severely truncated INS domain (mini-INS), whose relevance to the overall structure and function of the ProRS system has not been extensively studied before (**Chapter 2**), in addition to two *trans*-editing factors YbaK and ProXp-x. Our results show the mini-INS is essential for proper protein folding of the aminoacylation core. Although it lacks hydrolytic activity, the synthetic core has evolved compensatory editing mechanisms to ensure translational fidelity, in part, by increasing stringency in amino acid discrimination. Consistent with YbaK proteins characterized from various organisms (*E. coli, Haemophilus influenzae*), *Rp* Ybak displays robust editing activity against Cys-tRNA^{Pro} (**Chapter 3**). We also characterize the freestanding INS homolog ProXp-x, whose function has not been reported to date. Here we show ProXp-x lacks activity against Ala- and Cys-tRNA^{Pro}. Sequence and structural analysis of the active sites revealed the putative substrate binding pocket of ProXp-x is larger compared to INS, which suggests substrates larger than Ala may be accomodated. Indeed, ProXp-x possesses robust deacylase activity against Abu-tRNA^{Pro} (renamed ProXp-abu) and is likely a quality control checkpoint in maintaining translational fidelity. Furthermore, we delve into the molecular basis for amino acid and tRNA substrate discrimination by ProXp-abu (**Chapter 4)**. Our data shows ProXp-abu active site is not tunable, suggesting amino acid discrimination is not based on steric properties of substrates, unlike in the

case of INS. Moreover, comparison of aa-CCA MD simulated docking models of ProXp-abu and INS shows the substrates are bound 180 degrees with respect to each other, which suggests mechanism of deacylation as well as substrate selection is distinct to that of INS. We also show that ProXp-abu has a relaxed tRNA specificity with very weak recognition for the discriminator base. Over 50% of tRNAs encode for an A73 thus ProXp-abu may function as a general deacylase in the cell. Owing to the alternative function of ProXp-abu *in vitro*, we proceeded to establish a link between the observed Abu deacylase activity *in vitro* and its role in biological systems *in vivo* (**Chapter 5**). ProXp-abu and YbaK null stains in *Rp* have been generated as key tools in gaining a better understanding of the roles they play in maintaining cellular homeostasis. Finally, we have summarized all our findings and concluded that ProXp-abu and YbaK indeed are important for maintaining fidelity in Pro codon translation in *Rp* (**Chapter 6**).

Chapter 2 : Non-catalytic mini-INS domain is vital for structure and function in bacterial ProRS

2.1 Introduction

Translational fidelity is essential for cell survival as errors in translation may lead to toxic mutations in the proteome potentially leading to abnormalities, diseases, and in some cases, death (2). Hence, several translational quality control checkpoints are in place. One of the integral steps in protein synthesis is the correct formation of aminoacyl-tRNAs (1). This is achieved by aminoacyl-tRNA synthetases (aaRS), an ancient family of enzymes, which catalyze the addition of the cognate amino acid to its cognate tRNA. The aminoacylation step is error prone (~$1/10^4$), primarily due to the inherent similarities in size and volume of the amino acid substrates (1, 23). To overcome this limitation, aRS have evolved editing mechanisms to maintain translational fidelity (51). There are two major editing mechanisms employed by aRS; pre-transfer editing, hydrolysis of the misctivated amino acid to free amino acid and AMP, and post-transfer editing, which requires a separate editing domain to clear the misacylated amino acid from the 3'-end of the tRNA (25, 33).

In the case of prolyl-tRNA synthetases (ProRS), Ala and Cys are misacylated in addition to cognate Pro. Thus, several editing mechanisms are employed to ensure fidelity of Pro codon translation. In many bacterial systems, a triple-sieve editing

mechanism is employed, which consists of the ProRS active site that discriminates amino acids based largely on volume and size, the ProRS *cis*-editing domain (INS) that hydrolyzes Ala-tRNA^{Pro}, and a *trans*-editing factor called YbaK that clears Cys-tRNAPro via unique sulfhydryl side-chain chemistry (16, 21, 81).

ProRSs, across all taxonomic domains encode for an N-terminal class II catalytic core, comprised of 3 conserved motifs, linked to a C-terminal class IIa anticodon binding domain (13). Class II aRS generally exist as dimers or tetramers in nature (2). In the case of ProRS, it exists as a functional dimer formed by 2 monomers in an antiparallel orientation where oligomerization is primarily mediated by elements on motif 1 (13). Phylogenetics on ProRS across all domains of life reveal 3 prominent extra domains commonly found in ProRS (13, 22). The C-terminal extension, typically displayed on Eukaryotic and archaeon ProRS, is an ~80 residue domain encoded after the anticodon binding domain. It folds into a long α-helix terminated by a mixed β-sheet domain proximal to the catalytic domain (13, 113). The C-terminal extension displays Cysteines that fold into a tetrahedral zinc binding site generally conserved throughout eukaryotic ProRSs except for some organisms such as *Methanococcus jannaschii* and *Mycoplasma* (13).

On the other hand, prokaryotic ProRS lack a C-terminal extension and instead, encode for a ~180 amino acid insertion domain (INS) between motifs 2 and 3 (16). The INS is comparable in size to the CP1 domain in Class I ValRS, IleRS, and LeuRS and similarly possesses robust post-transfer editing activity against their respective misacylated tRNA (13, 15, 16). INS has been show to act as a *cis*-editing site in ProRS

highly specific for Ala-tRNA^{Pro} but not Cys or Pro-tRNA^{Pro} (16). A size exclusion mechanism is employed to discriminate between substrates. Smaller Ala fits in the catalytic pocket while larger Cys and Pro are rejected (65).

In some bacterial ProRS, the INS domain is truncated (mini-INS) (15). It is a small domain of about 25 amino acids, which folds into 2 α-helices joined by a loop. About 10% of bacterial ProRS encode for a truncated INS domain (mini-INS) (22). Previous studies suggest the mini-INS does not have any catalytic activity (15). However, the significance of the mini-INS has not been extensively characterized before. In this study we show the mini-INS is essential for preserving overall structure and function of *Rhodopseudomonas palustris* (*Rp*) ProRS. The availability of crystal structures (2I4L, 2I4M, 2I4N, 2I4O) presents advatages in studying structure and function of *Rp* ProRS (15). Furthermore, the lack of a full-length INS domain suggests the system may have evolved alternative editing strategies to maintain Pro codon translation. Survey of the *Rp* genome shows it encodes for 2 *trans*-editing factors YbaK and ProXp-x. YbaK has been shown to clear Cys-tRNAPro (79) while the function of the latter is currently unknown. Here we gain insight into the role mini-INS plays in this alternative editing system.

2.2 Experimental Procedures

2.2.1 Materials

All amino acids and chemicals were purchased from Sigma unless otherwise noted. DNA Primers were purchased from Integrated DNA Technologies (IDT).

2.2.2 Preparation of *Rp* ProRS Mutants

Rp ProRS pET15b was acquired from Stephen Cusack (EMBL Grenoble) and used as template for all mutagenesis procedures. Deletion mutants were prepared using overlap PCR (114) with the following primers: Δmini-INS 5'-TCCGGCGT CTACATCGACGTCTATGCGGCGACCGAG, Δmini-INS TES: 5'-GGCGTCTACATC GACACCGAATCTGTCTATGCGGCGACC, Δ235-240: 5'-CTGCCGGTGCCAGAT GGAGACCTGACCCCG, Δmini-INS GS: 5'-GGCGTCTACATCGACGGTAGCGGTA GCGGTTCTGTCTATGCGGCGACC. All single point mutations were prepared through site-directed mutagenesis using the following primers: D238A:5'-CCGGACGAGAACGTCGCATACGACGGCGACCTG, D240A:5'-GAGAACGTCG ATTACGCAGGCGACCTGACCCCG, D242A:5'-GTCGATTACGACGGCGCACTG ACCCCGATCATC, F178E: 5'-AACAAGATGGAGGTGGCTTACCTGCGCA CCTTCGCG, V179K:5'-TAACAAGATGTTCAAGGCTTACCTGCGCACCTTCG, L182S:5'-GCTTACTCTCGCAACTTCGCGCGGATGGG, A186E: 5'-CACCTTCGC GAAGATGGGTCTGAAGG, F178E V179K: 5'-AACAAGATGGAGAAGG CTTACCTGCGCACCTTCGC, F178E V179K L182S: 5'-AACAAGATGGAGAAGG CTTACAGTCGCACCTTCGC. Mutated residues are underlined as shown. A 125 ng : 50 ng primer:template ratio was used for all point mutations. All mutants were verified through plasmid DNA sequencing (Genewiz).

2.2.3 Enzyme Preparation

Rp ProRS in BL21-CodonPlus (DE3)-RIL (Agilent Technologies) (provided by Stephen Cusack, EMBL Grenoble) glycerol stock was revived and induced for protein expression with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for 4 hours at 37° C. *Escherichia coli* (*Ec*) ProRS in BL21 (DE3) were overexpressed and purified as *Rp* ProRS. Mutant plasmids were transformed into BL21-CodonPlus (DE3)-RIL cells. Overexpression was carried out either at 37° C for 4 h with final concentration of 1 mM IPTG or 18 and 25° C for 14-16 h with final concentration of 0.1 mM IPTG. Lysis was performed by incubation with 10 mg/mL lysozyme in the presence of protease inhibitors followed by sonication (output power of 5 and 8 cycles of 15s ON 45s OFF pulse). Cell pellet was separated from cell lysate via centrifugation for 45 min at 15,000 rpm and both fractions were analyzed on 12% SDS-PAGE to check for protein in inclusion bodies.

2.2.4 Co-expression of *Rp* ProRS Mutants with chaperone cocktails.

Mutant plasmids were co-expressed with 5 different chaperone cocktails (Takara) as previously described (115, 116). Cells were pelleted, lyzed and analyzed for refolding on a 12% SDS-PAGE gel visualized by coomasie stain.

2.2.5 In-solution Refolding of Inclusion Bodies

Following cell lysis, cell pellets were dissolved in 8.0 M guanidine-HCl. Denatured protein solutions were injected into 10 kDa MW cut-off dialysis cassettes. Refolding was initiated by a 24-hour gradient dialysis from 8.0 M guanidine-HCl to 0.0

M guanidine-HCl at 4° C in storage buffers A-C (storage buffer A: 300 mM NaCl, 50 mM Na-Pi pH 7.5; storage buffer B: 300 mM NaCl, 50 mM Na-Pi pH 7.5, 1 mM B-mercaptoethanol; storage buffer C: 300 mM NaCl, 50 mL Na-Pi pH 7.5, 10% Tween).

2.2.6 Solid-supported Refolding of Inclusion Bodies

Cell pellets were resuspended in 6.0 M urea. Unfolded proteins were loaded onto a His-Select column twice. On-column refolding was initiated by passing a decreasing gradient of 6.0 M urea to storage buffers A-C. Bound proteins were eluted with 200 mM imidazole in storage buffers A-C. Loading of denatured fractions was performed at room temperature (~25° C), while all refolding elutions were conducted at 4° C. Fractions were analyzed on 12% SDS-PAGE and visualized with coomasie brilliant blue R-250 stain.

2.2.7 Size Exclusion Chromatography

Size exclusion chromatography (SEC) was performed with a GE FPLC AKTA Purifier system on a Superdex 75 16/60 column. Varying concentrations of *Rp* ProRS and *Ec* ProRS in 100 µL sample volume were injected and eluted at a flow rate of 0.5 mL/min with storage buffer A at 4°C. SEC profiles were collected at 260 nm and 280 nm.

2.3 Results and Discussion

2.3.1 Role of mini-INS in Dimerization

Class II aaRS are generally found as dimers or tetramers in nature. ProRS has been shown to exist as a functional dimer composed of 2 monomers oriented in an anti-parallel manner (2, 13). To investigate the role of mini-INS in self-dimerization, we determined the SEC profiles of wild-type *Rp* ProRS and compared it to *Ec* ProRS. The SEC profile indicates that *Rp* ProRS mainly exists as a dimer at concentrations significantly below the *Ec* ProRS self-dimerization K_D of 1.5 ± 0.3 μM (M. Qualley, unpublished) (Figure 2.1).

Under similar concentrations, a small amount of monomeric *Ec* ProRS is observed. This suggests *Rp* ProRS has a lower self-dimerization K_D compared to wild-type *Ec* ProRS. Furthermore, sequence analysis of organisms that encode mini-INS reveal shows conserved acidic residues D238, D240, and D242 located on the loop bridging the two α-helices in mini-INS (Figure 2.2). Additionally, semi-conserved non-polar residues are identified that appear to constitute a hydrophobic patch shielded by the mini-INS. When viewed in the context of the crystal structure, the acidic residues are shown to project toward the dimerization interface, thereby allowing salt bridge formation with basic residues displayed on the anticodon binding domain of the other monomer (Figure 2.3).

To investigate the contribution of these highly conserved acidic residues on the mini-INS to the self-dimerization of *Rp* ProRS, deletion and point mutations were carried out (Table 2.1). Only one out of the 7 deletion and truncation mutants prepared, was

successfully expressed and purified. SEC profile of *Rp* ProRS D240A mutant indicates the mutation has no significant effect on self-dimerization (Figure 2.4) of *Rp* ProRS. This suggests a single point mutation is not sufficient to shift the oligomeric state of *Rp* ProRS to monomer. An alternate investigative route may be necessary to fully ascertain the role of mini-INS in *Rp* ProRS strong self-dimerization like probing the dimerization under varying salt concentrations. If indeed the mini-INS contributes to dimerization through salt-bridge formation, this effect should be reduced under high salt conditions.

2.3.2 Refolding Of Mutant *Rp* ProRS

The biggest caveat of this project was the inability to recover properly folded proteins. Successful preparation of Δmini-INS *Rp* ProRS mutant would allow us to investigate its significance, not only to the protein structure, but also for tRNA binding, amino acid activation, and pre-transfer editing. However, almost all *Rp* ProRS constructs resulted in misfolded proteins and were targeted into inclusion bodies (Figure 2.5). Several high success-rate refolding strategies were performed, as described in the methods section, all of which unsuccessfully recovered any refolded protein.

In some cases, rapid overexpression of recombinant proteins may cause targeting to inclusion bodies. Several strategies may be used to regulate overexpression and recover fully folded and functional proteins such as regulating temperature, length of overexpression, and varying the concentration of the inducer. Overexpression was performed at 18, 25, and 37°C with varying concentrations of IPTG as described in the methods section. Similarly, mutant proteins aggregated into inclusion bodies.

Furthermore, mutants were co-expressed with various chaperone cocktails (1: pGKJE8:dnaK-dnaJ-grpE groES-groEL; 2L pGro7:groES-groEL; 3: pKJE7:dnaK-danJ-grpE; 4: pGTf2:groES-groEL-tig; 5: pTf16:tig) shown to promote folding of unstable proteins at a very high success rate (115, 116). Analysis of cell lysis fractions revealed mutants were still targeted into inclusion bodies (Figure 2.6). Altogether, our observations demonstrate the sensitivity of *Rp* ProRS to any mutational perturbations in the mini-INS, which suggests the mini-INS is critical for proper protein folding.

2.3.3 Role of mini-INS on Protein Folding

Crystal structure of *Rp* ProRS (15) reveals a highly conserved hydrophobic patch between the mini-INS and motif 2 of *Rp* ProRS (Figure 2.7). Semi-conserved polar residues L243, I246, and I247 on mini-INS (Figure 2.2) are found to participate in the formation of the hydrophobic patch (Figure 2.7). Furthermore, sequence analysis of the motif 2 on representative organisms from the three subsets of ProRSs (determined by their extra domains: C-terminal extension, INS domain, mini-INS) reveals residues forming the hydrophobic patch (F178, V179, L182, A186 on *Rp* ProRS mini-INS) show fair conservation among ProRS that encode for mini-INS only. In ProRSs that contain a C-terminal extension or INS domain, the residues in positions that align with 178, 179, 182, and 186 on *Rp* ProRS are found to be polar (Table 2.2).

To investigate the role of these residues to overall protein folding, we mutated the non-polar residues on motif 2 of *Rp* ProRS mini-INS to the equivalent polar residues on ProRS that encode for INS (Table 2.3). We decided to perform all mutations on Δmini-

INS *Rp* ProRS-GS linker construct because it contains a longer flexible linker compared to the other deletion mutants (Table 2.1) thus may better promote protein folding. Analysis of cell lysis fractions showed all proteins were misfolded and separated into the cell pellet. Taken altogether, our observations suggest that the mini-INS is critical for *Rp* ProRS structure by stabilizing a hydrophobic pocket posterior to the catalytic domain.

2.4 Conclusions

We have investigated the role of mini-INS in ProRS oligomerization and in maintaining structural integrity. SEC profile shows *Rp* ProRS has a self-dimerization K_D lower than *Ec* ProRS 1.5 ± 0.3 μM (M. Qualley, unpublished). It is unclear whether mini-ins contributes to the stronger dimerization as all but one prepared mutant proteins resulted in protein misfolding. ProRS sensitivity to mutation of the mini-INS demonstrates its importance in maintaining structure. Indeed, we concluded that the mini-INS promotes proper protein folding by stabilizing a hydrophobic patch posterior to the aminoacylation domain.

2.5 Acknowledgements

We would like to thank Dr. Thomas Magliery for providing us with the chaperone cocktails (Takara).

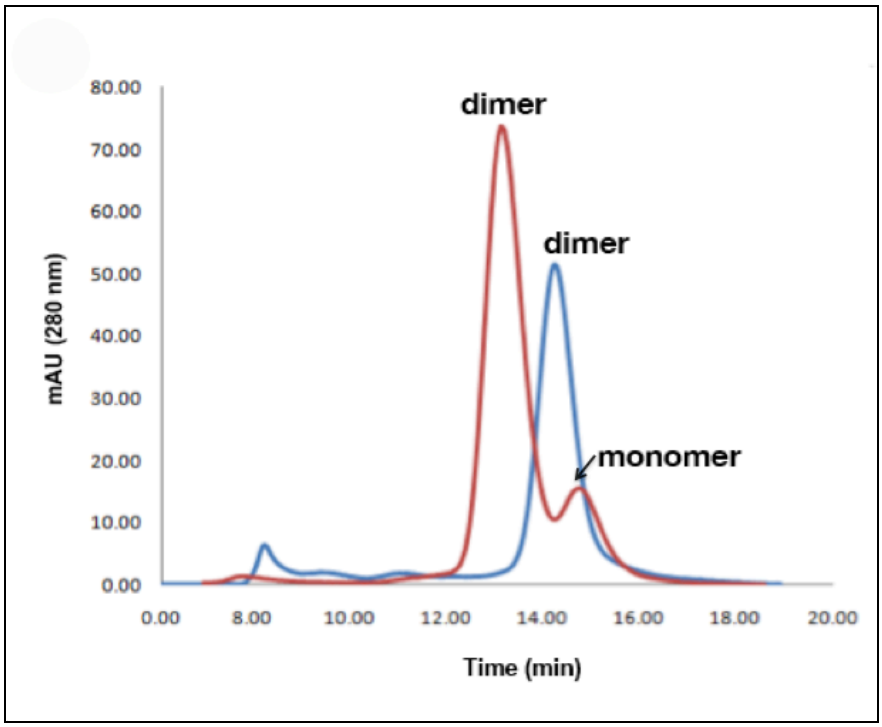


Figure 2.1. SEC Profile of *Ec* ProRS and *Rp* ProRS. SEC profiles of 1 μM *Ec* ProRS (red) and 1 μM *Rp* ProRS (blue) collected at 280 nm. Oligomeric form contained in each peak is indicated.

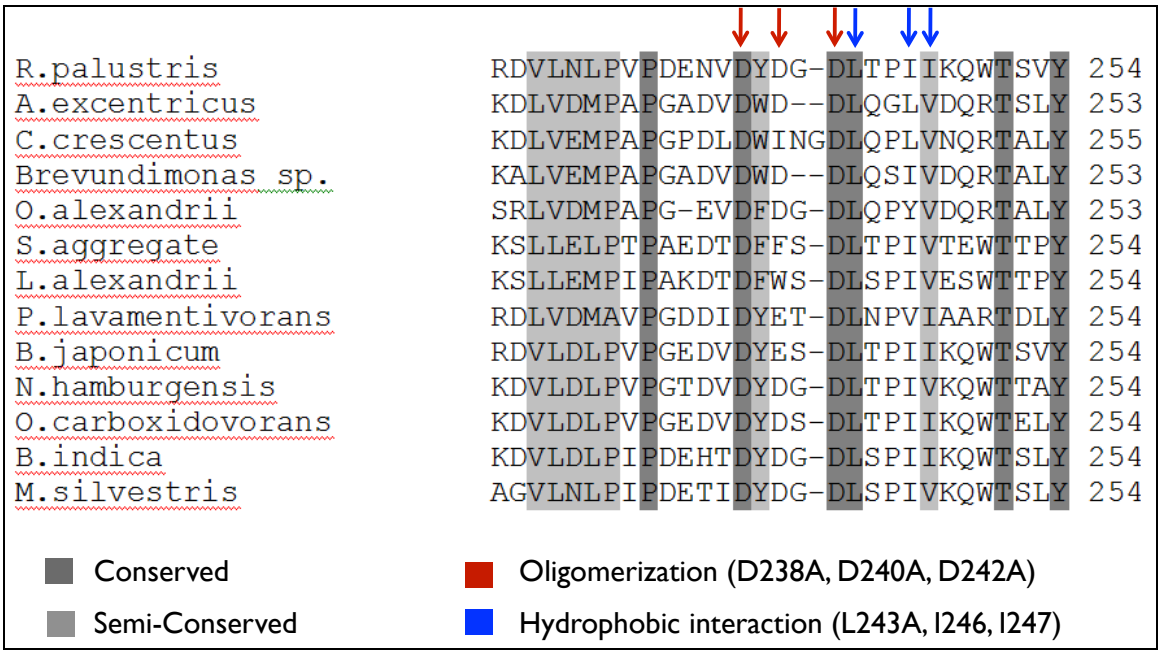


Figure 2.2. Sequence Alignment of ProRS mini-INS from various species. Semi-conserved residues are highlighted in light gray and strictly conserved residues are highlighted in dark gray. Residues that contribute to dimerization are indicated with red arrow while residues within the hydrophobic patch are indicated with blue arrows.

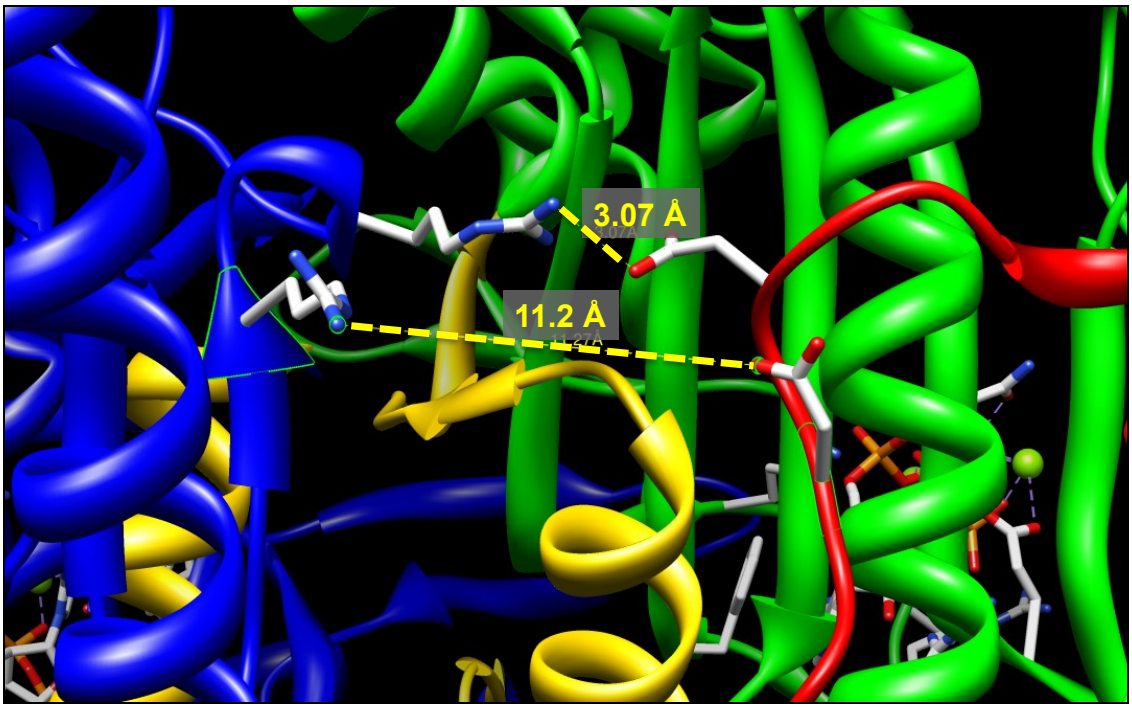


Figure 2.3.Dimerization interface of *Rp* ProRS. *Rp* ProRS monomers are shown in blue and green. Motif I is shown in yellow. Salt bridge interactions are shown in yellow lines with indicated distances.

Table 2.1. List of Deletion/Truncation Mutants carried out on *Rp* ProRS. Result of *Rp* ProRS mutant protein preparation reported as folded or misfolded. Only one (Rp ProRS D240A) out of the 7 is properly folded.

Mutation	**Result**
D238A	Protein misfolded
D240A	Properly folded
D242A	Protein misfolded
D240K	Protein misfolded
Δ 235-240	Protein misfolded
Δ mini-INS (Thr-Glu-Ser linker)	Protein misfolded
Δ mini-INS (Gly-Ser-Gly-Ser-Gly-Ser linker)	Protein misfolded

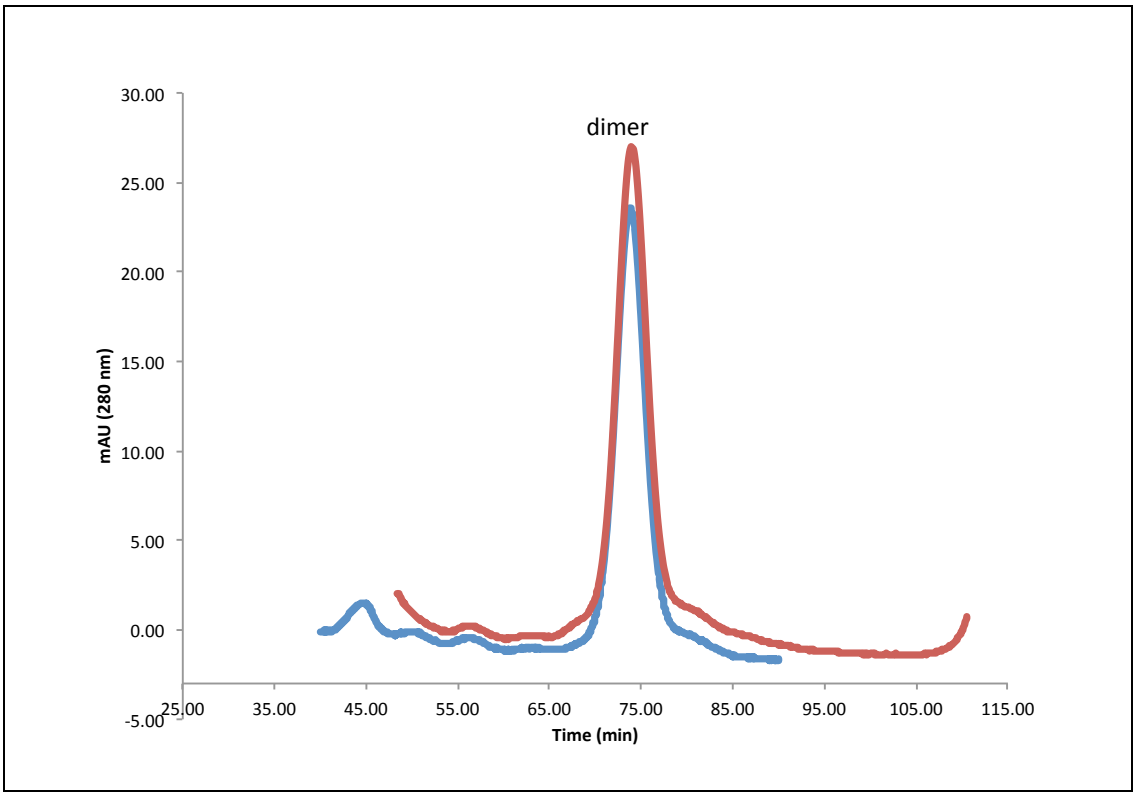


Figure 2.4. SEC Profile of WT *Rp* ProRS and *Rp* ProRS D240A. SEC profile of 1 μM WT *Rp* ProRS (blue) and 1 μM *Rp* ProRS D240A (red) collected at 280 nm.

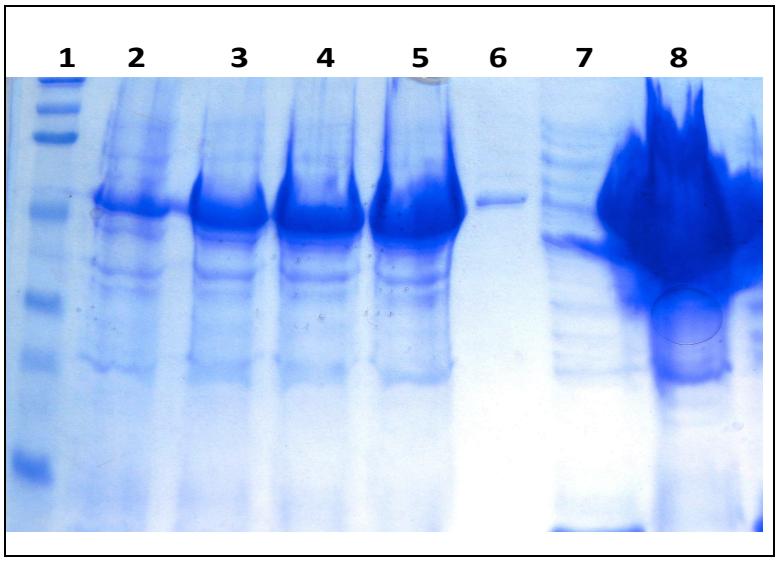


Figure 2.5. SDS-PAGE gel of *Rp* ProRS mutant construct. Here shown is a 12% SDS-PAGE of the overexpression profile of Δmini-INS GS linker over a 4 h time course. Time points collected prior to induction (lane 2), 1 h (lane 3), 2 h (lane 4), and 4 h (5). pET15b shows to be a leaky plasmid reflected in expression of mutant *Rp* ProRS prior to induction (lane 2). Analysis of cell lysis fractions is shown: cell lysate (lane 7) and pellet (lane 8). Samples were loaded with standard MW protein ladder (lane 1) and purified WT *Rp* ProRS as reference (lane 6).

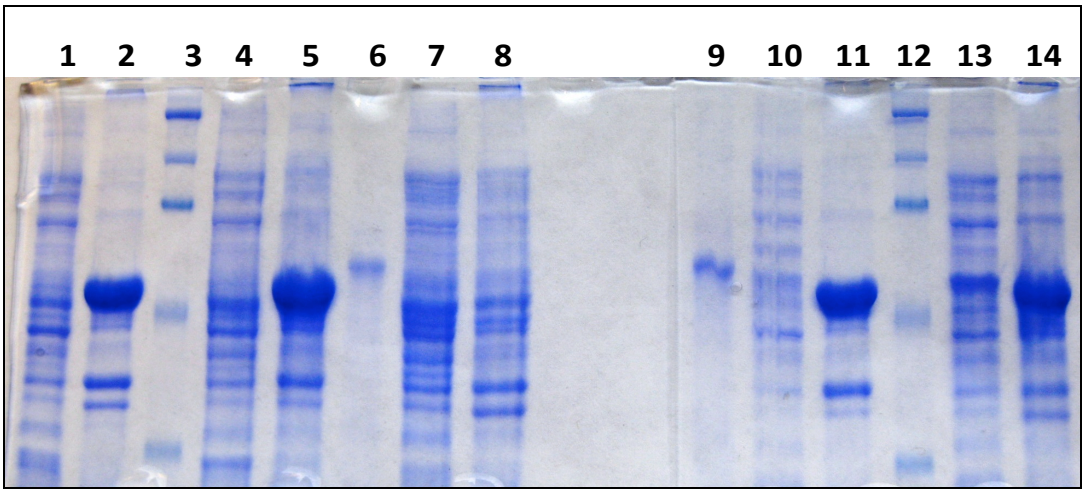


Figure 2.6. SDS-PAGE analysis of cell lysis fractions after chaperone cocktail co-expression. Shown is a 12% SDS-PAGE analysis of cell lysis fractions after co-expression of Δmini-INS GS linker with 5 different chaperone cocktails (Takara). Samples were loaded in pairs of lysate and pellet. GS-pGKJE8 lysate (lane 1) and pellet (lane 2); GS-pGRo7 lysate (lane 4) and pellet (lane 5); GS-pKJE7 lysate (lane 7) and pellet (lane 8); GS-pGTf2 lysate (lane 10) and pellet (lane 11); and GS-pTf16 lysate (lane 13) and pellet (lane 14). Samples were loaded with standard MW protein ladder (lanes 3 and 12) and purified WT *Rp* ProRS as reference (lane 9).

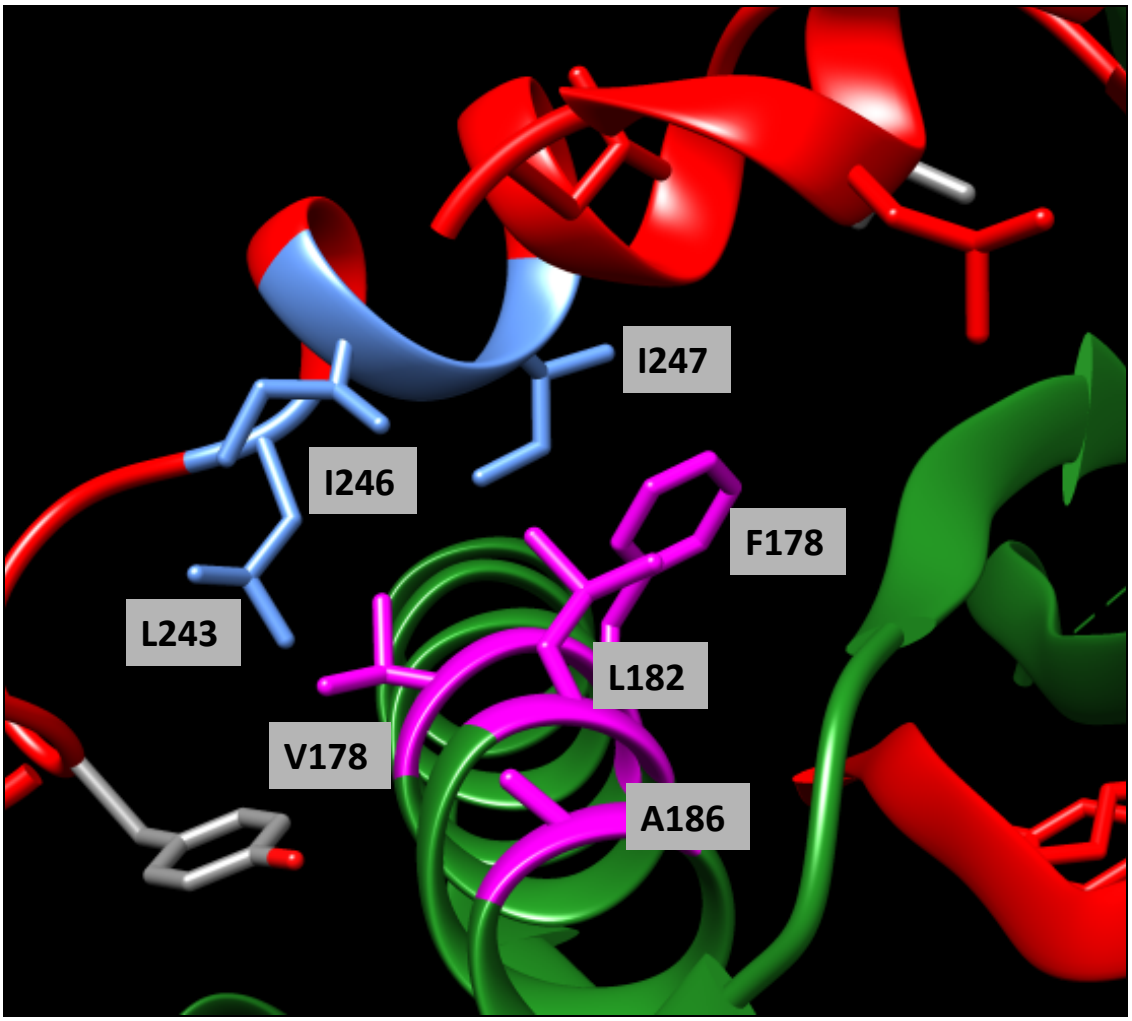


Figure 2.7. Crystal structure of *Rp* ProRS mini-INS. Structure reveals hydrophobic patch between mini-INS (red) and motif II (green) constituted by F178, V179, and L182 on motif II (residues highlighted in pink) and semi-conserved residues L234, I246, and I247 on mini-INS (highlighted in blue). Crystal structure of *Rp* ProRS pdb 2I4L is shown (15).

Table 2.2. Summary of Sequence Conservation on motif 2 among ProRS. Shown are conserved residues on mini-INS with residue number on *Rp* ProRS mini-INS indicated. Conserved residues on ProRSs that encode for INS and C-terminal extension equivalent to the indicated residue positions on *Rp* ProRS mini-INS are also shown.

ProRS subtype	**Residue # on *Rp* ProRS**			
	178	179	182	186
Mini-INS	F	V	L	A
INS	E	K	S	E
C-terminal extension	E	I	E	R

Table 2.3. List of point mutations carried out in *Rp* ProRS Δ mini-INS (GS-linker). Results of enzyme preparation of mutations performed on *Rp* ProRS are shown as either misfolded or folded. All mutants resulted in misfolded proteins.

Mutation	**Result**
F178E	Protein misfolded
V179K	Protein misfolded
A186E	Protein misfolded
F178E V179K	Protein misfolded
F178E V179K L182S A186E	Protein misfolded

Chapter 3 : Expanded function of *trans*-editing domains to non-protein amino acids prevents global mistranslation

3.1 Abstract

Aminoacyl-tRNA synthetases (aaRS) belong to an ancient family of enzymes that activate specific amino acids and attach them to cognate tRNAs for use in protein synthesis. aaRSs often misactivate isosteric proteinogenic amino acids due to similarities in size and volume of the side chains, potentially resulting in translational errors that threaten cell survival. Non-protein amino acids are abundant in the cell and potentially pose an even more deleterious impact on translation. The non-protein amino acid aminobutyrate (Abu) is a metabolite involved in various cellular processes. Structurally, Abu is intermediate in size between Ala and Val. Importantly, Abu is recognized by several aaRS such as ProRS, ValRS, and IleRS. Fidelity in Val and Ile codon translation is maintained by a *cis*-editing domain known as connective polypeptide 1 (CP1), which is a distinct domain inserted into the active site of ValRS and IleRS and hydrolyzes misacylated tRNAs including Abu-tRNA. In the case of ProRS which misactivates Ala and Cys in addition to Abu, a *cis*-editing domain (INS) is present in most Bacterial enzymes, which some Bacteria lack this domain but instead encode a single domain *trans*-editing homolog with the same activity. A distinct single-domain INS homolog, YbaK clears Cys-tRNA^{Pro} in *trans*. Although these mechanisms clear misactivated

protein amino acids, how Abu-tRNA^{Pro} is prevented from misincorporation is unclear. Interestingly, many different combinations of INS-like *cis*- and *trans*-editing domains exist in Bacteria. For example, the metabolically versatile bacterium *Rhodopseudomonas palustris* (*Rp*) encodes a ProRS containing a catalytically inactive, truncated INS domain (mini-INS), in addition to two distinct, INS homologs YbaK and ProXp-x. The function of the latter is unknown. Here, we show *Rp* ProXp-x does not deacylate Ala- and Cys-tRNA^{Pro} *in vitro*. Comparison of known crystal structures reveals that the putative substrate-binding pocket of ProXp-x is larger than that of INS, which suggests substrates larger than Ala are preferred. Indeed, ProXp-x weakly deacylates tRNAs charged with Val and Ile, but robustl edits Abu mischarged onto tRNA^{Pro}, tRNA^{Val}, and tRNA^{Ile}. Semi-promiscuous editing may offer advantages to cells and our data suggest ProXp-x, now renamed ProXp-abu, may act as a general Abu-tRNA deacylase. *Rp* ProRS specificity for activation of Pro over Abu is only about 1,000:1, which strongly suggests that editing of Abu-tRNA^{Pro} may be required *in vivo*. Taken together, these data suggest that Abu-tRNA editing by the *trans*-editing factor ProXp-abu is likely to be a critical checkpoint to ensure high fidelity in codon translation.

3.2 Introduction

Aminoacyl tRNA synthetases (aaRS) catalyze the two-step addition of the correct amino acid to cognate tRNAs. The first step involves activation of the amino acid via hydrolysis of ATP forming an aminoacyl adenylate intermediate. This is followed by an esterification step in which the activated amino acid is transferred to the 3'-end of the

cognate tRNA (1). Due to similarities in volume and size of related amino acids, aaRS are inherently error-prone (117). When left uncorrected, errors in translation may lead to mutations in the proteome, and in some cases, cell death. About half of the 20 known aaRS found in all three Domains of life misactivate isosteric, noncognate amino acids (2). For example, ValRS is unable to distinguish cognate Val from Thr, Ile, and the non-protein amino acid 2-aminobutyrate (Abu) (3, 31, 118). To overcome this limitation, aaRS have evolved several proofreading mechanisms. "Pre-transfer" editing is the hydrolysis of misactivated non-cognate amino acids into ATP and free amino acid (2), while "Post-transfer" editing employs an independent editing domain to clear misacylated tRNAs (2, 3).

In the case of ProRS, standard amino acids Ala and Cys are misactivated in addition to cognate Pro. To ensure fidelity in Pro codon translation, which is critical for protein structure and folding, organisms have evolved multiple proofreading strategies including ProRS pre- and post-transfer editing activity (14, 45, 64), and use of *trans*-editing factors (22, 79, 119). Furthermore, various combinations of these proofreading strategies are employed in a species-specific manner. For example, in some bacterial systems such as in *Escherichia coli (Ec*), a triple-sieve editing mechanism is employed (79). The first sieve consists of the ProRS catalytic core, which discriminates against non-cognate amino acids based on size. The second sieve is a *cis*-editing domain called INS, which is inserted between the class II conserved motifs 2 and 3, and hydrolytically clears misacylated Ala-tRNA^{Pro} but not Cys-tRNA^{Pro} (20, 64). The latter is cleared by the INS homolog YbaK, which represents the third sieve and functions in *trans*, via a thiol-

cyclization mechanism (21, 81, 119). However, not all species encode for an INS. Architectural diversity of ProRS across all three Domains of life is attributed to the expression of various extra domains appended to the universally conserved class IIa anticodon binding domain and the class II catalytic core (15). ProRS are classified as either Eukaryotic- and Archaeon-like or Prokaryotic-like (18, 120). The former is characterized by the presence of an ~80 residue C-terminal extension, demonstrated to be of structural rather than catalytic significance. Moreover, in some lower eukarya, an N-terminal domain homologous to INS is present in addition to the C-terminal extension (14, 17). Prokaryotic-like ProRS is generally characterized by the presence of an INS domain. Although, some bacterial ProRS (~22%) do not contain a full-length INS domain, a subset (~10%) encode for a severely truncated mini-INS (22). Although the mini-INS has been proposed to lack catalytic activity (15, 22), its relevance to the overall structure and function of the ProRS has not been investigated.

In addition to YbaK, four other distinct single-domain INS homologs have been identified through bioinformatics and are thus part of the INS Superfamily (14, 21, 22). This family of proteins is widely distributed in Bacteria and found throughout all three Domains of life in various species-specific combinations (22). Substrate specificities of three out of the six INS Superfamily members have been previously reported. INS and ProXp-ala recognize Ala-tRNA^{Pro} (14, 22), whereas YbaK is specific for Cys-tRNA (81, 119). Approximately 0.8% of Bacteria, including the gram-negative bacterium *Rhodopseudomonas palustris* (*Rp*), encode a ProRS with a mini-INS domain in addition to two *trans* editing factors YbaK and ProXp-x. The function of the latter remains

unknown. In this study, we substrate specificity and molecular basis for substrate discrimination of the *trans*-editing factor *Rp* ProXp-x. Our results demonstrate that the substrate specificities of *trans*-editing factors are not limited to standard amino acids, but encompass non-proteinogen amino acids, which may pose an even greater threat to the proteome.

3.3 Experimental Procedures

3.3.1 Materials

All amino acids and chemicals were purchased from Sigma unless otherwise noted. [α-^{32}P]-PP_i and [α-^{32}P]-ATP were from Amersham Biosciences and [^{35}S]-cysteine (1075 Ci/mmol) was from PerkinElmer Life Sciences.

3.3.2 Enzyme preparation

Rp ybaK, and *proX* (NCBI GenBank accession numbers: WP_011159832, and CAE29641, respectively) genes encoding YbaK, and ProXp-x, respectively, were PCR-amplified from *Rp* CGA009 genomic DNA (ATCC) using primers that included flanking restriction sites for BamHI and NdeI endonucleases (New England Biolabs). Each gene was subcloned into pCR2.1-TOPO (Invitrogen), and later cloned into pET15b (Novagen). Protein expression was carried out in *Ec* BL21-CodonPlus (DE3) RIL cells. *Rp* ProRS and ProXp-x overexpression was induced with 0.1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for ~16 h at room temperature. *Rp* YbaK overexpression was induced with 1 mM IPTG for 4 h at 37 °C. The N-terminally histidine-tagged

proteins were purified using HIS-select® nickel resin (Sigma-Aldrich). Wild-type (WT) *Ec* ProRS (21), and *Ec* tRNA nucleotidyltransferase (121) were prepared as previously described. Concentrations of *Rp* YbaK, *Rp* ProXp-x, and *Ec* tRNA nucleotidyltransferase were determined by the Bradford assay (122). The concentrations of *Rp* and *Ec* ProRS were determined by active site titration (123).

3.3.3 Preparation of tRNAs and aminoacyl-tRNA substrates

WT *Ec* $tRNA^{Pro}$ was prepared by *in vitro* transcription using T7 RNA polymerase as previously described (64). Aminoacylation of tRNA substrates with Pro, Ala, and Abu was carried out in buffer A (50 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), pH 7.5, 4 mM ATP, 25 mM $MgCl_2$ 0.1 mg/ml bovine serum albumin (BSA), 20 mM β-mercaptoethanol, 20 mM KCl) by incubating 10 μM of *Ec* K279A ProRS, 10 μM of $tRNA^{Pro}$, amino acid (10 mM Pro, 500 mM Ala, 200 mM Abu) and 0.029 mg/ml pyrophosphatase (Roche) for ~2 h at room temperature. Aminoacylation of tRNA substrates with Cys was carried out in buffer B (20 mM tris(hydroxymethyl)aminomethane chloride (TRIS-Cl), pH 8.0, 4 mM ATP, 10 mM $MgCl_2$ 0.1 mg/ml BSA, 25 mM dithiothreitol (DTT), 20 mM KCl) by incubating 10 μM of WT *Ec* ProRS, 10 μM of $tRNA^{Pro}$, 150 mM cysteine and 0.029 mg/ml pyrophosphatase for ~4 h at 37° C. Prior to aminoacylation, tRNA substrates were 3′-[^{32}P]-labeled using *Ec* tRNA nucleotidyltransferase as previously described (124). Following aminoacylation, aminoacyl-tRNAs (aa-tRNA) were phenol-chloroform

extracted followed by ethanol precipitation. Substrates for deacylation assays were stored at -80 °C for future use.

3.3.4. ATP:PP_i exchange assays

ATP:PP_i exchange assays assays were carried out using published conditions (8) with the following amino acid concentrations 0.025-5 mM Pro, 0.1-15 mM Cys, 50-1000 mM Ala, and 50-1000 mM Abu. *Rp* ProRS (10 nM) was used for Pro activation and 250 nM enzyme was used for Cys, Ala and Abu activation. SigmaPlot (Systat Software, San Jose, CA) was used to generate graphs for all assays. Error is reported as standard deviation (S.D.) of triplicate data.

3.3.5 Aminoacylation assays

Mischarging of $tRNA^{Pro}$ by *Rp* ProRS was carried out in buffer A using 0.5 μM ProRS, 6 μM *Ec* $tRNA^{Pro}$, 0.1-1.5 μM *Rp* ProXp-x, 1.5 μM *Rp* K45A ProXp-x and 500 mM Abu.

3.3.6. Pre-transfer editing assays

ATP hydrolysis assays were performed as previously described (125). The *Rp* ProRS concentration was 0.5 μM and the amino acid concentrations were as follows: 3 mM Pro, 3 mM Cys, and 500 mM Ala.

3.3.7 Deacylation assays

Cys-, Ala-, and Abu-tRNA deacylation reactions were performed using published conditions at 37 °C and 25 °C, respectively. Briefly, reactions containing ~0.7 μM aminoacyl-tRNA and buffer C (300 mM KPO_4 (pH 7), 0.2 mg/ml BSA, and 9.6 mM $MgCl_2$) were initiated by addition of enzyme. For Cys-$tRNA^{Pro}$ deacylation, 0.5 μM *Rp* YbaK, *Rp* ProRS, or *Rp* ProXp-x were used. For Ala-tRNA deacylation, 0.5 μM *Rp* ProRS, *Rp* ProXp-x, or 1 μM *Ec* ProRS were used. At the indicated time points, reaction aliquots (2 μl) were quenched into 4 μl of a solution containing 0.4 U/μl P1 nuclease in 200 mM NaOAc (pH 5). Deacylation levels were monitored using polyethyleneimine-cellulose TLC and analyzed as previously described (124).

3.3.8. Molecular modeling of ProXp-x

The X-ray crystal structure of *Thermus thermophilus* ProXp-x bound to alanyl sulfamoyl adenylate (PDB = 2Z0K, Murayama et al, unpublished) was used as the starting structure. Due to the fact that the alanyl moiety is attached to the 5'-position of the adenosine, this structure is likely not to mimic the correct binding mode of misacylated tRNA; however, the location of the active site pocket matches that of the homologous ProRS INS domain (15). Since there may be differences in the positions of active site residues from that of the crystal structure, a molecular dynamics (MD) approach was utilized to represent flexibility during the docking simulations. MD simulations were carried out by first manually removing all water and ligand atoms from the PDB file of the crystal structure. Using the tleap module of AmberTools12 (126), the

positively charged *Tt*ProXp-x was neutralized with Cl^- counterions and solvated in an octahedral box of TIP3P (127) water molecules such that no less than 8.0 Å separated the edge of the box and any solute atom. All simulations were carried out with an 8.0 Å cutoff for nonbonded interactions and utilized forcefields based on the ff99SB library (128, 129). All MD simulations were performed using Amber12 (126). Energy minimization was accomplished via a two-stage approach in which the solute atoms were fixed in the first 2000 step stage, followed by a stage consisting of 5000 steps. The particle mesh Ewald (PME) implementation of constant volume periodic boundaries (130) and 1000 steps of steepest-descent minimization was used for each stage of minimization. Additionally, two stages of pre-production simulation ensured thermal equilibration of the system. In the first stage, the system was slowly heated from 0-300 K over 100 ps using a time step of 2 fs with weak positional restraints (10 kcal/mol*$Å^2$) on the solute atoms. The SHAKE algorithm constrained all hydrogen-involved bonds (131) and a collision frequency of 1.0 ps^{-1} for Langevin dynamics was used for temperature control. For the second stage of equilibration and all production simulations, the PME implementation was used for constant pressure periodic boundary conditions and the solute atoms were left unrestrained for a total of 900 ps. Finally, production simulations were carried out for 15 ns using Amber12 with the conditions described for the second stage of pre-production equilibration. To eliminate synchronization artifacts caused by the use of Langevin dynamics, a random seed was used to start each equilibration and production simulation (132, 133). Trajectories were analyzed using the ptraj module of

AmberTools12 (126) and visualized with Visual Molecular Dynamics software (134) and PyMol (135, 136).

From the 15 ns MD trajectory, 25 snapshots were extracted at equidistant time intervals and each was used as an independent receptor for docking simulations. Ligands were prepared in Avogadro (137) with Abu, Ala, Cys or Pro attached to the 3'O of the three 3' terminal tRNA residues, CCA. Receptor and ligand molecules were prepared for docking using AutoDockTools (138), with the receptor residues within 5 Å of the active site left flexible. For each ligand, AutoDock Vina (139) was used for docking to each receptor snapshot. The resultant poses were clustered using a cutoff RMSD of 2.5 Å and one conformer representative of each of the top four clusters was manually chosen for further analysis. Each pose was then used as the starting structure for a subsequent 15 ns MD simulation (minimization, equilibration and production) using the parameters described above. An estimation of the relative binding free energies of each conformer was then calculated using the Molecular Mechanic-Poisson Boltzmann Surface Area (MM-PBSA) algorithm of AmberTools12 (126) for the 15 ns of production simulation only. The conformer exhibiting the lowest binding free energy was chosen as the most likely binding mode for each ligand.

3.4 Results

3.4.1. Amino Acid Specificity of *Rp* ProRS

Most bacterial ProRSs misactivate Ala and Cys and therefore require *cis*- or *trans*-editing domains to prevent mistranslation. To establish whether this was the case

for *Rp* ProRS, ATP:PP_i exchange assays were conducted to measure activation of Pro, Cys, and Ala. As shown in Table 3.1, the K_M for Ala (1020 mM) is significantly larger compared to cognate Pro with an ~16-fold lower turnover rate. Therefore, the overall Ala activation efficiency is ~10^5 fold less than cognate Pro. This observation is consistent with previous work on *Caulobacter crescentus* (*Cc*) ProRS, which also lacks an INS domain (22). As previously reported, the k_M for Cys is only slightly lower than that of cognate Pro (140), with an ~320 fold reduced k_{cat}. Thus, the overall activation efficiency is only ~200 fold lower than that of Pro. Based on the activation efficiencies for Cys and Ala, only the Cys discrimination factor is below the 3000 threshold where editing is likely to be required (141). However, this discrimination factor does not account for the relative concentration of Ala and Pro in *Rp*. Under certain conditions where Ala is present in vast excess over Pro, the "effective discrimination factor" may be significantly lower.

3.4.2. Pre-transfer Editing by *Rp* ProRS

We investigated the pre-transfer editing activity of *Rp* ProRS by performing ATP hydrolysis assays in the absence of tRNA (22, 45). As expected, no accumulation of AMP is observed with cognate Pro (Figure 3.1). Only marginal accumulation of AMP is observed with Cys. In contrast, AMP formation is stimulated in the presence of Ala reflecting robust pre-transfer editing activity of *Rp* ProRS for Ala.

3.4.3. Post-transfer editing by *Rp* ProRS, ProXp-x, and YbaK

We next investigated post-transfer editing activities of *Rp* ProRS, YbaK, and ProXp-x using deacylation assays to monitor the disappearance of aa-tRNA over time. *Rp* ProRS does not show any deacylase activity due to the lack of a full length INS domain (Figure 3.2.A-C). As expected, *Rp* YbaK shows robust deacylation activity against Cys-tRNA^{Pro} (Figure 3.2.C), in agreement with previous work on YbaK isolated from other species (21, 22, 81, 119). Moreover, ProXp-x showed no deacylation activity for Ala-, Cys-, and Pro-tRNA^{Pro} (Figure 3.2.A-C).

3.4.4. Sequence analysis of ProXp-x and INS catalytic sites

To investigate the lack of hydrolytic activity in ProXp-x, we compared the amino acid sequences of INS and ProXp-x. Previous studies show conserved residues I263, L266, T277, and H366 constitute the INS active site pocket in *Ec* ProRS (20). Comparison of INS and ProXp-x sequences reveal most of the bulky residues in INS are present as Ala or a smaller amino acid in ProXp-x (Figure 3.3). This suggests the putative substrate binding pocket in ProXp-x is larger compared to INS and likely accepts substrates larger than Ala.

3.4.5. Substrate Specificity of ProXp-x

To investigate the substrate specificity of ProXp-x, we tested its deacylase activity against amino acids larger than Ala (Figure 3.4). ProXp-x showed minimal deacylase activity towards similar size non-polar Val-tRNA^{Val} and Ile-tRNA^{Pro} but not larger Leu-

$tRNA^{Pro}$. Furthermore, no deacylation activity was shown against slightly polar substrates Ser-$tRNA^{Pro}$ and Thr-$tRNA^{Val}$. Strikingly, robust deacylation activity against Abu-$tRNA^{Pro}$ is observed. Altogether, our data suggests Abu is the preferred substrate for ProXp-x. We determined the activation parameters of Rp ProRS for Abu and found that the K_M for Abu is ~10^3 fold higher than cognate Pro (Table 3.1). Nevertheless, its turnover rate is ~2 fold higher relative to Pro, resulting in a ~10^3-fold lower overall activation efficiency. This is lower than the 3000 threshold where editing is likely required (141).

3.4.6. Molecular Dynamics (MD) simulated docking model of Abu-CCA in ProXp-x in putative substrate binding site

To gain further insight into the alternative substrate specificity of ProXp-x, we employed molecular dynamics to generate a model of Abu-CCA docked in the putative active site of ProXp-x and compared it to the docking model of Ala-CCA bound in *Ef* ProRS INS domain (Figure 3.5) (20). We note several differences and similarities. We observe the universally conserved Lys, critical for activity in INS (16), is maintained in the same orientation in both structures and both interact with the phosphate group on A76. However, the Abu substrate side chain is docked (in ProXp-x) in the opposite orientation (~180°) compared to Ala docked in INS. In ProXp-x, the Abu side-chain projects toward a pocket constituted by A129, L136, I48, and the methyl side chain on T31. Of these residues, only I48 is maintained between INS and ProXp-x active sites (Figure 3.3). Mesh representation of the INS and ProXp-x active sites reveal a larger

active site pocket for ProXp-x compared to INS (Figure 3.6). In INS, the amino group of the Ala substrate is proposed to interact with H366 for proper orientation of the substrate in the binding pocket (20). In contrast, our docking model reveals the amino group of Abu substrate is stabilized by 3 hydrogen bonds formed with the backbone carbonyl of T31, A34, and G102.

3.4.7 Docking simulations reveal an energetic bias toward Abu binding

In order to better understand the structural mechanisms underlying substrate specificity, docking simulations were performed to compare relative binding free energies of each of the substrates to the active site. Model ligands were designed with the 3′ terminal CCA residues acylated with Abu, Ala, Cys or Pro at the 3′O. To sufficiently sample the conformation space of ProXp-x prior to docking, a 15 ns MD simulation was performed. Twenty-five equally-spaced snapshots were extracted and used as starting structures for docking simulations. Active site residues were assigned as flexible during docking simulations while all remaining ProXp-x residues were kept static. From each set of 500 docked poses for each ligand, one representative conformer of the four most populated clusters were chosen for further analysis. A subsequent 15 ns simulation was used to calculate relative binding free energies using MM-PBSA and the best pose was chosen as the most likely bound structure. In agreement with our *in vitro* observations, Abu is the preferred substrate by ~3.2 kcal/mol, ~20.1 kcal/mol and ~7.5 kcal/mol for Ala, Cys and Pro, respectively (Figure 3.7). Even for the smallest difference of 3.2 kcal/mol between Abu and Ala, this corresponds to a population difference of nearly

5000:1 at 300 K. To further investigate the structural basis for these energetic differences, the docked structures were analyzed. Visualization of the shape and size of the active site pocket for each docking model, it is clear that the flexible active site residues are significantly perturbed upon CCA-Cys binding, possibly explaining the poor binding free energy (Figure 3.8). Despite the binding free energy difference between Abu and Ala suggested by the MM-PBSA calculations ($\Delta\Delta G$ = 3.2 kcal/mol) positioning of Abu and Ala is nearly identical. The additional methyl group on Abu appears to cause a slight restructuring of the active site residues and also more completely fills the volume of the pocket.

3.4.8. Effect of ProXp-x on Abu misacylation by *Rp* ProRS

To simulate cellular conditions where ProRS is constantly present, we investigate the functionality of ProXp-x in the presence of ProRS. Aminoacylation experiments were performed in the absence and presence of varying amounts of ProXp-x. Decreasing Abu misacylation is observed as the concentration of ProXp-x increases (Figure 3.9). This ProXp-x concentration-dependent behavior demonstrates the ability of ProXp-x to prevent accumulation of Abu-tRNA^{Pro} during protein synthesis. Moreover, mutating the highly conserved K45 in *Rp* ProXp-abu to Ala significantly reduced deacylase activity. This supports the decrease in misacylation is attributed to ProXp-x deacylase activity.

3.5. Discussion

There are several proposed outcomes in the event that an aaRS loses an editing function: the organism is forced to evolve alternative editing mechanisms to compensate for the loss of editing or the organism will evolve higher substrate specificity, therefore eliminating the necessity of an editing site (14). Alternatively loss of editing may result in statistical randomization of the proteome, which may benefit the organism (112). For example, the *Cc* ProRS system lacks a full-length INS and appears to have evolved an alternate editing strategy through the use of two distinct *trans*-edtiting factors ProXp-ala and YbaK to clear Ala-and Cys-$tRNA^{Pro}$ respectively (22). In the case of *Rp* ProRS, it appears the aminoacylation core has evolved to increase stringency in discriminating against Ala, as reflected by the elevated K_M for Ala coupled with robust pre-transfer editing. This implies Ala post-transfer editing may not be required in *Rp*. To clear Cys-$tRNA^{Pro}$, *Rp* encodes for YbaK to proofread Cys misacylations.

Misincorporation of non-protein amino acids has been studied over the past decades but their impact and relevance are often overlooked. Non-protein amino acids have been shown to threaten protein amino acid function due to their inherent structural relatedness. Non-protein amino acids have been shown to hijack translational machinery and have been implicated in several physiological dysfunction and diseases (26, 28). In *Rp*, kinetic data for amino acid activation suggests editing of the non-protein amino acid Abu may be required *in vivo*. Structurally, Abu may be considered as the intermediate molecule between Ala and Val, which only differs from both by one methyl group. Consequently, Abu is misactivated by several synthetases that recognize Ala or Val such

as ProRS, ValRS, and IleRS (3, 29). Abu is a metabolite heavily involved in various amino acid catabolic or metabolic pathways as a side product or intermediate, hence is generally present in the cell (30, 142–145). It is likely under certain conditions Abu levels are elevated; thus, a housekeeping strategy may be required. Class I aaRS ValRS and IleRS encode for a *cis*-editing domain (CP1), which has been shown to possess robust activity against misacylated Abu-tRNA$^{Val/Ile}$. However, in the case of an editing defective aaRS, the system may be vulnerable to Abu misincorporation as has been demonstrated in previous work (30). Interestingly, efficient Abu editing has not been demonstrated for ProRS.

In this study, we show that ProXp-x recognizes and robustly deacylates Abu-tRNA. Therefore, we now rename this factor ProXp-abu. Comparison of the INS and ProXp-abu active sites reveal a significantly larger active site for ProXp-abu; thus, is able to accommodate Abu. Additionally, MD simulated docking model of Abu-CCA in the ProXp-abu active site reveal the catalytic pocket is highly hydrophobic where residues surrounding Abu substrate are mostly non-polar. Deacylation assays show ProXp-abu also accepts Val- and Ile-tRNA substrates albeit with weak hydrolytic activity. Cys, Ser, and Thr are rejected due to the polar nature of their sidechains. *In vitro* substrate specificity of ProXp-abu is supported by the observed trend in calculated binding free energies of the substrates in the ProXp-abu binding pocket.

Interestingly, the Abu sidechain is oriented 180° compared to Ala docked in INS. This observation suggests a different editing mechanism may be employed than for INS. In the case of ProXp-abu, no catalytic water was found in the active site. More rigorous

computational analysis coupled with *in vitro* assays is required to determine the exact mechanism of deacylation of Abu-tRNA by ProXp-x.

Deacylation assays show ProXp-abu has similar activity for Abu on $tRNA^{Pro}$ and $tRNA^{Val}$. Aceeptor stem features show they maintain the same discriminator base (A73) but differ in their 1^{st} and 2^{nd} base pairs. This observation suggests ProXp-x may have relaxed tRNA specificity and may act as a general deacylase. This feature is advantageous to the cell as misincorporation of a non-protein amino acid is undesirable. Biophysical experiments are underway to investigate possible protein and RNA interacting partners of ProX *in vivo* that may play a role in its recruitment and function.

Rp is a metabolically versatile organism capable of several different metabolic modes; aerobic chemoheterotrophic, anaerobic phototrophic, carbon fixing, nitrogen fixing, etc (146, 147). Due to the ability of this organism to tolerate various living conditions, it is likely ProXp-abu is a requirement for one or more metabolic states where a possible increase in Abu cellular concentration occurs. The baseline proteome of *Rp* grown in its various metabolic modes was determined (147). Interestingly, ProXp-abu was detected only under nitrogen fixing conditions. Phylogenetics reveals ProXp-abu is mainly present in alpha-proteobacteria with some representation in beta and gamma proteobacteria and a small subset of archaea (22). Out of over a thousand bacterial species whose genomes have been completely sequenced, ~5% of the organisms encoded for ProXp-abu/YbaK/INS while ~0.8% encode for ProXp-abu/YbaK/mini-INS. Interestingly, 64% of the organisms encoding for ProXp-abu/YbaK/mini-INS are found to be nitrogen fixing bacteria. It is likely there is a link between the nitrogen fixing

metabolic mode of *Rp* and the cellular levels of Abu. Generation of a null strain is currently underway and will be used to focus on the role of ProXp-abu *in vivo*.

3.6 Acknowledgment

We would like to thank Dr. William Cantara for performing all computational modeling and calculations of ProXp-abu and for their valuable advice on interpretation of computational data.

Table 3.1. Steady-state kinetic parameters for amino acid activation by *Rp* ProRS. All values are the average of three trials with the ± standard deviation indicated.

Amino Acid	k_{cat} (s-1)	K_M (mM)	k_{cat}/K_M	Discrimination factor[a]
Pro	8 ± 0.1	0.271 ± 0.07	30	1
Cys[b]	0.025	0.17	0.15	197
Ala	36 ± 0.2	1020 ± 240	4.7×10^{-4}	6.2×10^{5}
Abu	14 ± 0.3	558 ± 160	0.025	1.1×10^{3}

[a] k_{cat}/K_M (Pro/Cys) or (Pro/Ala)

[b] Data from (140)

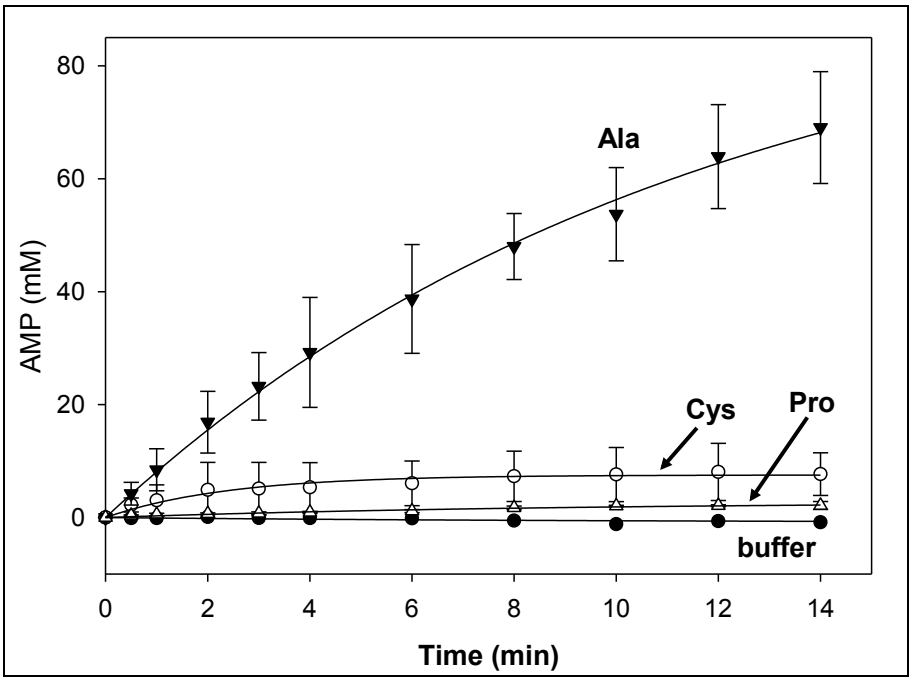


Figure 3.1. Pre-transfer editing activities of *Rp* ProRS. Pre-transfer editing time course showing formation of AMP in the presence of proline (△), alanine (▼), and cysteine (○). Reaction conditions were as described in "Experimental Procedures".

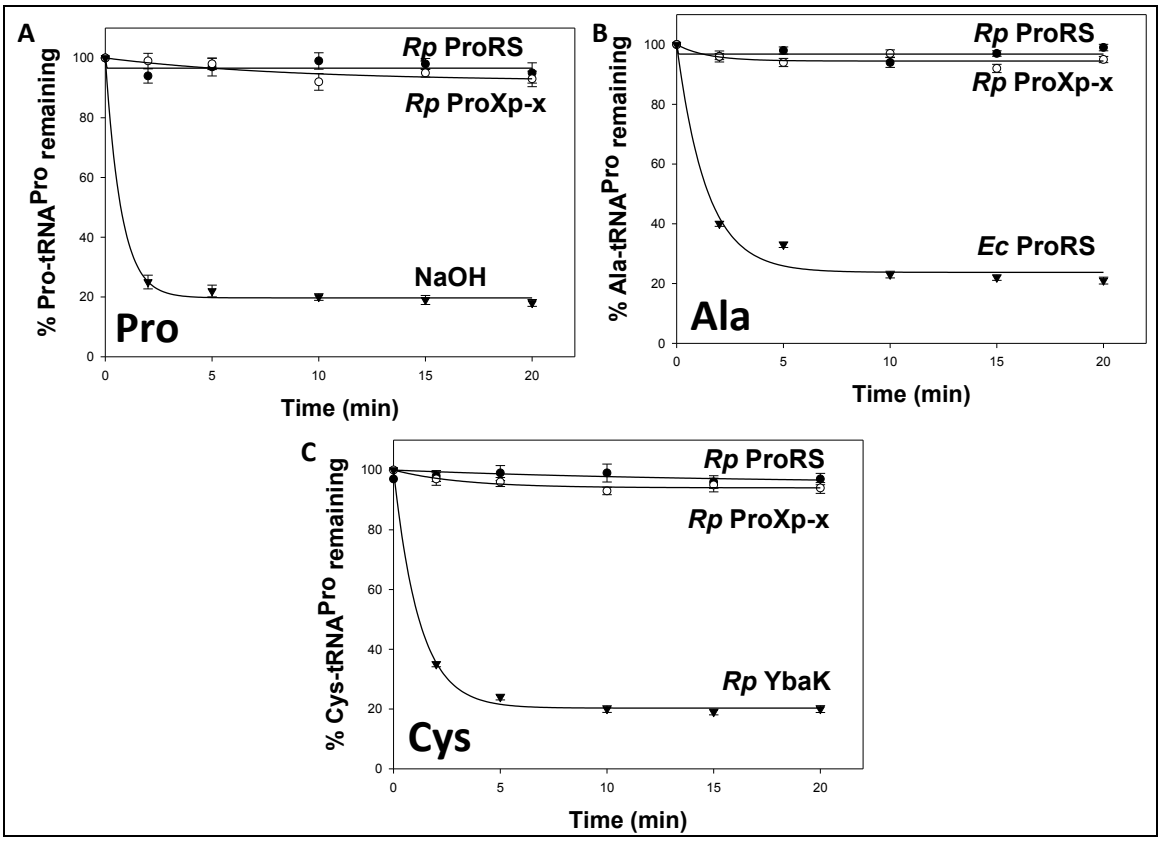


Figure 3.2. Post-transfer editing activities of *Rp* ProRS, ProXp-x, and YbaK. *A,* Pro-tRNA^{Pro} (0.5 μM) deacylation by 0.5 μM *Rp* ProRS (●), 0.5 μM *Rp* ProXp-x (○), and 120 mM NaOH (▼). *B*, Ala-tRNA^{Pro} (0.5 μM) deacylation by 0.5 μM *Rp* ProRS (●), 0.5 μM *Rp* ProXp-x (○), and 1 μM *Ec* ProRS (▼). *C*, Cys-tRNA^{Pro} (0.5 μM) deacylation by 0.5 μM *Rp* ProRS (●), 0.5 μM *Rp* ProXp-x (○), and 0.5 μM *Rp* YbaK (▼).

Figure 3.3. Sequence alignment of INS and ProXp-x. Shown are the INS domain sequences of *Enterococcus faecalis* (*Ef*), *Shigella dysenteriae* (*Sd*), and *Ec*. ProXp-x from *Rp*, *Aeopyrum pernix* (*Ap*), and *Thermus thermophillus* (*Tt*) are also shown. Conserved Lys and GXXXP motif are highlighted in cyan. Residues that constitute INS active site pocket are highlighted in magenta and aligned residues on ProXp-x are highlighted in yellow.

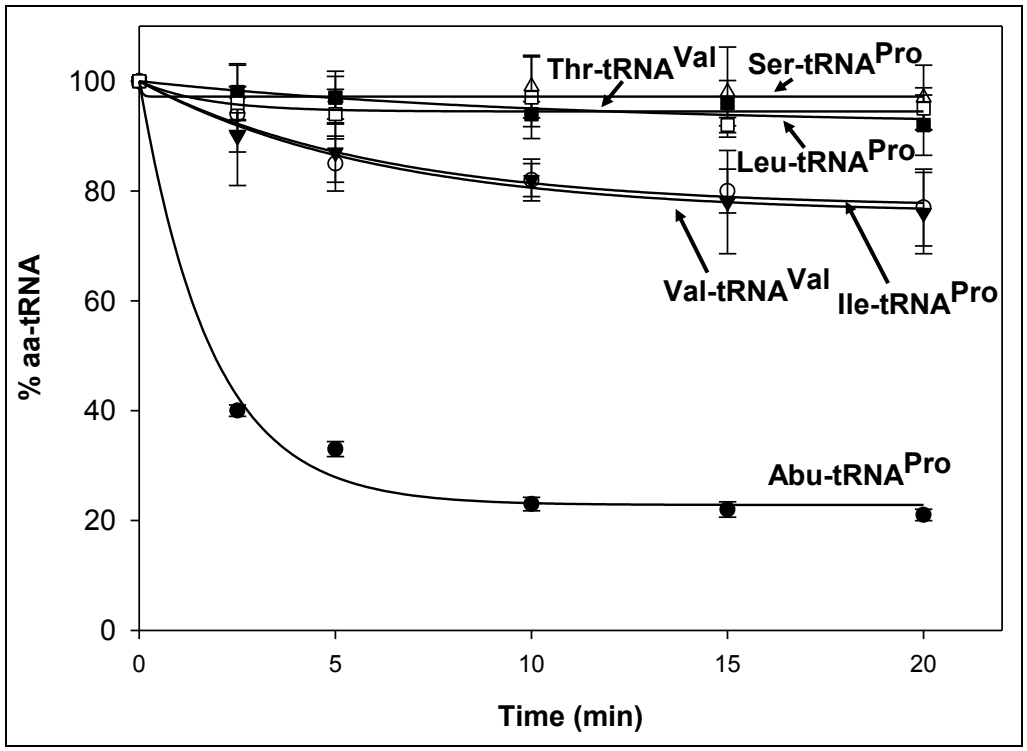


Figure 3.4. Deacylase activity of *Rp* ProXp against various substrates. Deacylation of (△) 0.5 μM Ser-tRNA^{Pro}, (□)Thr-tRNA^{Val}, (■) Leu-tRNA^{Pro}, (○) Ile-tRNA^{Pro}, (▼) Val-tRNA^{Val}, and (●) Abu-tRNA^{Pro} by 0.5 μM *Rp* ProXp-x (renamed ProXp-abu).

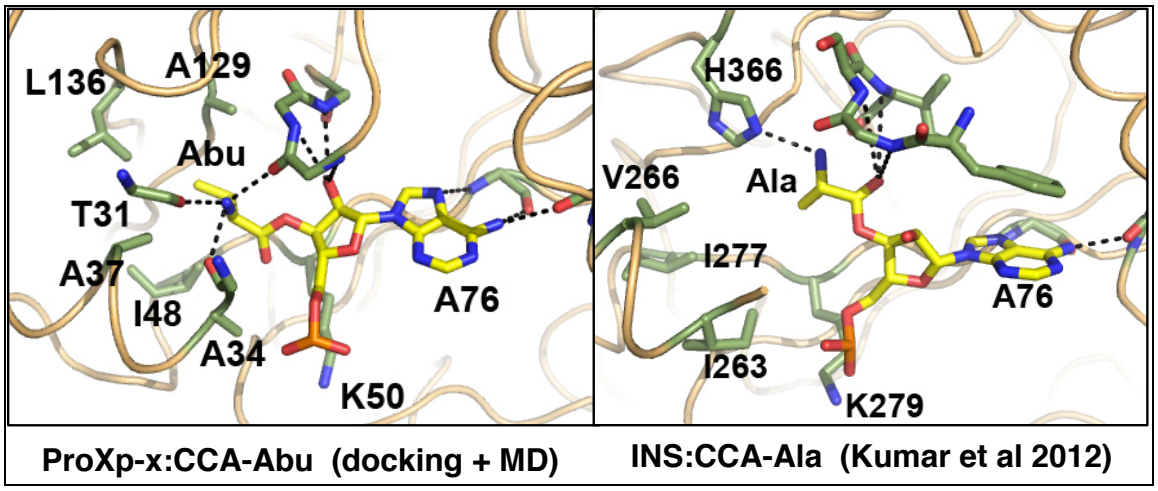


Figure 3.5. Docking models of CCA-Abu in *Rp* ProXp-x binding pocket and CCA-Ala in *Ef* ProRS INS binding pocket. Ribbon representation of ProXp-x in complex with CCA-Abu and *Ef* ProRS in complex with CCA-Ala. H-bonding interactions of substrate amino group and A76 are shown.

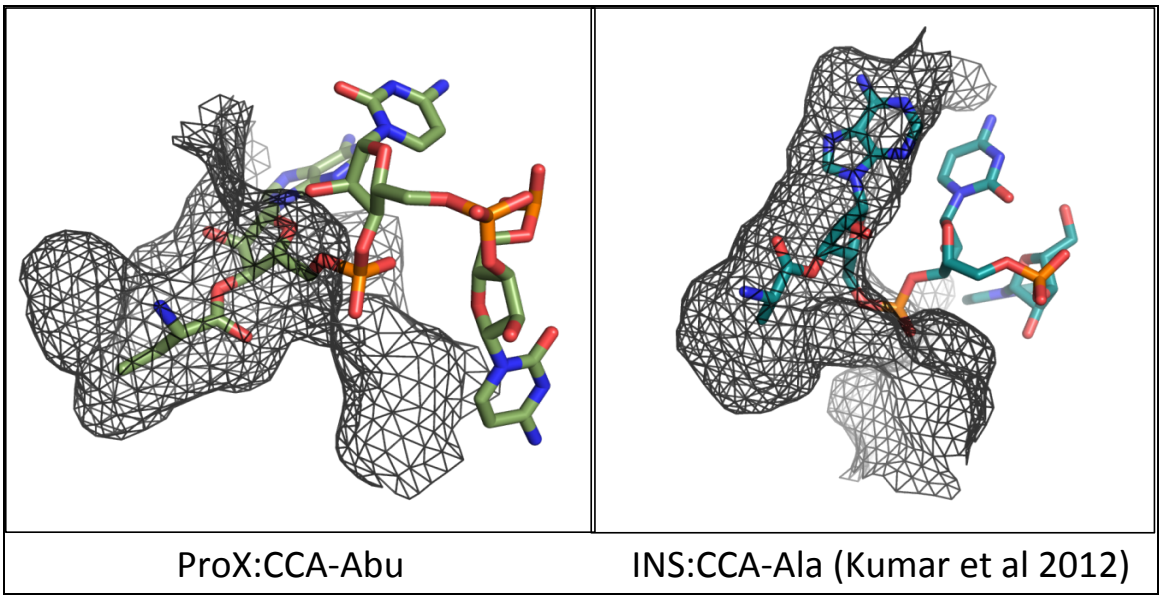


Figure 3.6. Mesh representation of ProXp-x and INS binding pockets. Shown are MD simulated docking models of Abu-CCA in *Tt* ProXp-x and Ala-CCA in *Ef* ProRS INS

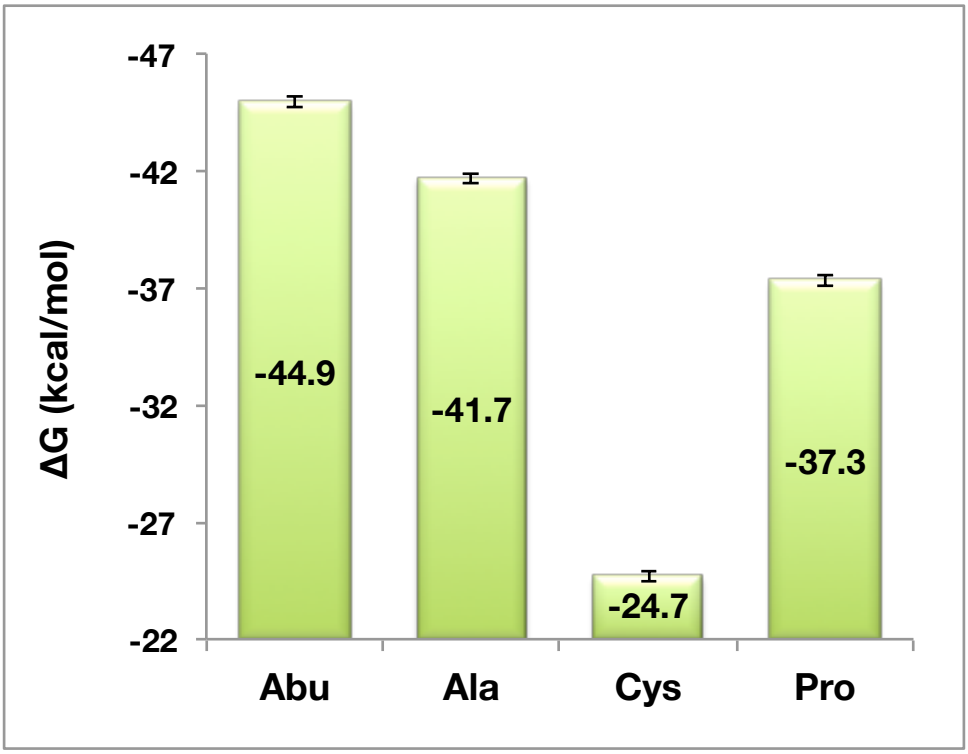


Figure 3.7. Chart of binding free energies. Graphical representation of MMPBSA calculated binding free energies of Abu, Ala, Cys, and Pro into the putative ProXp-x binding pocket.

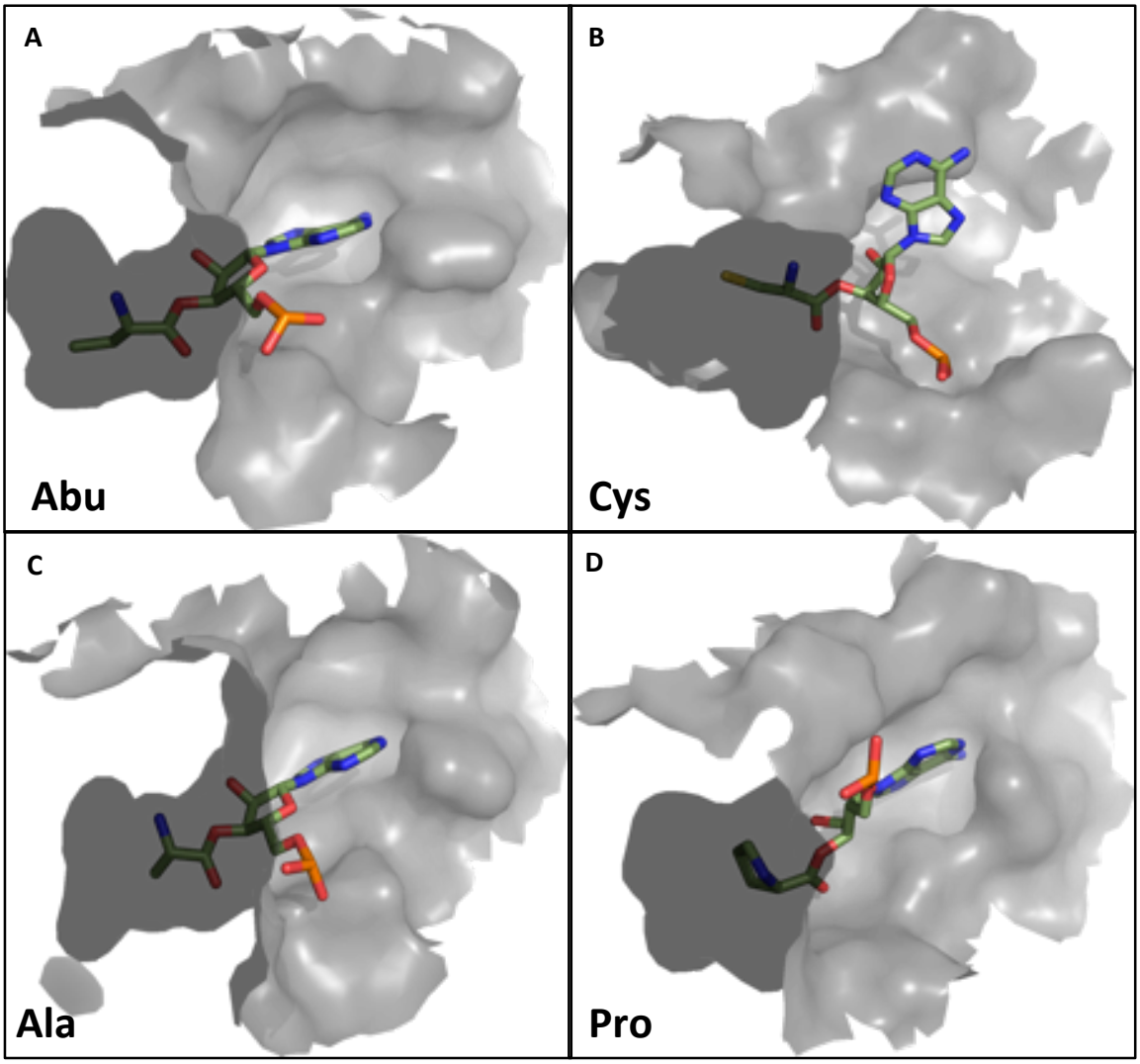


Figure 3.8. Depictions of ligand bound active site pockets of ProXp-x. Shown are depictions of (A) Abu, (B) Cys, (C) Ala, and (D) Pro bound in the putative active site of *Tt* ProXp-x (dark gray silhouette).

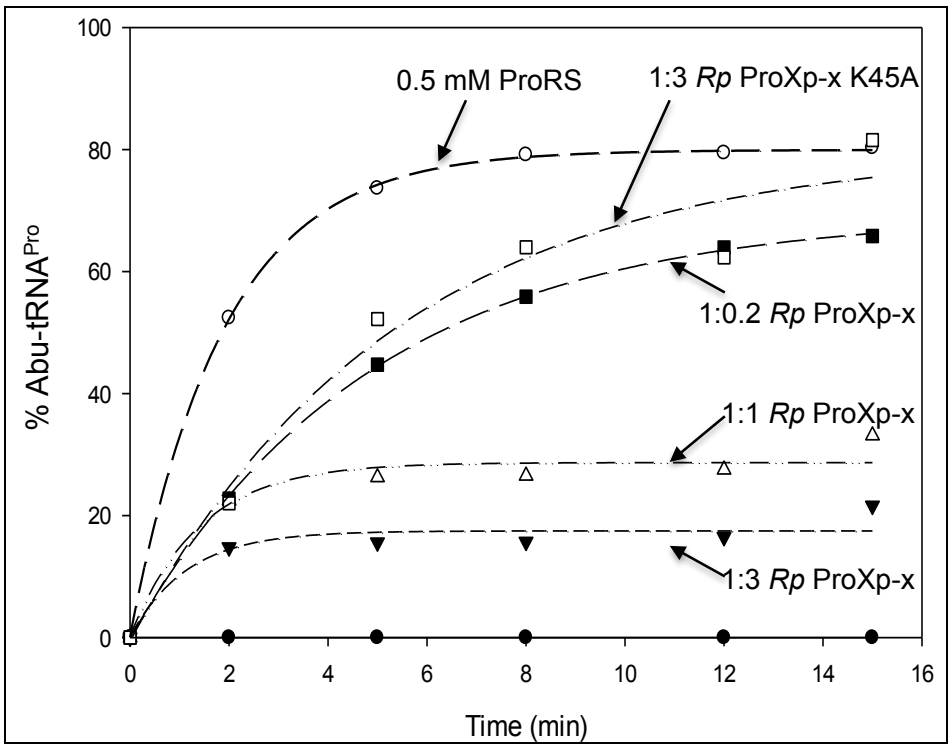


Figure 3.9. Effect of ProXp-x on Abu misacylation by *Rp* ProRS. Formation of mischarged Abu-tRNA^{Pro} by *Rp* ProRS alone (▼), 0.5 μM *Rp* ProRS and 0.1 μM ProXp-x (△), 0.5 μM *Rp* ProRS and 0.5 μM ProXp-x (○), 0.5 μM *Rp* ProRS and 1.5 μM ProXp-x (●), 0.5 μM *Rp* ProRS and 1.5 μM ProXp-x K45A. Reactions were carried out at 30 °C with 6 μM *Ec* tRNA^{Pro}.

Chapter 4 : Insights into the molecular basis of post-transfer editing and tRNA acceptor stem interactions of *Rhodopseudomonas palustris* ProXp-abu

4.1 Introduction

Translation of the genetic code into proteins is a central biological process to maintaining cellular homeostasis. Errors in translation cause dysfunction of cellular processes, and in some cases, cell death. Synthesis of the correct aminoacyl-tRNA (aa-tRNA) by aminoacyl-tRNA synthetases (aaRS) is an integral step in ensuring translational fidelity. aaRS are inherently error-prone due to the similarities in size and volume of their amino acid substrates (1). In order to overcome this limitation, aaRS have evolved various editing mechanisms to ensure formation of the correct aa-tRNA (5). Pre-transfer editing is the hydrolysis of the misactivated non-cognate amino acid releasing free amino acid and AMP (33). Conversely, post-transfer editing edits misacylated tRNAs resulting in free tRNA and amino acid (25).

ProRS mischarges Ala and Cys in addition to cognate Pro. In most bacterial systems, a triple-sieve editing mechanism is employed for accurate Pro codon translation (21, 79). The synthetic core acts as the primary sieve by screening against much larger and much smaller amino acids, while similar sized Ala, Cys, and non-protein amino acid aminobutyrate (Abu) are misactivated and mischarged along with cognate Pro (19). An editing domain fused to the ProRS core (INS domain) acts as a fine sieve and specifically

edits Ala-tRNA^{Pro} in *cis* (16). The *trans*-editing factor YbaK, which has been shown to possess robust editing activity against Cys-tRNA^{Pro}, acts as the third sieve (79, 81, 119).

YbaK belongs to a family of freestanding proteins homologous to the ProRS INS domain. Members of this family include YbaK, ProXp-ala, ProXp-abu, ProXp-y, and ProXp-z (14, 21, 22, 80). ProXp-ala has been shown to clear mischarged Ala-tRNA^{Pro} and has been demonstrated as an alternative editing domain in ProRSs that lack the full-length INS domain (14, 15, 22). In these ProRS systems, an alternative triple-sieve editing model is employed where proofreading of mischarged tRNAs is accomplished by two distinct *trans*-editing factors (22). *Rhodopseudomonas palustris* (*Rp*) encodes for a ProRS containing a severely mini-INS and two *trans*-editing factors YbaK and ProXp-abu. *Rp* YbaK has been shown to posess editing activity against Cys-tRNA^{Pro} consistent with YbaK proteins isolated from other species (21, 81, 119).

Although ProXp-abu shares structural and some sequence homology with INS, it is functionally distinct from the INS domain. ProXp-abu does not possess any deacylase activity for Ala-tRNA^{Pro} but recognizes Abu-tRNA^{Pro} instead. Structural and sequence alignments reveal that bulky residues constituting the active site pocket in INS are mostly replaced with smaller Ala in ProXp-abu. A MD-simulated model of Abu-CCA docked in the putative active site of ProXp-abu shows that the Abu moiety of the substrate is flipped ~180 degrees relative to Ala docked into the INS binding pocket. Similar to INS, residues constituting the putative binding pocket of ProXp-abu are primarily non-polar. This is consistent with *in vitro* experiments (Chapter 3) that demonstrate rejection of substrates with polar side-chains and preference for hydrophobic amino acids such as Val

and Ile. Here, we investigate the molecular basis of substrate discrimination of ProXp-abu by mutagenesis. Additionally, we explore tRNA acceptor stem recognition elements that may be critical for correct aa-tRNA recognition by ProXp-abu.

4.2 Experimental Procedures

4.2.1 Materials

All amino acids and chemicals were purchased from Sigma unless otherwise noted. [α-^{32}P]-ATP was purchased from Perkin Elmer Life Sciences. Abu-3,5-dinitrobenzyl ester (Abu-DBE) was supplied by Hiro Suga (University of Tokyo). Biotinylated dinitro-flexizyme (dFx) was acquired from ThermoScientific.

4.2.2 Preparation of ProXp-abu mutants

Rp ProXp-abu A131L, A131S, A131F mutants were prepared through site-directed mutagenesis using pET15b *Rp* ProXp-abu as template. Following transformation into *Ec* XL1-Blue, plasmid DNA was isolated and DNA sequencing (Genewiz) confirmed mutations.

4.2.3 Enzyme expression and purification

N-terminally histidine-tagged WT *Rp* ProX-abu and *Escherichia coli* (*Ec*) nucleotidyl-tRNA transferase were prepared as previously described (124). *Rp* ProXp-abu mutants were transformed into BL21-CodonPlus (DE3)-RIL (Stratgene).

Overexpression was induced with 0.1 mM IPTG at 25° C for 14-16 hours. The N-terminally His-tagged proteins were purified using HIS-select® nickel resin (Sigma-Aldrich). Enzyme concentrations were determined using a Bradford assay.

4.2.4 Preparation of aminoacyl-tRNA substrates

3′-[^{32}P]-labeled Pro-, Ala-, and Cys-tRNA substrates were prepared as detailed in Chapter 3. Abu-tRNA substrates were prepared using flexizyme catalyzed aminoacylation of Abu-DBE onto *Ec* tRNA. Briefly, *Ec* tRNA and dFx (42 mM each in 18 μL reaction volume) and trace amounts of 3'-[^{32}P]-labeled *Ec* tRNA were heated to 95° C for 1 min, followed by addition of $MgCl_2$ and Abu-DBE to a final concentration of 25mM and 5mM, respectively. dFx-catalyzed aminoacylation reactions were carried out for 2 h on ice (~4° C). Reactions were quenched with 120 μL of 0.3 M sodium acetate (NaOAc, pH 5) to a final concentration of 0.25 mM. Biotinylated dFx was removed by incubating the reaction mixture with 150 μL of streptavidin agarose resin (Novagen) for ~30 min on ice followed by centrifugation in a table top centrifuge at 4° C for 2 min at 13,000 rpm. The supernatant containing the aminoacylated tRNA was removed. Following aminoacylation, aa-tRNAs were phenol-chloroform extracted and ethanol precipitated. The aa-tRNA pellets were dissolved in diethylpyrocarbonate (DEPC)-treated water and stored at 80° C for future use.

4.2.5 Deacylation assays

Cys-, Ala-, Pro-, and Abu-tRNA deacylation reactions were performed as described in Chapter 3. Briefly, reactions containing ~0.5 μM aminoacyl-tRNA and buffer (300 mM KPO_4 (pH 7), 0.2 mg/ml BSA, and 9.6 mM $MgCl_2$) were initiated by addition of 0.5 μM enzyme. For Cys-$tRNA^{Pro}$ deacylation, 0.5 μM *Rp* YbaK was used as control. For Ala-tRNA deacylation, 0.5 μM *Ec* ProRS (INS) was used as control. For Pro-tRNA deacylation, 0.5 μM *Caulobacter crescentus* (*Cc*) V29A ProXp-ala was used as control. Finally, for Abu-tRNA deacylation, 0.5 μM WT *Rp* ProXp-abu was used as positive control. At the indicated time points, reaction aliquots (2 μl) were quenched into 4 μl of a solution containing 0.4 U/μl P1 nuclease in 200 mM NaOAc (pH 5). Deacylation levels were monitored using polyethyleneimine-cellulose TLC and analyzed as previously described (124).

4.2.6 Analytical Ultracentrifugation (AUC) sedimentation velocity assays

N-terminal His-tag was cleaved off *Rp* ProXp-abu with thrombin and purified on nickel-affinity column (HIS-select® nickel resin) followed by FPLC purification (GE FPLC AKTA Purifier System) on Superdex 75 16/60 column to separate His-tag cleaved and uncleaved proteins. His-tag cleaved *Rp* ProXp-abu was conjugated to Alexa Fluor® 488 5-TFP (Life Technologies) as described (148). Prior to sedimentation velocity experiments, all samples were dialysed for at least 20 h in buffer containing 1 mM $MgCl_2$, 50 mM HEPES pH 6.8, 20 mM NaCl. Analytical ultracentrifugation experiments were performed using ProteomeLab XL-I (Beckman Coulter). Sedimentation velocity

experiments were performed by protein samples (420 μL) into a double sector Epon charcoal-filled centerpiece and subjected to an angular velocity of 50,000 rpm at 20° C. Absorbance scans were collected every 10 min at 260 nm for tRNA, 280 nm and 495 nm for *Rp* ProXp-abu, 280 nm for *Rp* ProRS. Sedimentation velocity data was analysed using SedFit, where SedFit generates a c(s) distribution of Lamm equation solutions (149).

4.3 Results and Discussion

4.3.1 ProXp-abu active site is not tunable

To investigate the molecular basis of substrate discrimination by ProXp-abu and gain mechanistic insights into its deacylase activity, we performed mutational studies on residues in the putative active side pocket. A MD-simulated docking model of Abu-CCA bound in the *Thermus thermophllus* (*Tt*) ProXp-abu substrate binding pocket revealed that Leu 136 (Ala 131 in *Rp* ProXp-abu) projected directly towards the Abu substrate and may be critical in modulating substrate discrimination. To investigate its role in substrate specificity, mutated this residue. The ProXp-abu binding pocket is highly hydrophobic, thus, rejecting amino acids with polar side chains such as Cys. We hypothesize that mutating A131 on *Rp* ProXp-abu into Ser may allow accomodation of moderately polar substrates like Cys by reducing the overall hydrophobicity of substrate binding pocket (Figure 4.1.A). In Chapter 3, we show the ProXp-abu substrate binding pocket is larger compared to INS. We predict that reducing the active site volume of ProXp-abu may provide a better fit for smaller substrates like Ala. To test this, we mutated A131 on *Rp* ProXp-abu into larger Phe (Figure 4.1.B). Finally, we prepared *Rp* A131L ProXp-abu to

recreate the *Tt* ProXp-abu binding pocket (Figure 4.1.C). Deacylase activity against Cys-, Ala-, Pro-, and Abu-tRNA^{Pro} was determined for each of the 3 *Rp* ProXp-abu mutants.

As expected, WT *Rp* ProXp-abu does not show editing activity for Cys-tRNA^{Pro} (Figure 4.2.A). Furthermore, none of the ProXp-abu variants demonstrated deacylation activity for Cys-tRNA^{Pro}. Mutants A131F and A131L maintain hydrophobicity in the active site pocket and are expected to reject Cys. Lack of Cys editing by the A131S variant suggests Ser may not be sufficient to significantly alter the overall hydrophobicity of the active site pocket. Mutation to a more polar amino acid or mutation of more than one residue may be required to alter the substrate specificity of *Rp* ProXp-abu from Abu to Cys. None of the ProXp-abu mutants, including WT *Rp* ProXp-abu, showed deacylation activity against cognate Pro-tRNA^{Pro} (Figure 4.2.B). This is a positive result as deacylation of cognate aa-tRNA is unwanted. Interestingly, the A131S mutant shows Abu-tRNA^{Pro} deacylation activity comparable to that of WT *Rp* ProXp-abu (Figure 4.2.C). This result agrees with our hypothesis that Ser substitution in the 131 position of *Rp* ProXp-abu may not be sufficient enough to alter the hydrophobicity of the catalytic core. Mutant A131F lacks Abu-tRNA^{Pro} deacylation activity. Substitution to a bulky residue likely reduced the active site pocket size resulting in exclusion of Abu. As expected, WT, A131S, and A131L *Rp* ProXp-abu failed to deacylate Ala-tRNA^{Pro} deacylation. Contrary to our expectations, the A131F mutant also lacks activity for Ala-tRNA^{Pro}. Altogether, our results suggest the ProXp-abu catalytic pocket is not tunable, and that size exclusion may not be the determining factor for substrate discrimination. This contrasts with INS, where modulating the size of the active site switches its

substrate specificity from smaller Ala to larger Cys (20). The inability to alter substrate specificity by modulating the active pocket size is characteristic of YbaK, which discriminates substrates based on the chemical identity of the amino acid side chain (21). However, a similar mechanism is likely not applicable to ProXp-abu due to the chemically inert Abu side chain. Mutation analysis of other conserved residues constituting the active site pocket is required to fully understand the mechanism for substrate discrimination by ProXp-abu.

4.3.2 tRNA specificity of *Rp* ProXp-abu

Recognition of aa-tRNA substrates is not solely based on the amino acid moiety but also on the tRNA. A good example is Ef-TU whose affinity for aa-tRNA is dictated by both the amino acid and the tRNA moieties (150). Abu is not only mischarged by ProRS but also by ValRS and IleRS (29, 30). Here, we investigate various features of the tRNA acceptor stems of the three tRNAs to determine elements that may contribute to *Rp* ProXp-abu recognition and substrate discrimination.

Examination of the acceptor stem features of $tRNA^{Val}$, $tRNA^{Pro}$, and $tRNA^{Ile}$ show all three tRNAs share the same discriminator base, A73, but contain different first base pairs (Figure 4.3). *Rp* and *Ec* tRNAs share identical acceptor stem elements (from N73 through the third base pair, N3:N70) except for $tRNA^{Ile}$, which differ at the first base pair (A1:U72 for *Ec* $tRNA^{Ile}$ and G1:C72 for *Rp* $tRNA^{Ile}$) (Figure 4.3). To investigate the importance of the first base pair for recognition by *Rp* ProXp-abu, deacylation assays with Abu-tRNAs were performed. Results show that all three tRNAs are deacylated at

comparable rates (Figure 4.4). This suggests the first base pair may not be critical for tRNA recognition.

We then proceeded to investigate the role of discriminator base and generated $tRNA^{Pro}$ variant A73C. Deacylation assays show the activity is moderately reduced (3.9-fold) when the discriminator base is mutated from an A to a C (Figure 4.4). In contrast, the triple mutant C73/G1:C72, is reduced 7.1-fold in k_{obs} relative to WT $tRNA^{Pro}$, and is ~2-fold less active than the A73C variant. Altogether, our observations suggest the discriminator base contributes weakly to tRNA recognition by ProXp-abu. In *Rp*, over 50% of the tRNAs encode for an A73. This implies that despite the moderate tRNA specificity of ProXp-abu, it is capable of acting as a general deacylase *in vivo*. The relaxed tRNA specificity of ProXp-abu may be advantageous to cells due to the non-protein nature of its amino acid substrate.

4.3.3. Interactions of *Rp* ProXp-abu with potential binding partners

Due to the relaxed tRNA selectivity of ProXp-abu, we hypothesize that it may gain specificity through interaction partners. To test our hypothesis, we decided to characterize possible interactions with ProRS and $tRNA^{Pro}$ using analytical ultracentrifugation (AUC). AUC sedimentation velocity (SV) analysis shows *Rp* ProRS exists as a single species with a sedimentation coefficient (S) of ~5.7 (Figure 4.5). We conjugated *Rp* ProXp-abu with Alexa Fluor 488 (AF488) to allow exclusive monitoring of *Rp* ProXp-abu at 495 nm. The SV profile of *Rp* ProXp-abu AF488 alone shows a single peak with maxima at ~1.8 S (Figure 4.6). To investigate the interaction of *Rp*

ProXp-abu with *Rp* ProRS, we compared SV profiles of ProRS alone, ProXp-abu alone, and ProRS + ProXp-abu collected at 280 nm and 495 nm. Overlay of the three profiles at 280 nm shows one peak at 1.8 S corresponding to ProXp-abu and another at 5.8 S correnssponding to ProRS (Figure 4.7). Appearance of a new peak is not observed. Furthermore, we compared scans of ProRS + ProXp-abu collected at 280 nm and 495 nm to check whether ProXp-abu (monitored at 495 nm) is shifted to the apparent peak observed at 5.8 S (Figure 4.8). The SV profile at 495 nm shows no shift of ProXp-abu species to the higher molecular weight peak, which suggests ProXp-abu does not interact with ProRS.

We also probed the interaction of *Rp* ProXp-abu and *Ec* $tRNA^{Pro}$ by collecting SV profiles of ProXp-abu alone, tRNA alone, and ProXp-abu + tRNA at 260 nm and 495 nm. Overlay of scans collected at 260 nm shows peaks at 1.8 S consistent with ProXp-abu, 3.8 S corresponding to tRNA alone, and 7.1 S corresponding to tRNA-ProXp-abu complex formation (Figure 4.9). Comparison of SV profiles of ProXp-abu + tRNA collected at 260 nm and 495 nm shows a shift of ProXp-abu species to a higher molecular weight corresponding to the new peak (7.1 S) observed in the 260 nm trace (Figure 4.10). This observation shows ProXp-abu interacts with tRNA and is consistent with our *in vitro* studies demonstrating ProXp-abu has moderate tRNA specificity. Interaction of *Rp* ProXp-abu with A73C $tRNA^{Pro}$ should also be characterized with AUC to complement our *in vitro* experiments on ProXp-abu tRNA recognition. In this case, we expect no formation of the ProXp-abu-tRNA binary complex.

Although ProXp-abu does not interact with ProRS, it is possible interaction is mediated indirectly via the tRNA. To investigate possible formation of a ProRS-tRNA-ProXp-abu ternary complex, SV analysis was performed for each of the three components separately for a mixture of all three components. Overlay of SV profiles collected at 260 nm shows peaks at 1.8 S consistent with ProXp-abu, 3.8 S corresponding to tRNA, 5.8 S corresponding to ProRS, and 7.1, which is likely the tRNA-ProXp-abu binary complex (Figure 4.11). This data suggests there is no ternary complex formed with ProRS, ProXp-abu, and tRNA. Comparison of SV profiles of ProRS+tRNA+ProXp-abu collected at 260 nm, 280 nm, and 495 nm shows ProXp-abu is present in a higher molecular weight species corresponding to that of tRNA-ProXp-abu complex but no evidence for a ternary complex with ProRS (Figure 4.12). Taken altogether, our data suggests ProXp-abu interacts with tRNA, in agreement with the observed moderate tRNA specificity, but not with ProRS. Moreover, no formation of a ternary complex was observed, thus, tRNA-mediated interaction between ProRS and ProXp-abu is unlikely. Interactions of ProXp-abu with ValRS and IleRS should also be investigated as Abu is also mischarged onto $tRNA^{Val}$ and $tRNA^{Ile}$ and may likely interact with these aaRS for function and recruitment.

4.4. Conclusions

Tunability of the substrate binding pocket of *Rp* ProXp-abu was investigated by modulating the active site pocket size. These studies showed that substrate specificity is not readily altered. This suggests a different model for substrate discrimination may be

employed by ProXp-abu. Further characterization of various active site features is required to determine the mechanism of substrate discrimination and deacylation of ProXp-abu. Exploring tRNA acceptor stem features reveals A73 is a weak recognition element for *Rp* ProXp-abu. Mutating this residue to a C slightly reduces deacylation by ProXp-abu. In contrast, the enzyme was less sensitive to changes at 1:72 base pair. The relaxed tRNA specificity of ProXp-abu suggests it may act as a general deacylase *in vivo*. AUC sedimentation velocity to experiments complex formation behavior of ProXp-abu, tRNA, and ProRS showed interactions with tRNA but not with ProRS.

4.5 Acknowledgements

We would like to thank Dr. Marina Bhaktina for assistance in conducting and data analysis of all AUC experiments. We would also like to recognize all the hard work of Brian Bizub in preparing all *Rp* ProXp-abu variants and performing all deacylation assays as part of his Research Experience for Undergraduates (REU) Summer project.

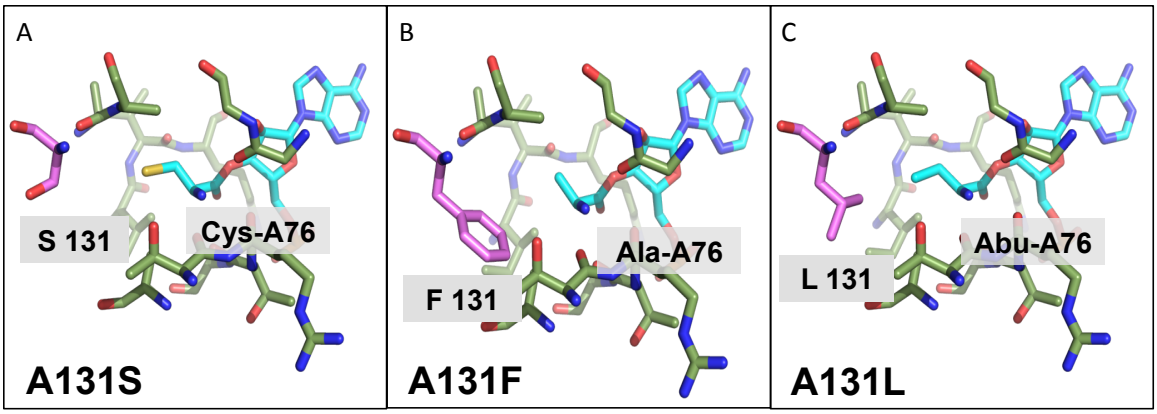


Figure 4.1. Proposed *Rp* ProXp-abu active site mutations. Depiction of residues in the active site of *Tt* ProXp-abu (151). Proposed mutations (A) A131S, (B) A131F, and (C) A131L on *Rp* ProXp-abu are highlighted in pink. Predicted substrate accommodation for each mutation is indicated (aa-A76).

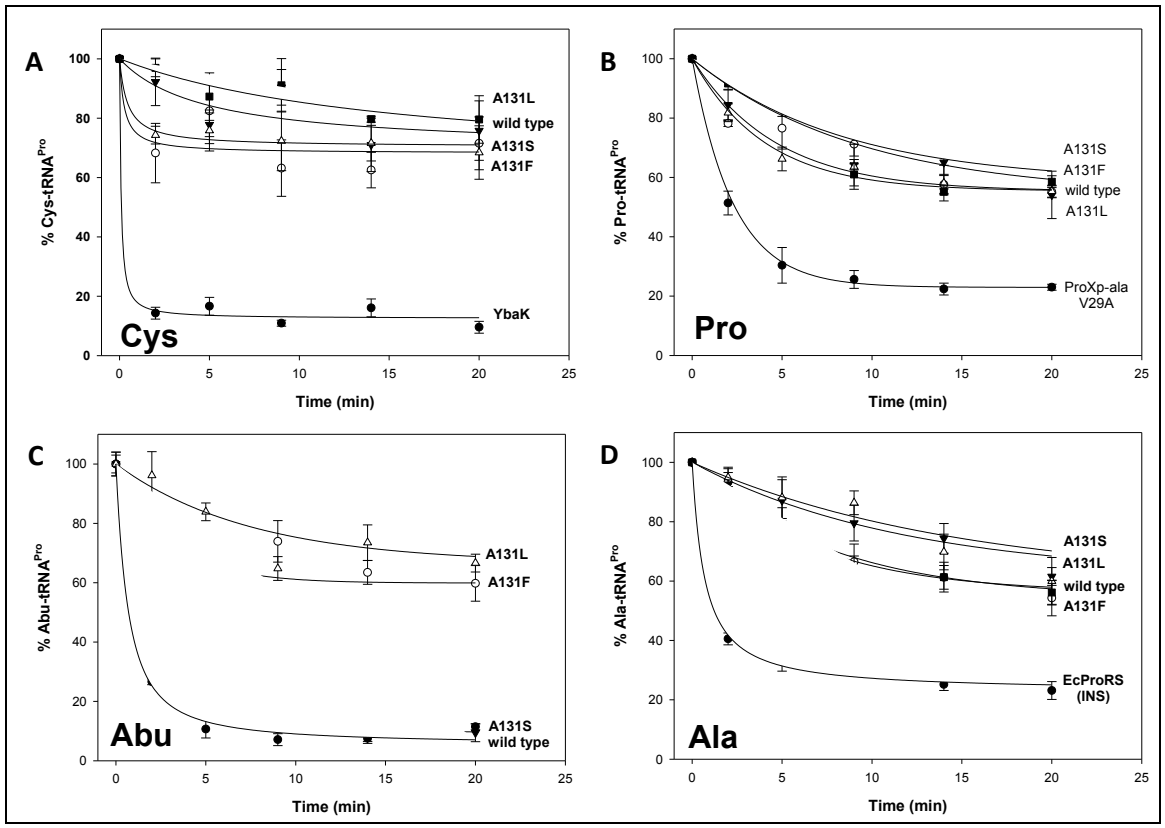


Figure 4.2. Deacylation assays by ProXp-abu variants. (A) Deacylation of 0.5 μM Cys-tRNA^{Pro} by (△) 3 μM A131S, (●) 3 μM A131L, (○) 3 μM A131F, (■) 3 μM *Rp* ProXp-abu, and (●) 0.5 μM *Rp* YbaK at 37º C. (B) Deacylation of 0.5 μM Pro-tRNA^{Pro} by (■) 3 μM A131S, (△) 3 μM A131L, (○) 3 μM A131F, (▼) 3 μM *Rp* ProXp-abu, and (●) 0.5 μM *Cc* V29A ProXp-ala at 30º C. (C) Deacylation of 0.5 μM Abu-tRNA^{Pro} by (▼) 3 μM A131S, (△) 3 μM A131L, (○) 3 μM A131F, and (○) 0.5 μM *Rp* ProXp-abu at 30º C. (D) Deacylation of 0.5 μM Ala-tRNA^{Pro} by (△) 3 μM A131S, (▼) 3 μM A131L, (■) 3 μM A131F, (○) 3 μM *Rp* ProXp-abu, and (●) 1 μM *Ec* ProRS (INS) at 30º C.

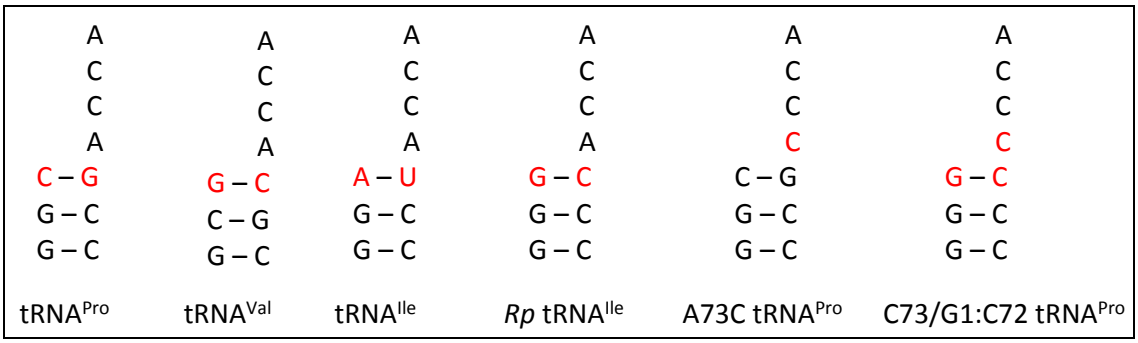


Figure 4.3. Acceptor stem elements of *Ec* tRNAs tested for deacylation by ProXp-abu. The first three base pairs of the acceptor stem, discriminator base, and the conserved CCA end of each tRNA are shown. Highlighted in red are the first base pairs and the elements in $tRNA^{Pro}$ tested for recognition: A73C and combined change of A73/C1:G72 to C73/G1:C72.

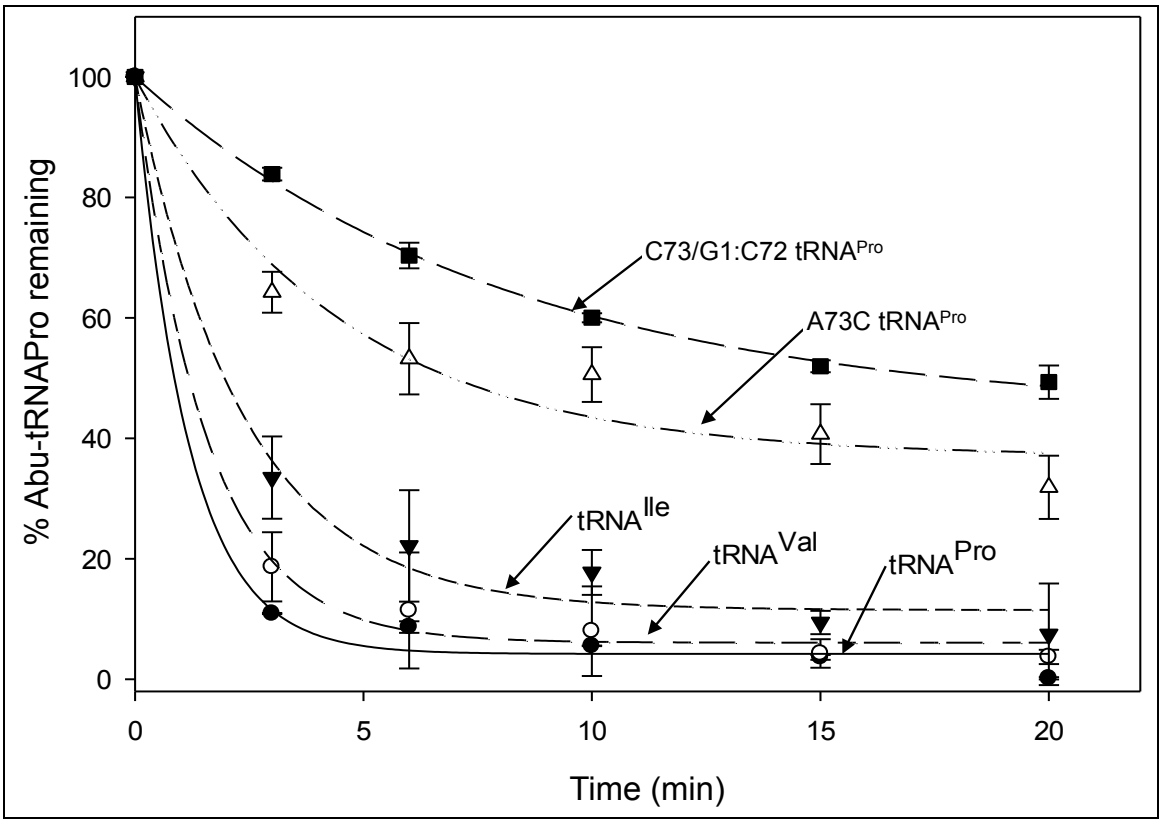


Figure 4.4. Deacylation of Abu-tRNA variants by *Rp* ProXp-abu. Deacylation of 0.5 μM (●) Abu-tRNA^{Pro}, (○) Abu-tRNA^{Val}, (▼) Abu-tRNA^{Ile}, (△) A73C Abu-tRNA^{Pro}, (■) C73/G1:C72 Abu-tRNA^{Pro} by 0.5 μM *Rp* ProXp-abu.

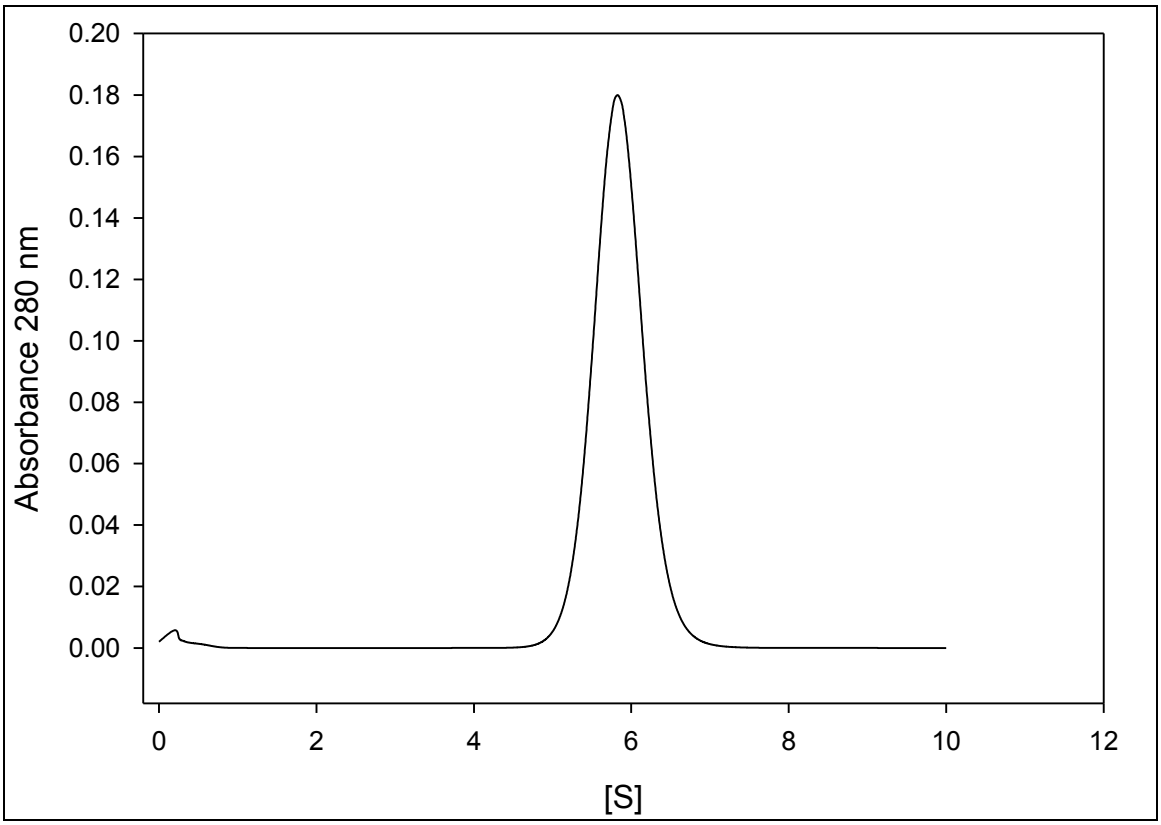


Figure 4.5. Sedimentation velocity profile of *Rp* ProRS. Shown is the SV profile of 0.3 μM *Rp* ProRS alone, where sedimentation coefficient ([S]) is plotted against absorbance at 280 nm.

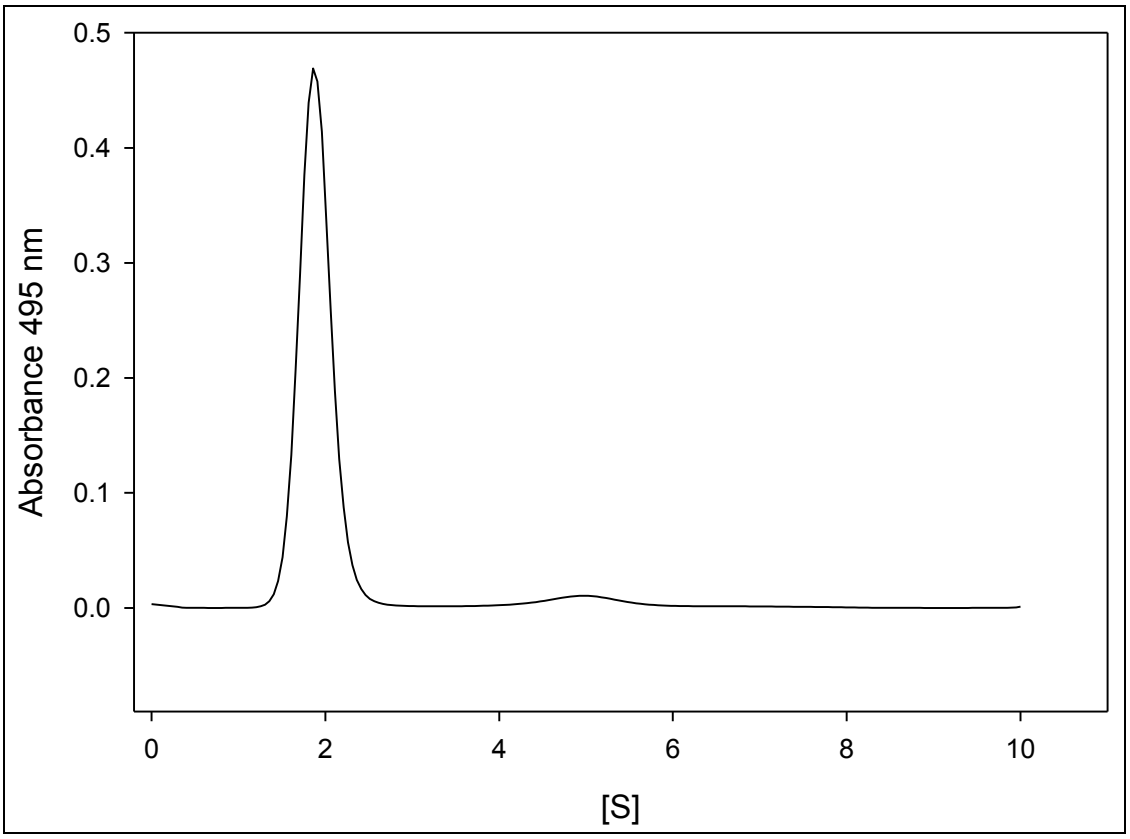


Figure 4.6. Sedimentation velocity profile of *Rp* ProXp-abu AF488. Shown is the SV profile of 2 μM *Rp* ProXp-abu AF488 alone, where sedimentation coefficient ([S]) is plotted against absorbance at 495 nm.

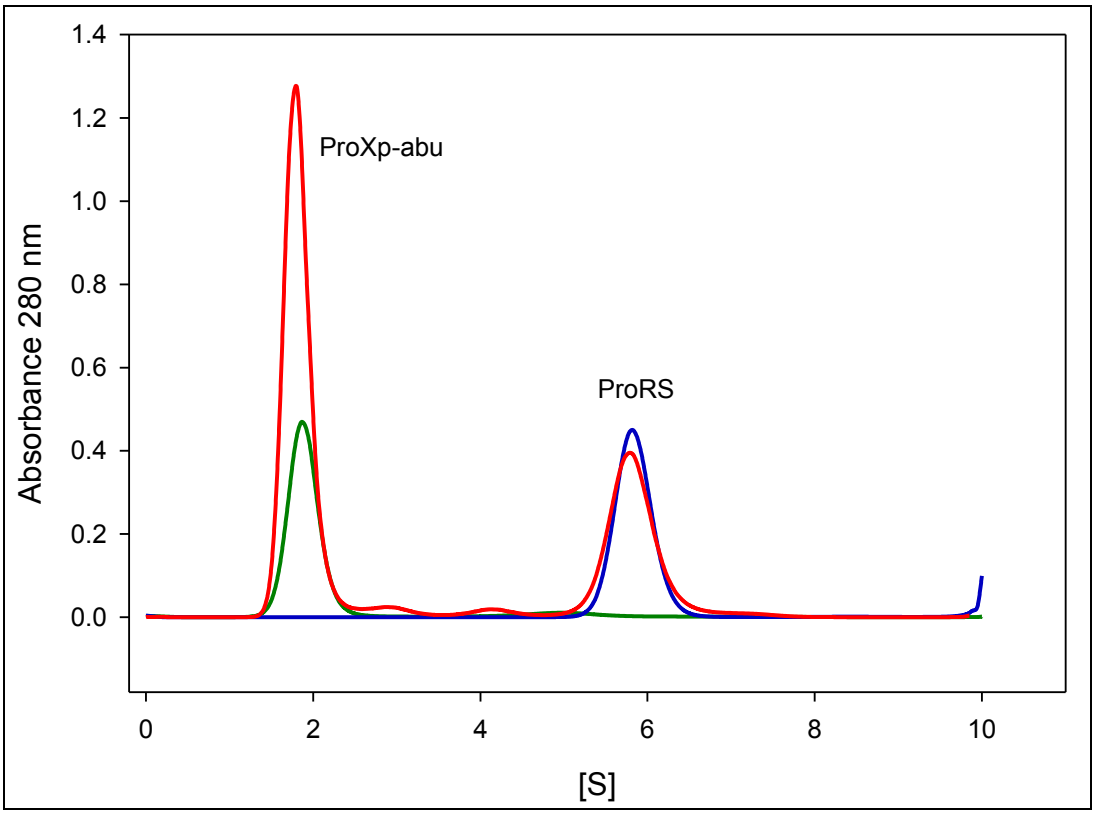


Figure 4.7. Sedimentation velocity profile of *Rp* ProXp-abu AF488 and *Rp* ProRS interaction. Shown are the SV profiles of *Rp* ProXp-abu AF488 and *Rp* ProRS where sedimentation coefficient ([S]) is plotted against absorbance at 280 nm. SV profiles of 2 µM *Rp* ProXp-abu AF488 alone (green), 0.3 µM ProRS alone (blue), and ProXp-abu AF488 + ProRS complex (red) are shown. Corresponding species in each peak are indicated.

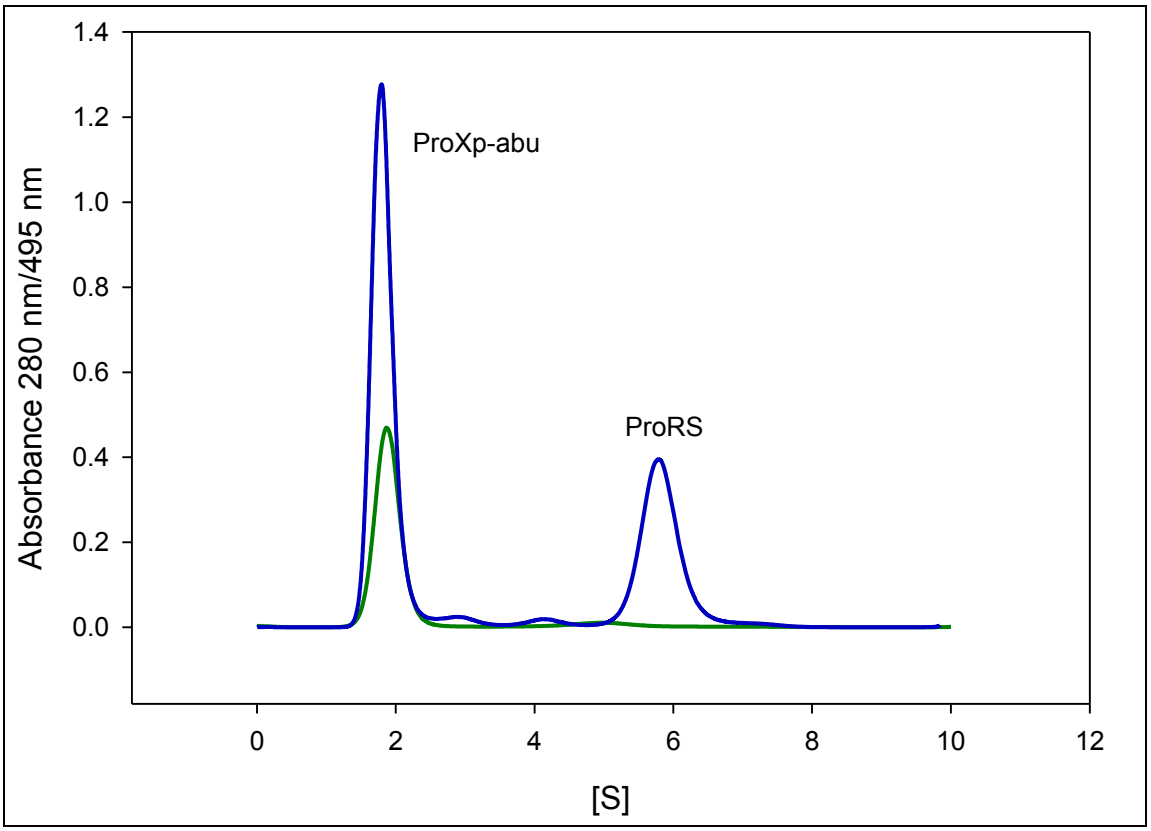


Figure 4.8. Sedimentation velocity profile of Rp ProXp-abu AF488 and Rp ProRS interaction. Shown are the SV profiles of 2 μM Rp ProXp-abu AF488 and 0.3 μM ProRS complex monitored at 280 nm (blue) and 495 nm (green). Sedimentation coefficient ([S]) is plotted against respective absorbances (280 nm or 495nm). Corresponding species in each peak are indicated.

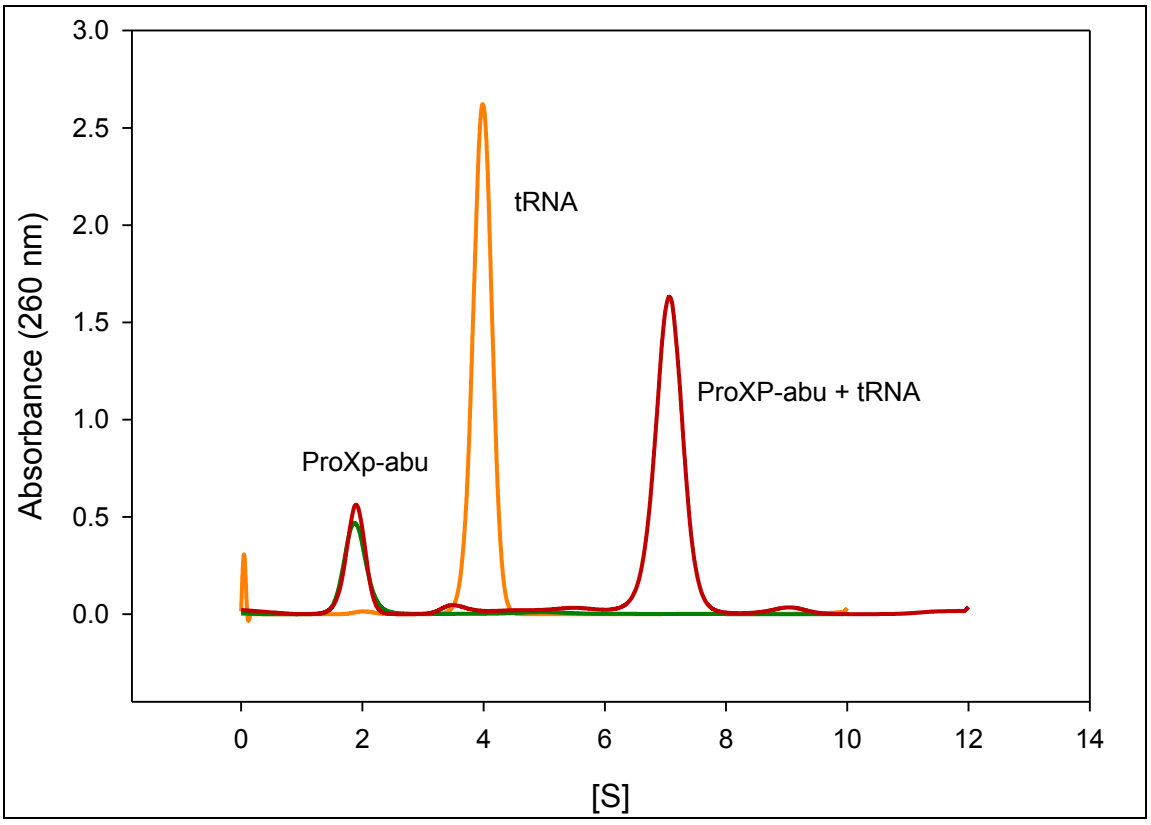


Figure 4.9. Sedimentation velocity profile of *Rp* ProXp-abu AF488 and *Ec* tRNA interaction. Shown are the SV profiles of *Rp* ProXp-abu AF488 and *Ec* tRNA where sedimentation coefficient ([S]) is plotted against absorbance at 260 nm. SV profiles of 2 µM *Rp* ProXp-abu AF488 alone (green), 1 µM *Ec* tRNA alone (orange), and ProXp-abu AF488 + tRNA complex (red) are shown. Corresponding species in each peak are indicated.

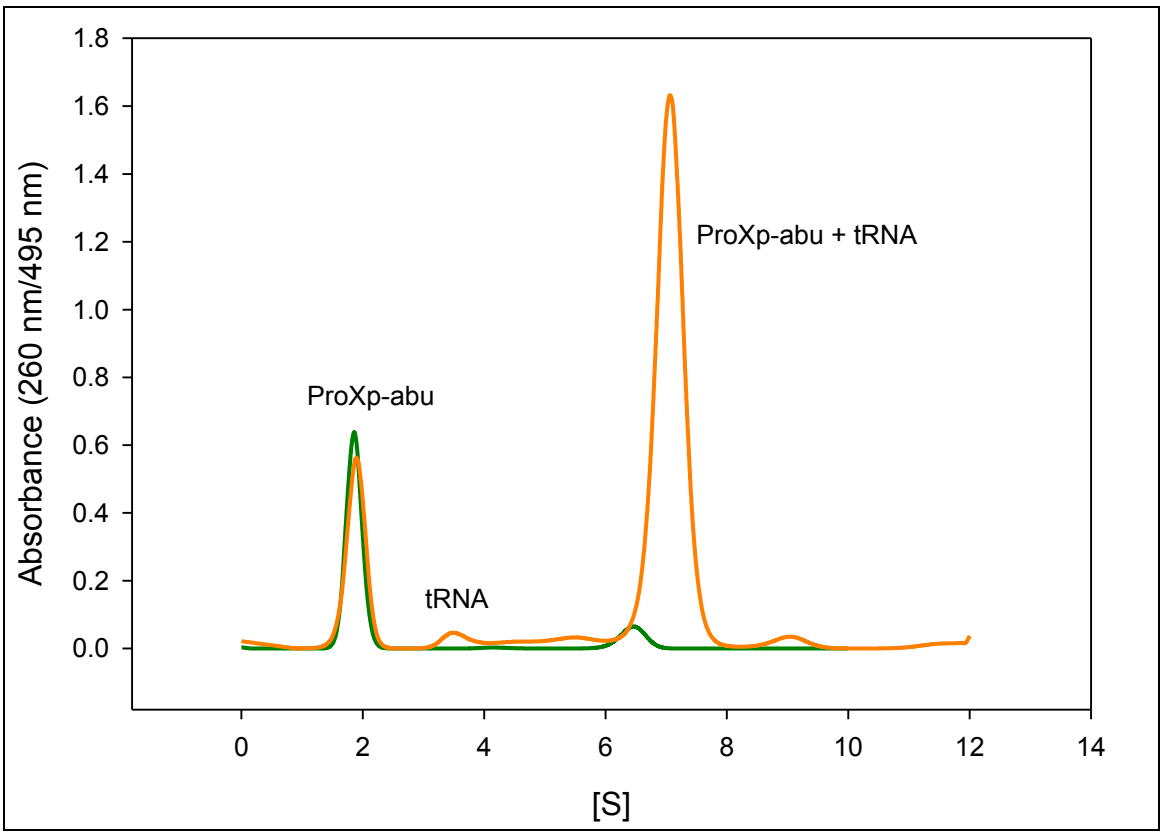


Figure 4.10. Sedimentation velocity profile of *Rp* ProXp-abu AF488 and *Ec* tRNA interaction. Shown are the SV profiles of 2 μM *Rp* ProXp-abu AF488 and 1 μM *Ec* tRNA complex collected at 260 nm (orange) and 495 nm (green). Sedimentation coefficient ([S]) is plotted against respective absorbance (260 nm/495nm). Corresponding species in each peak are indicated.

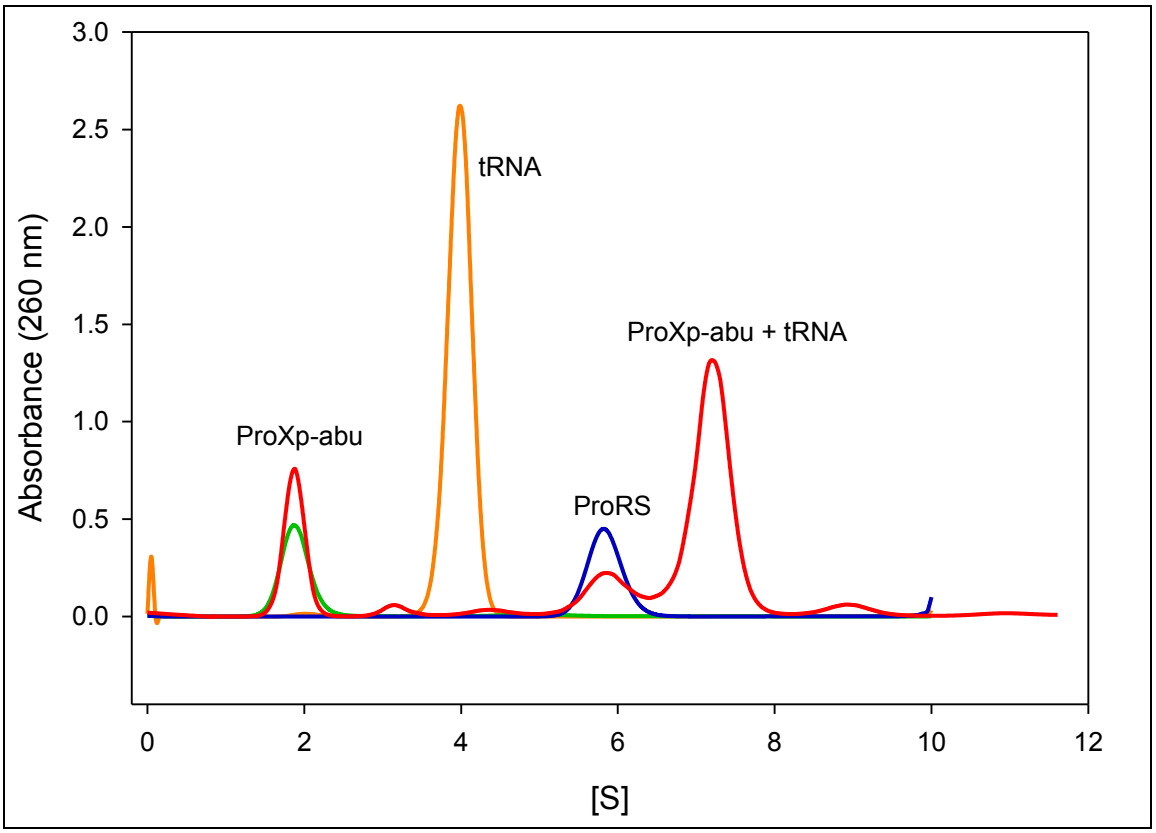


Figure 4.11. Sedimentation velocity profile of *Rp* ProXp-abu AF488, *Rp* ProRS, and *Ec* tRNA interaction. Shown are the SV profiles of *Rp* ProXp-abu AF488, *Rp* ProRS, and *Ec* tRNA where sedimentation coefficient ([S]) is plotted against absorbance at 260 nm. SV profiles of 2 µM *Rp* ProXp-abu AF488 alone (green), 0.3 µM *Rp* ProRS (blue), 1 µM *Ec* tRNA alone (orange), and mixture of 2 µM *Rp* ProXp-abu AF488 + 0.3 µM *Rp* ProRS + 1 µM *Ec* tRNA complex (red) are shown. Corresponding species in each peak are indicated.

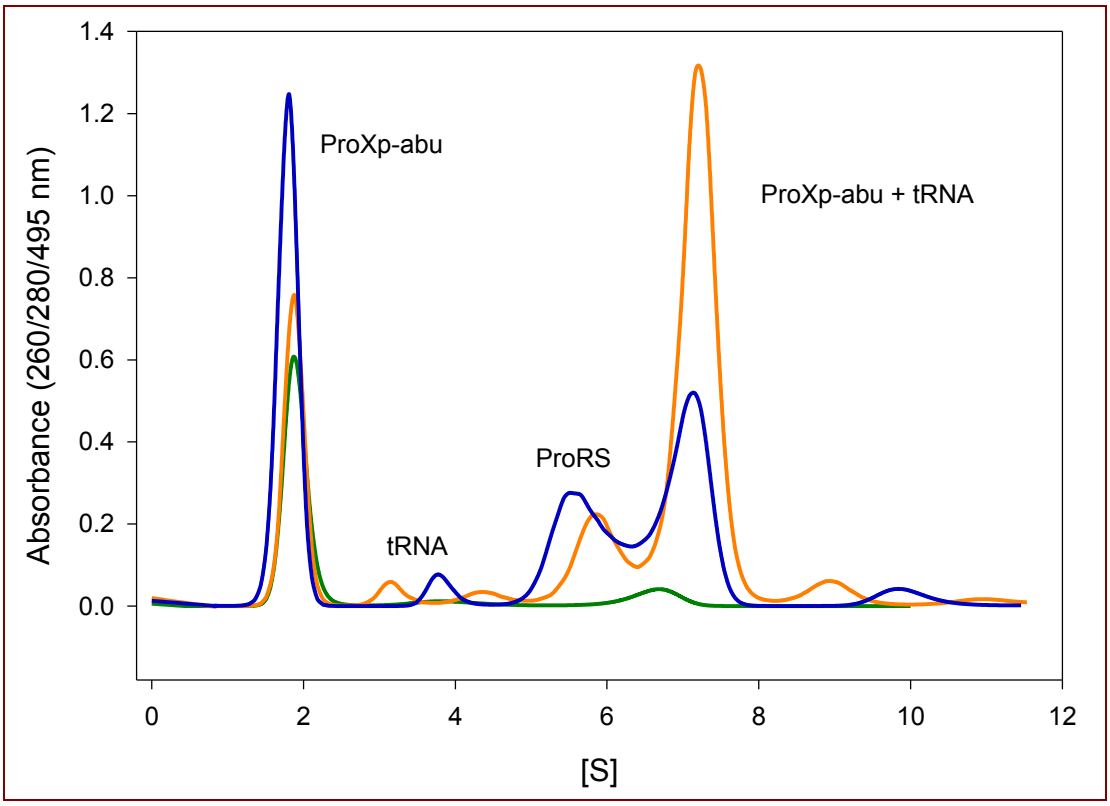


Figure 4.12. Sedimentation velocity profiles of *Rp* ProXp-abu AF488, *Rp* ProRS, and *Ec* tRNA interaction. Shown are the SV profiles of 2 µM *Rp* ProXp-abu AF488, 0.3 µM *Rp* ProRS, and 1 µM *Ec* tRNA complex collected at 260 nm (orange), 280 nm (blue), and 495 nm (green). Sedimentation coefficient ([S]) is plotted against respective absorbances (260 nm or 280 nm or 495 nm). Corresponding species in each peak are indicated.

Chapter 5 : Investigating the role of *Trans*-editing factor ProXp-abu in maintaining cellular homeostasis in *Rhodopseudomonas palustris*

5.1 Introduction

Aminoacyl-tRNA synthetases (aaRS) play a key role in protein synthesis by catalyzing the covalent attachment of the correct amino acid onto its cognate tRNA. aaRS are error prone due to their inability to distinguish between similarly sized amino acid substrates (1). In order to overcome this limitation, aaRS have evolved various editing mechanisms (pre-transfer and/or post-transfer) (3, 5). In the case of Prolyl-tRNA synthetase (ProRS), noncognate Ala and Cys are misactivated together with cognate Pro. In many bacterial systems such as in *Escherichia coli*, accurate synthesis of Pro-tRNA^{Pro} is ensured by a triple-sieve editing mechanism, which consists of the ProRS active site that discriminates amino acids based primarily on volume and size (coarse sieve), the ProRS editing domain (INS) that hydrolyzes Ala-tRNA^{Pro} in *cis* (fine sieve), and a single-domain INS homolog, YbaK that clears Cys-tRNA^{Pro} in *trans* via sulfhydryl cyclization of the thiol side chain (chemical sieve) (21, 66, 79). INS and YbaK belong to a family of 6 homologous, but functionally distinct, enzymes collectively called the INS Superfamily. Members of this family are widely distributed across all Domains of life and include ProXp-ala (22), ProXp-abu (Chapter 3), ProXp-y, and ProXp-z (21, 22). These editing factors are found in various species-specific combinations. For example,

Caulobacter crescentus encodes for a ProRS lacking a full-length INS domain and encodes for ProXp-ala and YbaK to clear Ala- and Cys-tRNA^{Pro} respectively (22).

Rhodopseudomonas palustris (*Rp*) is a gram negative, purple, non-sulfuric bacterium and belongs to the alpha proteobacteria (146). It is widely distributed in natures reflected by its isolation from diverse sources such as swine waste lagoons, industrial waste areas, marine coastal sediments, earthworm droppings, and pond water (146, 147). It is a metabolically versatile organism capable of several metabolic modes, including aerobic chemoheterotrophic, anaerobic photoheterotrophic, nitrogen fixing, carbon fixing, etc. It demonstrates exceptional flexibility among these metabolic modes and is able to grow with or without oxygen. It can utilize energy from light and organic compounds and is able to degrade lignin monomers, diverse fatty acids, and complex organic compounds such as benzoate, under both aerobic and anaerobic growth conditions (152). It generates hydrogen gas as a byproduct of nitrogen fixation, thereby acting as a potential biofuel source. *Rp* is capable of converting carbon dioxide into cell mass, making it a potential sink for greenhouse gases (147). Altogether, *Rp* presents an ideal model system for studying how the intricate web of cellular processes that operates within a single cell adapts in response to various stimuli (changes in light, carbon source, etc.) under diverse metabolic modes. *Rp* ProRS contains a mini-INS domain and the *Rp* genome also encodes two *trans*-editing factors YbaK and ProXp-abu. Although mini-INS is catalytically inactive, it is required for maintaining the overall structure of ProRS (Chapter 2). YbaK has been shown to clear mischarged Cys-tRNA^{Pro} while ProXp-abu shows robust editing activity against aminobutyrate (Abu)-tRNAs (Chapter 3).

Abu is an intermediate molecule in various cellular processes, such as amino acid catabolism and metabolism and is therefore generally present in the cell. The most common Abu biosynthetic pathway involves a series of deamination reactions in Thr metabolism producing Abu as a side-product (143, 144, 153) (Figure 5.1). Interest in Abu biosynthesis stemmed from the challenges presented by synthesizing enantiomerically pure compounds in high quantities. This led to exploitation of biological systems, which are generally efficient in enantiomeric discrimination, for large scale synthesis of compounds such as Abu (143, 154). Genetically engineered bacterial strains were generated to readily catalyze synthesis of Abu from extracellularly supplied Threonine. Aside from involvement in the Threonine biosynthetic pathway, Abu production in the cell is also closely related to the catabolic pathway of branched chain amino acids like Ile (155) where they share the same precursor molecule alpha-ketobutyrate (Figure 5.1). Interestingly, ValRS and IleRS recognize and misactivate Abu (29). Structurally, Abu is considered as an intermediate molecule between Ala and Val. It is larger than Ala by a methylene group and smaller than Val by the same functional unit. Consequently, Abu is recognized and mischarged by several aaRS.

Owing to the nonprotein nature of Abu and its close biosynthetic and structural relationship with several standard amino acids, in this chapter, we explore behavior of cellular levels of Abu and expression of ProXp-abu under various metabolic states. The baseline proteome expression of *Rp* grown under various metabolic modes was previously determined (147). This study revealed ProXp-abu is expressed only under nitrogen fixing conditions. Moreover, bioinformatics analysis of organisms encoding for

ProXp-abu reveals about 64% of the organisms that encode the combination mini-INS/YbaK/ProX are nitrogen fixing organisms (22). ProXp-abu is likely linked to the nitrogen fixation activity of these organisms. Here we establish the foundation and tools for studying this system in *Rp*, the relevant cellular pathways ProXp-abu may be involved in and the role it plays in maintaining proper cellular function.

5.2 Experimental Procedures

5.2.1 Materials

All materials were purchased from Sigma unless otherwise noted.

5.2.2. Media preparation

All *Rp* liquid cultures were grown in defined minimal media (156). Carbon sources were provided to a final concentration of 10 mM. Luria-Bertani (LB) media was prepared with the proportions 10 g tryptone, 5 g NaCl, 5 g yeast extract. Where appropriate, final concentrations of 10% sucrose and 50 mg/mL kanamycin were added.

5.2.3. *Rp* CGA009

Rp CGA009 was acquired from ATCC in lyophilized powder form. *Rp* was revived in liquid media and plated onto LB-agar for isolation of single colonies. Isolated colonies were transferred to liquid defined minimal medium and maintained in this form for future use. For growth studies, *Rp* was pre-grown for 24-30 h phototrophically

(anaerobic/light) in 5 mL minimal medium supplied with 10 mM succinate as sole carbon source and incubated under a light (3,000 lux, lx) at 30° C. Pregrown cultures were inoculated into anaerobic stoppered screw cap (Hungate) tubes containing 5 mL minimal medium supplemented with an appropriate carbon source. Growth was monitored at OD_{578}. For conjugation, *Rp* was grown aerobically in LB medium at 30° C in the dark.

5.2.4. Amino acid rescue studies

Defined minimal medium contained in anaerobic Hungate tubes was supplemented with 10 mM of the appropriate amino acid both in the presence and absence of 5 mM Abu. Succinate (10 mL) was provided as a carbon source and NH_4Cl as the nitrogen source. Pre-grown *Rp* culture was inoculated into each sample toa starting $O.D._{578}$ of ~ 0.03. Culture samples were grown anaerobically with light at 30° C and growth was monitored as O.D. at 578 nm over a course of ~60 hours.

5.2.5. Generation of *Rp* ProXp-abu and *Rp* YbaK Null Strains

The general scheme shown in Figure 5.2 was used to generate *Rp* ProXp-abu and YbaK null strains. Two fragments containing 1.0 - 1.5 kb flanking regions on either side of ProXp-abu/YbaK were amplified and cloned from *Rp* CGA009 genomic DNA (Figure 5.2.Step 1). For the ProXp-abu upstream region, primers 5'-AGTACCGGAATTCGATT GGTGCGCTACATCAAC and 5'-CGGCACGGTACCCGAACTCATTGTC were used, and for the downstream region, primers 5'-GAGCTGGGTACCGCCGAATGG GTCGATGTCTG and 5'-GAACTGGGATCCTCGGCCAACGAGATCCTCGGCAG

were used. For the YbaK upstream region, primers 5'-CTCAA CGAATTCAAGCTCGGGCTCGGCACTTTCAAG and 5'-GGGCACGGTACCCA GCGAGCGAACCCGGACTCTTG were used, and for the downstream region, primers 5'-GGCCAGCGCGGTACCCAGATCGAACTCGATCCGAGC and 5'-GGTC AGGG ATCCCGCGTAGCCGACGGCATGGTAG were used. Each of the fragments was cloned into pUC19 in the multiple cloning site (MCS) region using EcoRI and KpnI for the upstream fragments yielded pJMB01 (ProXp-abu) and pJMB05 (YbaK), and KpnI and BamHI for the downstream region yielded pJMB02 (ProXp-abu) and pJMB06 (YbaK) (Figure 5.2.Step 2). Plasmid sequences were verified through DNA sequencing using universal M13 primers provided by Genewiz. Upstream fragments were digested from pJMB01 and pJMB05 and cloned into pJMB02 and pJMB06 respectively, yielding pJMB03 (ProXp-abu) and pJMB07 (YbaK) (Figure 5.2.Step 3). Plasmids were verified with diagnostic digests using BglI, AclI, and EcoRI, which yield unique fragmentation patterns for each plasmid. Upstream and downstream regions in pJMB03 and pJMB07 were then digested with EcoRI and BamHI, and cloned into the suicide vector pK18mobSacB to yield pJMB04 (ProXp-abu) and pJMB08 (YbaK) respectively (Figure 5.2.Step 4). pJMB04 was verified using diagnostic digest with AccI while pJMB07 was verified using double digestion with AccI and NcoI. Plasmids were transformed into *E. coli* S17 and screened for Kan resistance and sucrose sensitivity. Plasmids were transferred into R. palustris by conjugation. ProXp-abu single crossovers were selected and verified using primers 5'-GCGGCATGGGCATGCGGCTGAAGTATTCCAACC and 5'-GCTGCTGGTGGCGTCCGAACTACTATGGCGCTTAC for upstream single

crossover, and using primers 5'-AAGCGTTCGGCTTCAACGTCGCCTGGATCG and 5'-CATTCAACGCCAACACCTCCGAGTCCGAGAC for downstream single crossover (Figure 5.3). YbaK single crossovers were selected and verified using primers 5'-CGCATTCGCATTCACGATGCGTTGACCTATCC and 5'-ACAGCGAGGTCAGG CTGAGCAGCAGATAGGCTTGG for upstream single crossover, and using primers 5'-GGCTGATCGCTGCAACACAGCCGGGCACGTATTAG and 5'-ATGATGCGGTCG CCGGAATATCTCAGCCAGTGC for downstream single crossover. Single crossovers were inoculated into 100 uL of minimal media and incubated at 30 C in the dark for 16-20 hours. 100x and 1000x dilutions were plated onto minimal media-sucrose plates and incubated anaerobically for 7-10 days with light. Double crossover was selected for by patching onto minimal media-kanamycin and minimal media-sucrose plates, and confirmed by the sequencing of the appropriate chromosomal region using primers 5'-CATTCAACGCCAACACCTCCGAGTCCGAGAC and 5'-GCTGCTGGTGGCGTCCG AACTACTATGGCGCTTAC for ProXp-abu, and using primers 5'-GGCT GATCGCTGCAACACAGCCGGGCACGTATTAG and 5'-ACAGCGAGGTCAGGCT GAGCAGCAGATAGGCTTGG for YbaK.

5.2.6. Extraction of Intracellular Metabolites

Rp was grown anaerobically in light at 30 °C in defined minimal media. Succinate (10 mM) was supplied as the carbon source and NH_4Cl as the nitrogen source or N_2 gas for nitrogen fixing conditions. Cultures were grown to an $O.D._{580}$ of 0.6-0.8. Cells were harvested and washed 3 times with cold distilled water. Extraction of metabolites was

accomplished through boiling of cells resuspended in 5 mL water for a total of 10 min with vortexing for 30 sec after the first 5 min of boiling. Lysates were filtered through 0.22 uM filter to remove cell debris. Samples were freeze dried, lyophilized, and resuspended in 50 or 100 uL water for anaylsis on LC-MS/MS.

5.2.7 Liquid Chromatography-Tandem Mass Spectrometry

Separation and detection of metabolites was achieved using a Hypercarb column (100×2.1 mm, 5 μm pore, Thermo Fisher Scientific) and a triple-quadrupole QTRAP 5500 (AB Sciex) in positive ion mode (Figure 5.4). Multiple reaction monitoring (MRM) mode was used to allow simultaneous monitoring of all metabolites. For each amino acid and Abu, commercial standards were individually injected to determine the optimal mass spectrometric parameters.

5.3 Results and Discussion

5.3.1. Amino acid rescue studies

We tested the sensitivity of *Rp* to Abu by exposing liquid cultures to various concentrations of Abu. Cells showed an increase in doubling time as the concentration of Abu increased (

Table 5.1). This indicates Abu is taken up by cells. Conversely, we observe no effect on the doubling time upon exposure to elevated concentrations of Ala up to 10 mM.

We hypothesize that the sensitivity to Abu (but not Ala) may be due to misincorporation of Abu into the proteome. To test this hypothsis, we performed amino acid rescue experiments to determine which codons may be susceptible to Abu misincorporation. We chose to test Pro, Val, and Ile because Abu has been shown to be mischarged onto their respective tRNAs. We also tested Thr, a precursor to Abu biosynthesis. We first measure growth rate in the presence of excess of a specific amino acid to assess toxicity. Addition of Val showed some toxicity to the cell. This is likely due to feedback inhibition when Val is supplemented as free amino acid (155). Addition of Thr arrested growth of *Rp*. This lethal effect may be partially caused by bioconversion of Thr to Abu. Further investigation on the lethality of Thr in *Rp* needs to be conducted to test our hypothesis. Ile (10 mM) alone did not result in a change in doubling time and Pro had a minor effect at 10 mM. None of the tested amino acids rescued cell growth in the presence of 5 mM Abu. Rates of amino acid cell uptake may have affected the results. It is possible Abu is taken up faster than the introduced standard amino acids.

5.3.2. Quantificaton of Intracellular Abu in *Rp*

To quantify cellular concentrations of free Abu in *Rp* under various metabolic modes, we isolate intracellular metabolites for LC-MS/MS MRM analysis as detailed in the experimental procedures. Previous studies show expression of ProXp-abu under nitrogen fixing conditions. Moreover, ~64% of organisms that encode for ProXp-abu are nitrogen fixing organisms. We hypothesize that there is a link between ProXp-abu expression and Abu concentration in the cell. To investigate whether Abu concentration

is elevated under the same conditions upregulation of ProXp-abu is observed, we compared Abu cellular concentrations of *Rp* grown under its preferred photoheterotrophic conditions (anaerobic provided with light, supplemented with succinate as carbon source and NH_4Cl as nitrogen source), and under nitrogen fixing (anaerobic provided with light, supplied with succinate as carbon source, and N_2 gas as nitrogen source) growth conditions. Additionally, cultures grown in the presence and absence of 0.2 mM Abu for both growth conditions were also prepared. As a positive control, we prepared a sample spiked with 10 μM Abu immediately before extraction of cellular metabolites. Two aminobutyrate isomers are generally present in the cell, γ-aminobutyrate (GABA) and α-aminobutyrate (Abu), with the former found in higher abundance. Aminobutyrate isomers were successfully resolved through LC-MS/MS as shown inFigure 5.5.

Ion chromatograms were extracted for each sample at 104/58 Da parent/daughter *m/z* ion transition which corresponds to Abu, determined by analysis of commercially available Abu standard. Standard Abu elutes at ~1.55 min (Figure 5.6.A) while control sample spiked with Abu prior to cell lysis shows Abu standard peak slightly shift to ~1.4 min (Figure 5.6.B.). Analysis of metabolites extracted from *Rp* cultures grown under photoheterotrophic (Figure 5.6.C) and nitrogen fixing conditions (Figure 5.6.E) do not show an elution peak at ~1.55 min, indicating Abu is not present in detectable amounts under these growth conditions. Similarly, metabolites extracted from cultures grown in the presence of Abu do not show detectable amounts of Abu (Figure 5.6.D and F) suggesting Abu is not taken up by the cell. However, amino acid toxicity experiments clearly show that Abu cell uptake occurs as evidenced by the observed growth defect of

Rp upon exposure to Abu. Likely, Abu is taken is and readily utilized in various cellular functions therefore no accumulation is observed.

We also monitored standard amino acid levels under photoheterotrophic and nitrogen fixing conditions. We inspected the relative levels of the three amino acids misactivated by ProRS, Ala, Cys and Pro as well as Ile, Val, and Thr. None of the amino acid levels varied significantly under these two growh conditions, except for Ala, which increased 2-fold under nitrogen fixing conditions (Figure 5.7).

5.3.3. *Rp* null strains

Rp ProXp-abu and YbaK null strains were generated using an unmarked clean deletion strategy, which deletes ~80% of the gene leaving short, in-frame upstream and downstream fragments to avoid disrupting neighboring genes. Successful generation of ProXp-abu and YbaK null strains indicates these genes are nonessential; however, they may be required under certain growth conditions. For example, an *Ec* YbaK null strain supports growth under normal conditions but suffers growth defects when exposed to increased levels of Cys under oxidative stress (Ziwei Liu and Karin Musier-Forsyth, unpublished data). In some organisms, mistranslation due to loss of editing function is well tolerated. *Ec* is able to tolerate the presence of about 10% of mistranslated proteins by triggering cellular responses like heat-shock, thereby reducing overall growth rate. On the other hand, Mycoplasma parasites enjoy phenotypic plasticity and exploit proteomic randomization to counter host defense mechanisms. Insights into the degree of mistranslation tolerated by *Rp* and the resulting cell responses triggered (if any) should be

investigated. Generation of these null strains provides the fundamental tools for further investigation of the significance of ProXp-abu to protein synthesis *in vivo*, and for the exploration of other cellular functions where ProXp-abu may be recruited.

5.4 Conclusions

In this study, we have shown that *Rp* suffers a growth defect when exposed to elevated levels of Abu, reflected in the increased doubling times. In contrast, high concentrations of Ala do not affect cell growth. This is in agreement with our in vitro experiments showing that *Rp* ProRS system is able to efficiently discriminate against Ala despite loss of the INS domain. Interestingly, Thr, a precursor molecule to Abu biosynthesis, is lethal to *Rp*. Targeted metabolomics shows no detectable amounts of free Abu in cell lysates grown under photoheterotrophic and nitrogen fixing conditions. It is possible Abu is immediately utilized in cellular functions upon uptake; thus, no significant accumulation is observed. A survey of organisms encoding for ProXp-abu reveals a significant number are metabolically flexible organisms; about 80% of the population are anaerobic organisms, including obligate anaerobes. Furthermore, about 64% are nitrogen fixing organisms and ~25% are photosynthetic bacteria. Interestingly, several highly pathogenic species, such as *Salmonella enterica* and *Klebsiella pneumonia*, encode for ProXp-abu. This emphasizes the need to further study the dynamic metabolic capabilities of these organisms, how this dynamism relates to ProXp-abu regulation, and its impact on protein synthesis and other core cellular functions.

5.5 Acknowledgements

We would like to thank our collaborator Dr. Birgit Alber for providing us with the necessary supplies and experimental set-up for *Rp* cultures. We would also like to thank Dr. Ana Alonso and Dr. Jean-Christophe Cocuron for use of the Targeted Metabolomics facilities.

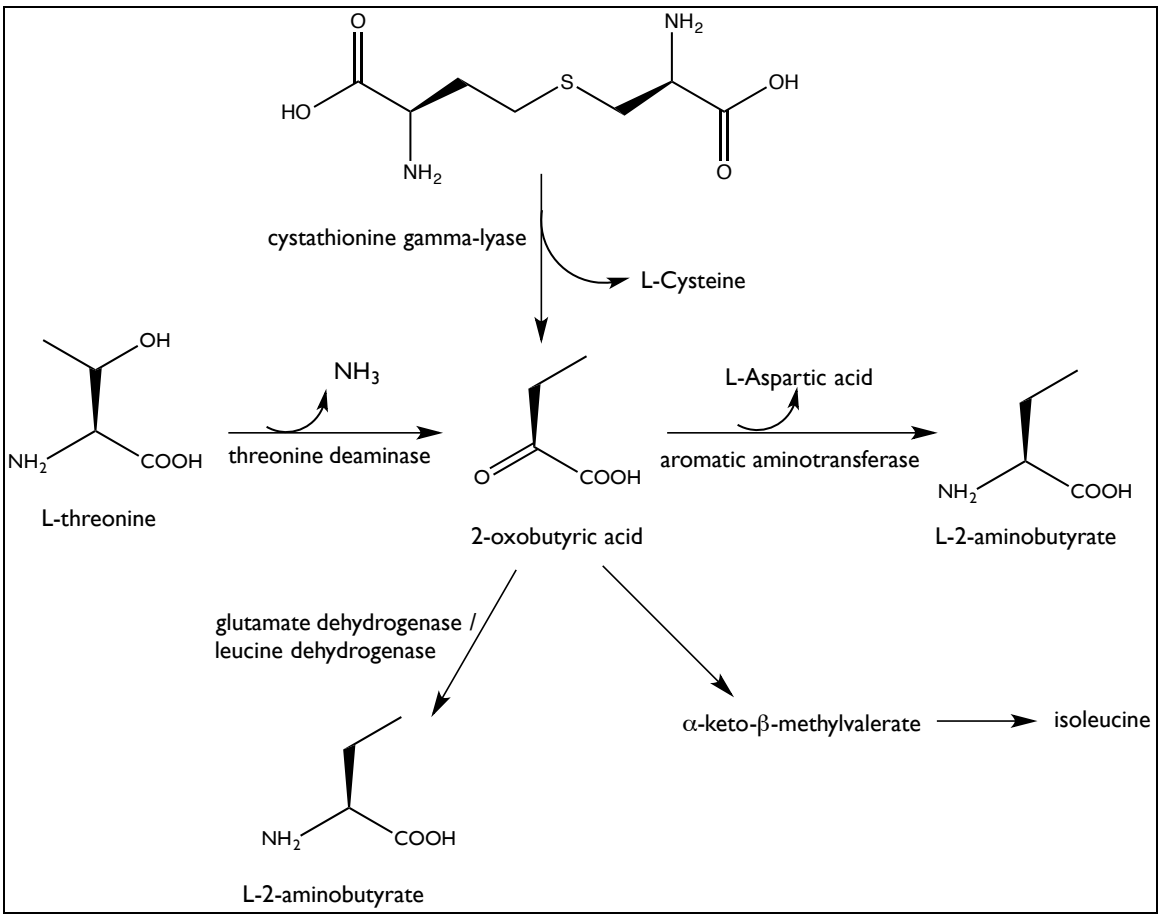


Figure 5.1. Abu biosynthesis scheme.

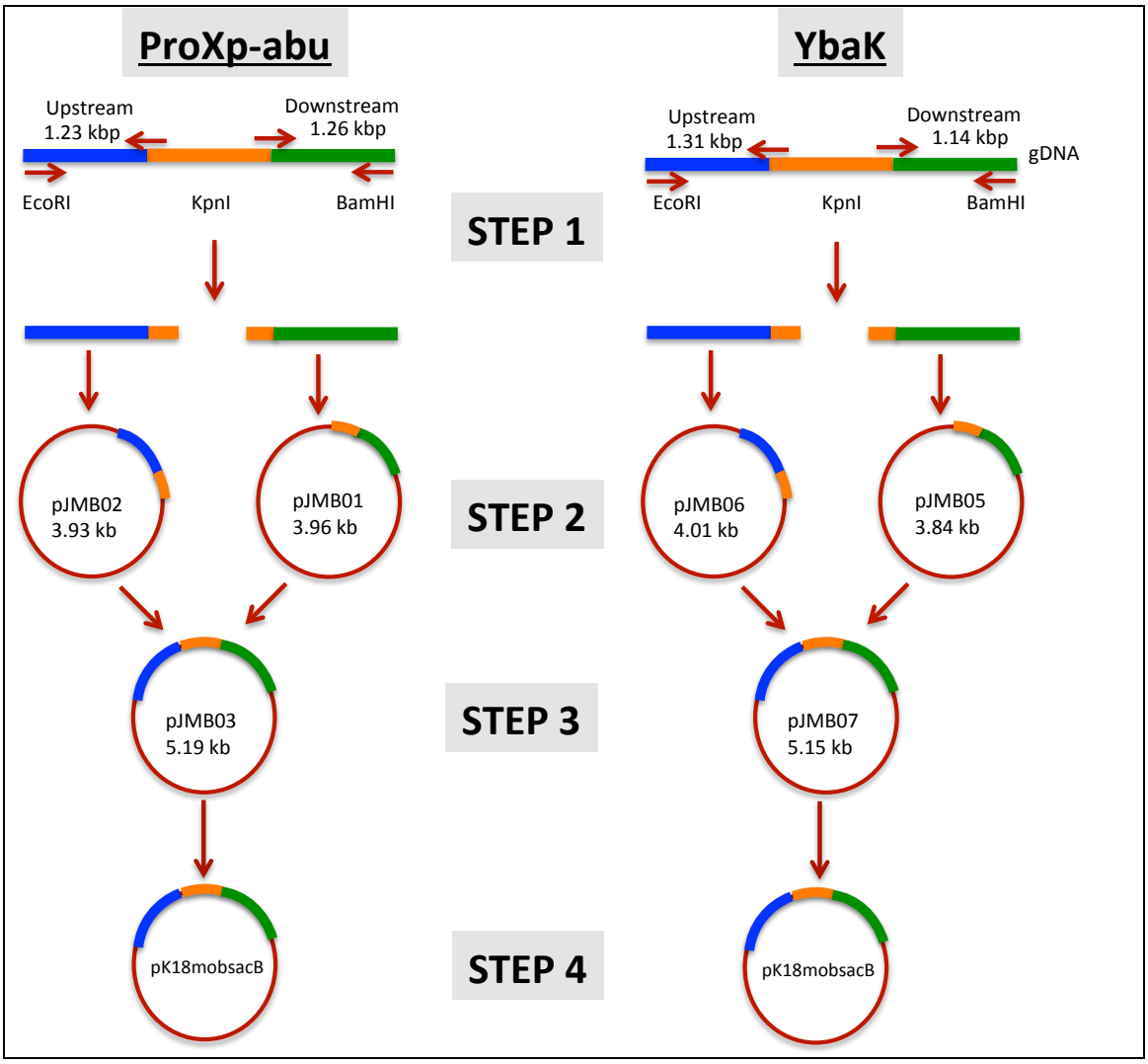


Figure 5.2. General scheme for generation of *Rp* ProXp-abu and YbaK null strains. Gene of interest is depicted in orange. Upstream and downstream products are depicted in blue and green respectively. Fragment and plasmid sizes are indicated.

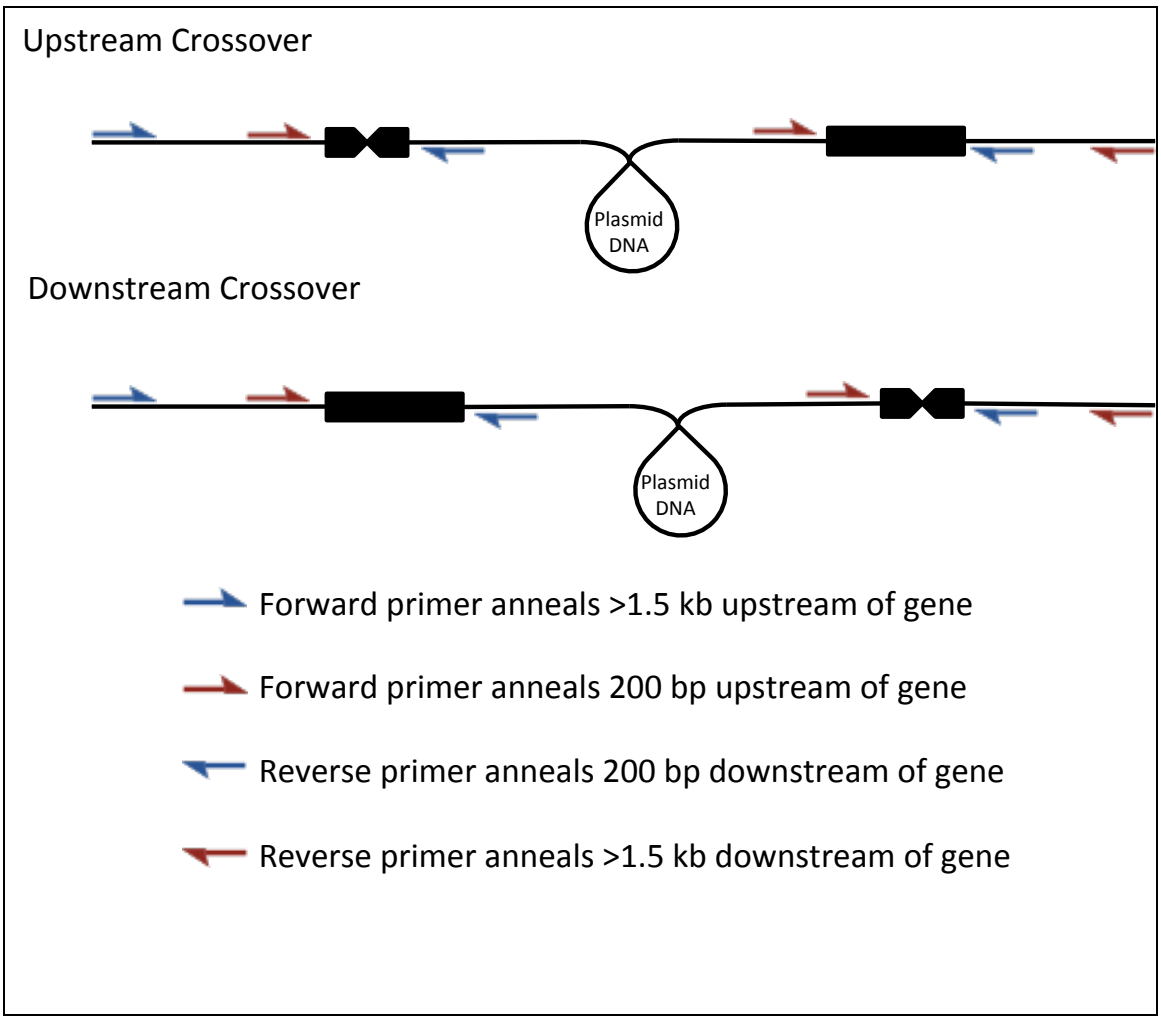


Figure 5.3. Strategy for genotyping crossover events.Upstream and downstream single crossover events are depicted. Primer pairs are depicted in identical colors. Regions where primers anneal are indicated.

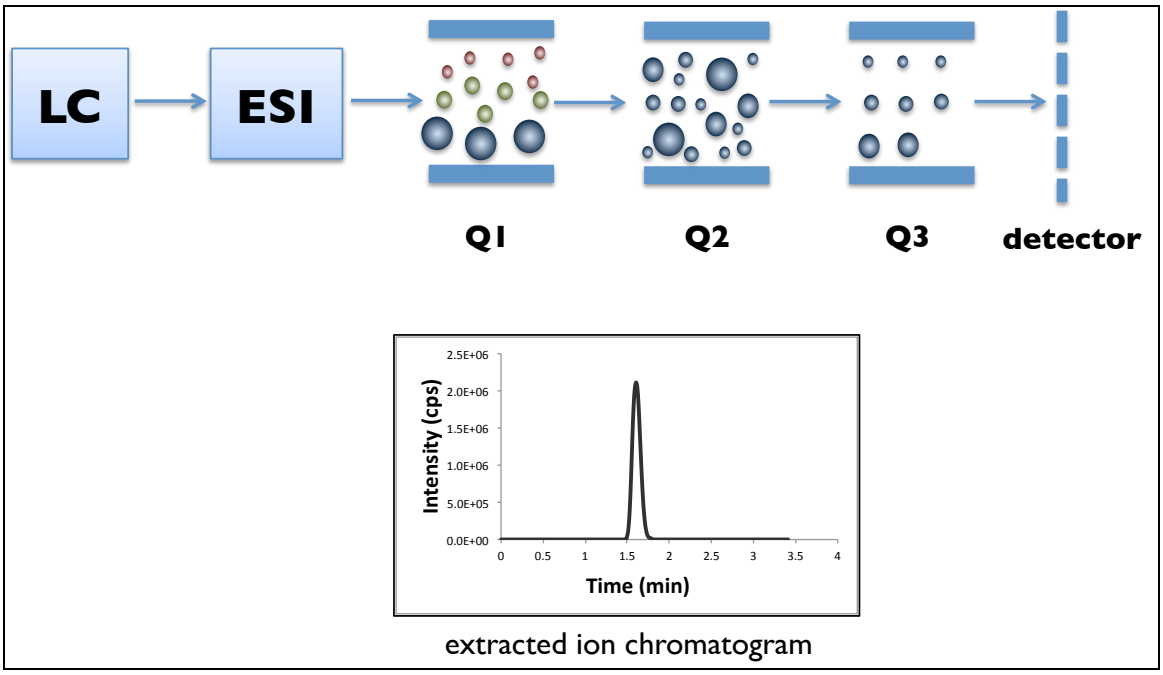


Figure 5.4. Schematic of LC-MS/MS MRM. Metabolites are initially resolved using liquid chromatography (LC) followed by conversion of compounds from liquid phase into gas-phase ions by electrospray ionization (ESI). Generated ions are resolved through triple quadrupole MS where ions are subjected to two mass filtration steps. In the first quadrupole (Q1) parent *m/z* of interest are selected. The second quadrupole (Q2) is a collision cell where selected ions from Q1 are fragmented into daughter ions. Fragment *m/z* are resolved in the third quadrupole (Q3). Data is reported as ion-specific chromatograms where retention time is plotted against intensity (ion count per second) for a specific MRM scan (parent/daughter *m/z* transition) (157).

Table 5.1. Doubling times of *Rp* in the absence and presence of 5 mM Abu. Here shown are the amino acids supplemented in minimal media and their respective growth rates are reported in doubling times (hours).

Growth Condition	Doubling time (hours)
No amino acid	8
1 mM Abu	9
5 mM Abu	11
5 mM Ala	8
10 mM Ala	8
10 mM Pro	9
10 mM Pro + 5 mM Abu	14
10 mM Val	10
10 mM Val + 5 mM Abu	14
10 mM Ile	8
10 mM Ile + 5 mM Abu	11
10 mM Thr	lethal
10 mM Thr + 5 mM Abu	lethal

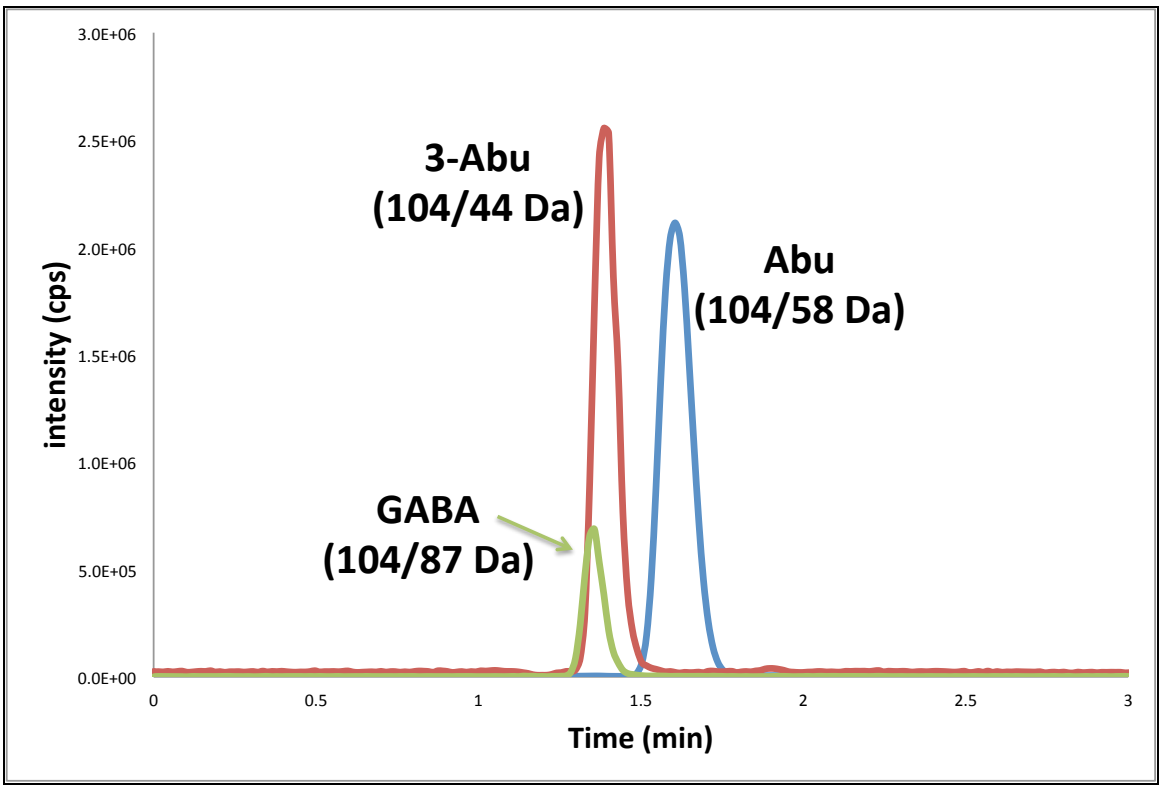
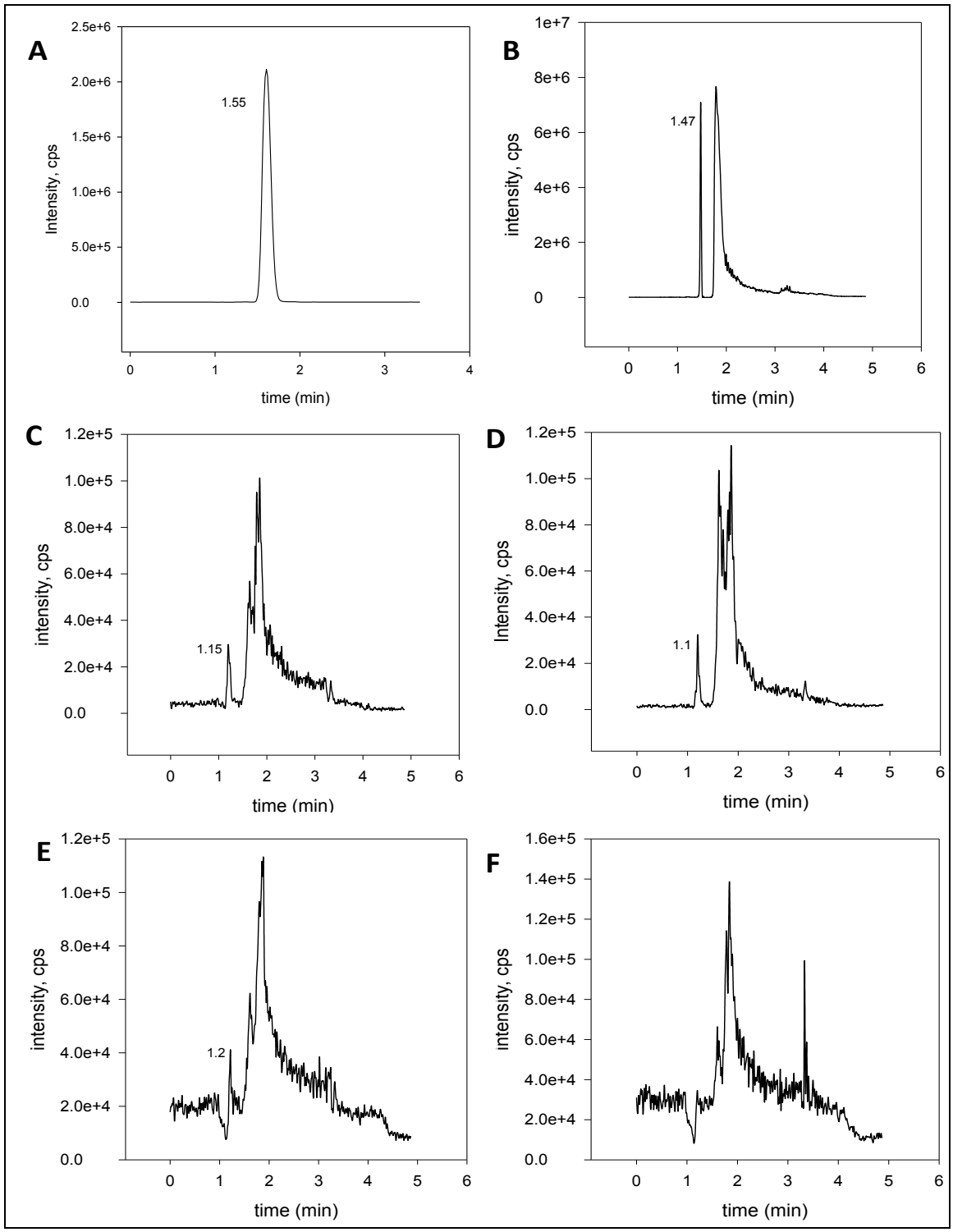


Figure 5.5.LC-MS/MS resolution of aminobutyrate isomers. Here shown is an overlay of extracted ion chromatograms of aminobutyrate isomers: α-aminobutyrate/Abu (blue), β-aminobutyrate/3-Abu (red), and γ-aminobutyrate/GABA (green). Retention time is plotted against intensity (counts per second). Individual parent/daughter *m/z* ion transitions are indicated.

Figure 5.6. Ion Chromatograms extracted at 104/58 Da parent/daughter *m/z* transition. Shown are the extracted ion chromatograms of (A) Abu standard, (B) control sample spiked with 10 μM Abu prior to cell lysis, and intracellular metabolites extracted from *Rp* grown under (C) photoheterotrophic, (D) photoheterotrophic + 0.2 mM Abu, (E) nitrogen fixing, (F) nitrogen fixing + 0.2 mM Abu growth conditions.

Figure 5.6


A
Intensity, cps
2.5e+6
2.0e+6
1.5e+6
1.0e+6
5.0e+5
0.0
1.55
0
1
2
3
4
time (min)
B
intensity, cps
1e+7
8e+6
6e+6
4e+6
2e+6
0
1.47
0
1
2
3
4
5
6
time (min)
C
intensity, cps
1.2e+5
1.0e+5
8.0e+4
6.0e+4
4.0e+4
2.0e+4
0.0
1.15
0
1
2
3
4
5
6
time (min)
D
Intensity, cps
1.2e+5
1.0e+5
8.0e+4
6.0e+4
4.0e+4
2.0e+4
0.0
1.1
0
1
2
3
4
5
6
time (min)
E
intensity, cps
1.2e+5
1.0e+5
8.0e+4
6.0e+4
4.0e+4
2.0e+4
0.0
1.2
0
1
2
3
4
5
6
time (min)
F
intensity, cps
1.6e+5
1.4e+5
1.2e+5
1.0e+5
8.0e+4
6.0e+4
4.0e+4
2.0e+4
0.0
0
1
2
3
4
5
6
time (min)

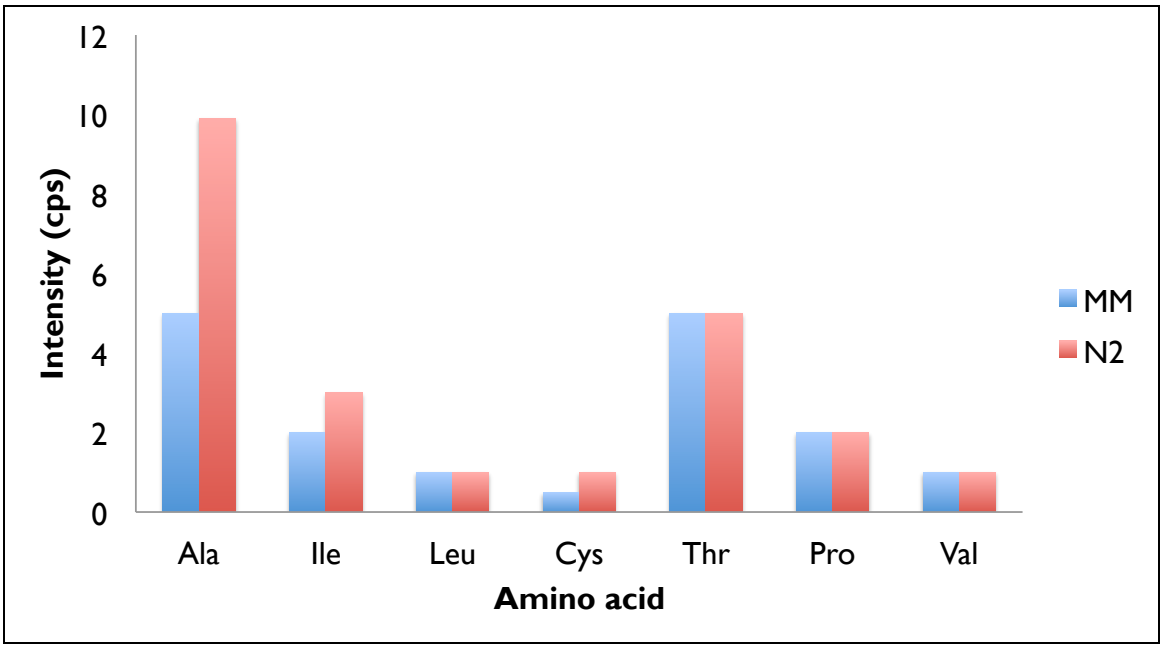


Figure 5.7. Relative amino acid levels in *Rp*. Grown in defined mineral media (MM, blue) and under nitrogen fixing conditions (N_2, red).

Chapter 6 : Conclusions and Future Directions

6.1 Conclusions

Accurate translation of the genetic code from DNA to protein is integral to maintaining cellular homeostasis. Aminoacyl-tRNA synthetases (aaRS) play a pivotal role in protein synthesis by catalyzing the formation of aminoacyl-tRNAs (aa-tRNA) for incorporation into the growing peptide in the ribosome. However, aaRS are error prone due to the similarities in size and volume of their amino acid substrates. For example, ProRS misactivates Ala and Cys in addition to Pro. Due to the high propensity of noncognate amino acid misactivation, aaRS have evolved editing mechanisms (pre-transfer or post-transfer) that correct errors in aa-tRNA synthesis. In ProRS, a triple sieve editing mechanism is employed to ensure accuracy in Pro codon translation. The aminoacylation synthetic core acts as the coarse sieve by rejecting larger amino acids through steric exclusion but allows misactivation of Ala and Cys in addition to Pro leading to synthesis of incorrect Ala- and Cys-tRNA^{Pro}. A family of related proteins (INS superfamily), widespread through all taxonomic domains of life, is employed in species-specific combinations to clear Ala- and Cys-tRNA^{Pro}. In *Escherichia coli* (*Ec*), a combination of *cis* (INS domain) and *trans* (YbaK) editing is used to edit Ala- and Cys-tRNA^{Pro} respectively. Comparatively, *Caulobacter crescentus* (*Cc*) lacks a full-length

INS domain (mini-INS) and uses two, related but functionally distinct, *trans*-editing factors ProXp-ala (Ala-tRNA^{Pro}) and YbaK (Cys-tRNA^{Pro}) instead.

In the case of *Rhodopseudomonas palustris* (*Rp*), it encodes for a mini-INS, similar to *Cc*, along with two *trans*-editing factors YbaK and ProXp-abu. The function of the latter was unknown. In this study, we establish the *in vitro* activity of ProRS and related editing domains from *Rp* and gain insight into the critical roles these enzymes play in ensuring translational fidelity.

In Chapter 2, we characterize the significance of the conservation of the mini-INS and determine its relevance to the structure and function of *Rp* ProRS. We show that the mini-INS plays a critical role in preserving overall structure and function of *Rp* ProRS by stabilizing a hydrophobic patch posterior to the aminoacylation domain. This observation is reflected in the demonstrated sensitivity of *Rp* ProRS to any mutational perturbations on the mini-INS. We show that the mini-INS lacks catalytic activity and is mainly conserved for structural purposes. The loss of a catalytically active *cis*-editing domain has forced the system to evolve a more specialized synthetic core with enhanced substrate specificity, reflected in the elevated K_M and robust pre-transfer editing of Ala. Interestingly, *Rp* ProRS has relaxed discrimination against Abu and misactivates it with an error rate of ~1/1000 relative to Pro, which suggests editing may be required *in vivo*.

In chapter 3, we functionally characterize *trans*-editing factors YbaK and ProXp-abu. As expected, *Rp* YbaK deacylates Cys-tRNA^{Pro} consistent with YbaK isolated from other species (*Ec, Haemophilus influenzae, Cc*) (21, 22, 79). Here we show that although ProXp-abu shares structural and sequence similarities to INS, ProXp-ala, and YbaK, it is

functionally distinct, evidenced by its lack of editing activity against Ala- and Cys-tRNA^{Pro}. Structural analysis reveals the ProXp-abu putative substrate binding pocket is larger compared to INS, suggesting substrates larger than Ala are preferred. Indeed, ProXp-abu edits Abu-tRNA^{Pro}. Further screening against similar sized proteinogenic aminoacyl-tRNA substrates affirms Abu-tRNA is the preferred substrate with moderate activity for nonpolar Val- and Ile-tRNA. Evaluation of substrate binding energies using Molecular mechanics Poisson-Boltzmann surface area (MMPBSA) calculations corroborates the observed trend in substrate specificity *in vitro*.

We investigated the molecular basis for substrate discrimination and determine tRNA acceptor stem recognition elements in Chapter 4. We show that the substrate binding pockets of ProXp-abu and the INS domain share similarities with regards to the degree of hydrophobicity and conservation of elements necessary for catalysis and proper orientation of the aa-tRNA substrate, such as the conserved Lys (K50 in ProXp-abu and K279 in INS domain) and GXXXP loop, but differ in their mode of substrate discrimination. Mutational analysis of the ProXp-abu active site pocket shows its substrate specificity is not tunable, suggesting that in contrast to INS, the basis for substrate discrimination of ProXp-abu is not dictated by steric properties. Moreover, comparison of Ala-CCA docked in INS and Abu-CCA docked in ProXp-abu, reveals that the amino acid substrates are bound in opposite orientations thus defining two distinct and non-superimposable substrate binding pockets. Taken altogether, our observations suggest ProXp-abu is mechanistically distinct from INS despite similarities in their active sites.

In addition to discrimination of the amino acid moiety, elements on the tRNA also contribute to correct aminoacyl-tRNA recognition. Our work shows Abu mischarged onto tRNAs specific for Pro, Val, and Ile are deacylated by ProXp-abu with similar rates. All three tRNAs share the same discriminator base (A73), but contain different first base pairs, which suggests the first base pair on the acceptor stem is not a major recognition element for ProXp-abu. Mutation of A73 to C shows a moderate decrease in editing activity, indicating the discriminator base is a weak determinant for tRNA recognition by ProXp-abu. In *Rp*, over 50% of tRNAs encode for A73 thus ProXp-abu appears to act as a general Abu deacylase. Owing to the nonprotein nature of Abu, the relaxed tRNA discrimination of ProXp-abu is advantageous to the cell.

In chapter 5, we investigate the role of ProXp-abu *in vivo*. We show that growth of *Rp* is not affected under elevated concentrations of Ala, but is highly sensitive to Thr, a precursor molecule in Abu biosynthesis. Targeted metabolomics studies were unable to detect Abu in cell lysate. It is likely Abu is readily processed in various cellular processes and therefore does not accumulate in its free form. Phylogenetic analysis reveals organisms that encode for ProXp-abu are capable of anaerobic metabolic growth mode with a majority (64%) capable of nitrogen fixation. Moreover, analysis of the baseline proteome of *Rp* under various metabolic modes show ProXp-abu expressed under nitrogen fixing conditions. *Rp* ProXp-abu and YbaK null strains have been generated, thereby providing the fundamental tools necessary in further investigations of the function of ProXp-abu *in vivo*. Interestingly, highly pathogenic species such as *Salmonella enterica* and *Klebsiella pneumonia* encode for ProXp-abu, highlighting the

significance of investigating its role *in vivo*, which could potentially lead to the basis for design of novel antimicrobials.

6.2 Future Directions

In chapter 4, we investigated the molecular basis for ProXp-abu substrate discrimination by performing mutations on the active site of *Rp* ProXp-abu guided by the MD-simulated model of Abu-CCA docked into the putative binding pocket of *Tt* ProXp-abu (Figure 3.5). We will confirm the identified putative substrate binding pocket by performing mutations on residues that constitute the active site: T31A, A129G, I48A, and A37G. Additionally, we will perform NMR studies to definitively map out the ProXp-abu binding pocket. We will use a stably-mischarged Abu-microhelix from Ronald Micura (University of Innsbruck) (158) to look at chemical shift perturbations of both bound and *apo* forms of ProXp-abu.

Our results in chapter 4 show ProXp-abu active site is not tunable. Moreover, MD modeling of Abu-CCA bound to ProXp-abu binding site shows no potentially catalytic water within the vicinity of the amino acid substrate. This suggests ProXp-abu may use a deacylase mechanism distinct from that of INS, which uses water-mediated hydrolysis. We will confirm these observations by performing competition assays using Ala-tRNA^{Pro} as the Abu-tRNA^{Pro} competitor. MMPBSA calculated binding free energies (Chapter 3) show Ala binds only ~3.2 kcal less relative to Abu, therefore is likely to compete with Abu for binding in the ProXp-abu active site pocket. We will titrate Ala-tRNA^{Pro} in various concentrations and observe for an Ala-tRNA^{Pro} concentration-dependent

reduction of Abu-tRNA^{Pro} deacylase activity. Successful inhibition of Abu-tRNA^{Pro} deacylation by Ala-tRNA^{Pro} indicates water-mediated hydrolysis is probably not the deacylase mechanism employed in ProXp-abu.

We also showed that ProXp-abu has a relaxed specificity for tRNA and acts as a general deacylase. It is likely ProXp-abu gains specificity through its interaction partners (RNA or protein). AUC experiments show ProXp-abu interacts with tRNA^{Pro} but to date, no evidence for ProRS interaction has been obtained. We will characterize possible complex formation with ValRS and IleRS using AUC. Additionally, we will explore possible interacting partners *in vivo*. We have generated polyclonal antibodies (Ab) specific for *Rp* ProXp-abu, *Rp* YbaK, and *Rp* ProRS (Figure 6.1). We will use the Abs for co-immunoprecipitation (co-IP) and pull-down assays to detect RNA and protein interaction partners. One of the caveats of this preoject is the ability to identify an interacting partner among the potentially wide variety and large number of RNA and protein species yielded from co-IP and pull-down assays. To overcome this, we will initially look at aaRS and tRNA binding partners due to the relatedness of these editing factors to the translational machinery. We will identify aaRS partners through western blot using commercially available Abs and identify possible tRNA partners using RNA sequencing. Information gathered from this project will contribute vital information to ProXp-abu recruitment and involvement in various cellular pathways *in vivo*.

Elongation factor thermo unstable (Ef-TU) plays a critical role in protein synthesis by binding and escorting the newly synthesized aa-tRNAs to the ribosome for incorporation into the growing peptide. EF-TU has been shown to discriminate based on

both amino acid and tRNA elements (150). Due to the nonprotein nature of Abu, it is vital to determine whether Abu-tRNA is recognized by Ef-TU, and how this recognition and binding of misacylated Abu-tRNA compares to that of correctly synthesized aa-tRNAs. Previous studies suggest that Abu-tRNA is recognized by Ef-TU and delivered into the ribosome for incorporation into the proteome. An *Ec* strain harboring an editing defective ValRS T222P mutant has been shown to misincorporate Abu into the proteome when supplemented with Abu in the culture media (29). Since EF-TU recognizes the tRNA moiety as well, we will conduct similar experiments probing Abu misincorporation in Pro and Ile codons using *Ec* strains harboring editing-defective ProRS K279A and IleRS T242P mutation respectively. Additionally, Ef-TU binding experiments will be conducted *in vitro* to compare Ef-TU binding to Abu-tRNA and cognate aa-tRNAs. This project will contribute information to the challenges Abu presents for accurate translation and potentially highlight the requirement of a specialized editing function.

In chapter 5, we generated ProXp-abu null *Rp* strain (*Rp ΔProX)*. All in vitro data support Abu-tRNA as the preferred substrate. We would like to confirm this through *in vivo* experiments by qualifying sensitivity of and *Rp ΔProX* to the various amino acid substrates tested for deacylation in Chapter 3.3.5. To accomplish this, we will perform halo assays (159), as well as establish growth profiles under varying amino acid concentrations. Pursuant to our *in vitro* results, we expect *Rp ΔProX* to show stronger sensitivity to Abu compared to WT *Rp* (*Rp* $ProX^{+}$). Other tested amino acids yielding similar results present candidates for susbtrates of ProXp-abu *in vivo*.

To complement amino acid sensitivity experiments, we will also perform proteomic analysis (peptide digest followed by LC-MS/MS) to quantify degree of misincorporation. We expect no misincorporation in *Rp ProX*$^{+}$, significant misincorporation in *Rp ΔProX*, and no misincorporation in *Rp ΔProX* complemented with ProXp-abu expressed on a vector.

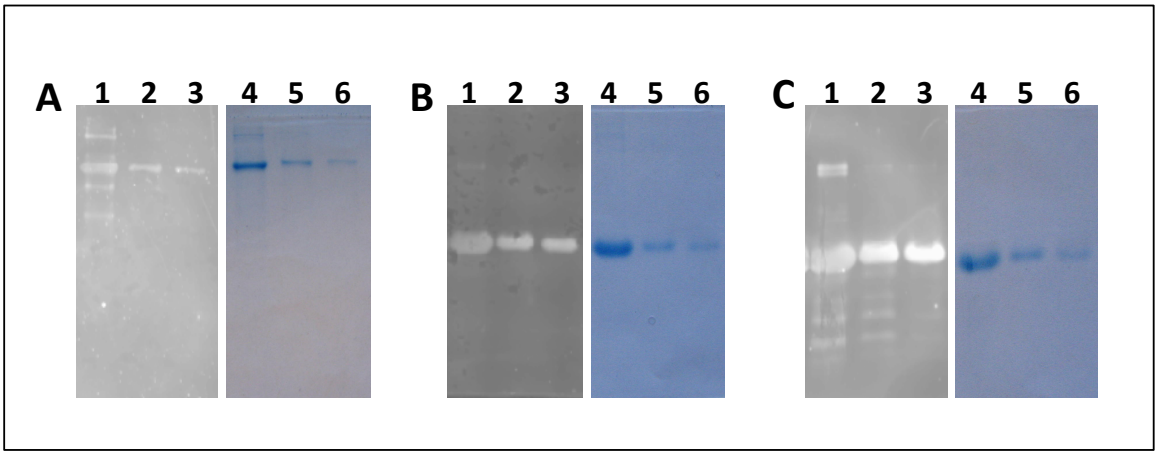


Figure 6.1. Immunoblot and coomasie stain visualized gels of *Rp* ProRS, *Rp* ProXp-abu, and *Rp* YbaK. Shown are (A) 10 µg (lane 1), 5 µg (lane 2), and 1 µg (lane 3) of *Rp* ProRS probed with 1000x diluted *Rp* ProRS polyclonal Abs (1° Ab) and visualized with Goat Anti-rabbit (Cy5, 2° Ab) and 10 µg (lane 4), 5 µg (lane 5), and 1 µg (lane 6) of *Rp* ProRS visualized with coomasie stain. (B) 10 µg (lane 1), 5 µg (lane 2), and 1 µg (lane 3) of *Rp* ProXp-abu probed with 1000x diluted *Rp* ProXp-abu polyclonal Abs (1° Ab) and visualized with Goat Anti-rabbit (Cy5, 2° Ab) and 10 µg (lane 4), 5 µg (lane 5), and 1 µg (lane 6) of *Rp* ProXp-abu visualized with coomasie stain. (C) 10 µg (lane 1), 5 µg (lane 2), and 1 µg (lane 3) of *Rp* YbaK probed with 1000x diluted *Rp* YbaK polyclonal Abs (1° Ab) and visualized with Goat Anti-rabbit (Cy5, 2° Ab) and 10 µg (lane 4), 5 µg (lane 5), and 1 µg (lane 6) of *Rp* YbaK visualized with coomasie stain.

LIST OF REFERENCES

1. Ibba M, Söll D (2000) Aminoacyl-tRNA synthesis. *Annu Rev Biochem* 69:617–650.

2. Yadavalli SS, Ibba M (2012) in *Advances in Protein Chemistry and Structural Biology*, ed Assen Marintchev (Academic Press), pp 1–43.

3. Jakubowski H (2012) Quality control in tRNA charging. *Wiley Interdiscip Rev RNA* 3:295–310.

4. Ling J, Reynolds N, Ibba M (2009) Aminoacyl-tRNA synthesis and translational quality control. *Annu Rev Microbiol* 63:61–78.

5. Ibba M, Söll and D (1999) Quality Control Mechanisms During Translation. *Science* 286:1893–1897.

6. Cusack S, Berthet-Colominas C, Härtlein M, Nassar N, Leberman R (1990) A second class of synthetase structure revealed by X-ray analysis of Escherichia coli seryl-tRNA synthetase at 2.5 A. *Nature* 347:249–255.

7. Cusack S (1993) Sequence, structure and evolutionary relationships between class 2 aminoacyl-tRNA synthetases: an update. *Biochimie* 75:1077–1081.

8. Eriani G, Delarue M, Poch O, Gangloff J, Moras D (1990) Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs. *Nature* 347:203–206.

9. Carter CW Jr (1993) Cognition, mechanism, and evolutionary relationships in aminoacyl-tRNA synthetases. *Annu Rev Biochem* 62:715–748.

10. Zhang C-M, Perona JJ, Ryu K, Francklyn C, Hou Y-M (2006) Distinct kinetic mechanisms of the two classes of Aminoacyl-tRNA synthetases. *J Mol Biol* 361:300–311.

11. Ling J et al. (2009) Resampling and editing of mischarged tRNA prior to translation elongation. *Mol Cell* 33:654–660.

12. Kim JH, Han JM, Kim S (2013) Protein-Protein Interactions and Multi-component Complexes of Aminoacyl-tRNA Synthetases. *Top Curr Chem*.

13. Yaremchuk A, Cusack S, Tukalo M (2000) Crystal structure of a eukaryote/archaeon-like prolyl-tRNA synthetase and its complex with tRNAPro(CGG). *EMBO J* 19:4745–4758.

14. Ahel I, Korencic D, Ibba M, Söll D (2003) Trans-editing of mischarged tRNAs. *Proc Natl Acad Sci* 100:15422–15427.

15. Crepin T, Yaremchuk A, Tukalo M, Cusack S (2006) Structures of Two Bacterial Prolyl-tRNA Synthetases with and without a cis-Editing Domain. *Structure* 14:1511–1525.

16. Wong F-C, Beuning PJ, Nagan M, Shiba K, Musier-Forsyth K (2002) Functional Role of the Prokaryotic Proline-tRNA Synthetase Insertion Domain in Amino Acid Editing†. *Biochemistry (Mosc)* 41:7108–7115.

17. SternJohn J, Hati S, Siliciano PG, Musier-Forsyth K (2007) Restoring species-specific posttransfer editing activity to a synthetase with a defunct editing domain. *Proc Natl Acad Sci* 104:2127–2132.

18. Cusack S, Yaremchuk A, Krikliviy I, Tukalo M (1998) tRNA(Pro) anticodon recognition by Thermus thermophilus prolyl-tRNA synthetase. *Struct Lond Engl 1993* 6:101–108.

19. Beuning PJ, Musier-Forsyth K (2001) Species-specific Differences in Amino Acid Editing by Class II Prolyl-tRNA Synthetase. *J Biol Chem* 276:30779–30785.

20. Kumar S, Das M, Hadad CM, Musier-Forsyth K (2012) Substrate Specificity of Bacterial Prolyl-tRNA Synthetase Editing Domain Is Controlled by a Tunable Hydrophobic Pocket. *J Biol Chem* 287:3175–3184.

21. So BR et al. (2011) Substrate-mediated Fidelity Mechanism Ensures Accurate Decoding of Proline Codons. *J Biol Chem* 286:31810 –31820.

22. Vargas-Rodriguez O, Musier-Forsyth K (2013) Exclusive Use of trans-Editing Domains Prevents Proline Mistranslation. *J Biol Chem* 288:14391–14399.

23. Reynolds NM, Lazazzera BA, Ibba M (2010) Cellular mechanisms that control mistranslation. *Nat Rev Microbiol* 8:849–856.

24. Jakubowski H, Goldman E (1992) Editing of errors in selection of amino acids for protein synthesis. *Microbiol Rev* 56:412–429.

25. Fersht AR (1977) Editing mechanisms in protein synthesis. Rejection of valine by the isoleucyl-tRNA synthetase. *Biochemistry (Mosc)* 16:1025–1030.

26. Rodgers KJ (2014) Non-protein amino acids and neurodegeneration: The enemy within. *Exp Neurol* 253C:192–196.

27. Rodgers KJ, Shiozawa N (2008) Misincorporation of amino acid analogues into proteins by biosynthesis. *Int J Biochem Cell Biol* 40:1452–1466.

28. Hendrickson TL, Crécy-Lagard V de, Schimmel P (2004) Incorporation of Nonnatural Amino Acids into Proteins. *Annu Rev Biochem* 73:147–176.

29. Döring V et al. (2001) Enlarging the Amino Acid Set of Escherichia coli by Infiltration of the Valine Coding Pathway. *Science* 292:501–504.

30. Nangle LA, Lagard V de C, Döring V, Schimmel P (2002) Genetic Code Ambiguity Cell Viability Related to Severity of Editing Defects in Mutant tRNA Synthetases. *J Biol Chem* 277:45729–45733.

31. Jakubowski H, Fersht AR (1981) Alternative pathways for editing non-cognate amino acids by aminoacyl-tRNA synthetases. *Nucleic Acids Res* 9:3105–3117.

32. Tsui W-C, Fersht AR (1981) Probing the principles of amino acid selection using the alanyl-tRNA synthetase from Escherichia coli. *Nucleic Acids Res* 9:4627–4637.

33. Dulic M, Cvetesic N, Perona JJ, Gruic-Sovulj I (2010) Partitioning of tRNA-dependent editing between pre- and post-transfer pathways in class I aminoacyl-tRNA synthetases. *J Biol Chem* 285:23799–23809.

34. Fersht AR (1998) Sieves in Sequence. *Science* 280:541–541.

35. Baldwin AN, Berg P (1966) Transfer Ribonucleic Acid-induced Hydrolysis of Valyladenylate Bound to Isoleucyl Ribonucleic Acid Synthetase. *J Biol Chem* 241:839–845.

36. Hendrickson TL et al. (2002) Mutational separation of two pathways for editing by a class I tRNA synthetase. *Mol Cell* 9:353–362.

37. Nomanbhoy TK, Hendrickson TL, Schimmel P (1999) Transfer RNA-dependent translocation of misactivated amino acids to prevent errors in protein synthesis. *Mol Cell* 4:519–528.

38. Bishop AC, Nomanbhoy TK, Schimmel P (2002) Blocking site-to-site translocation of a misactivated amino acid by mutation of a class I tRNA synthetase. *Proc Natl Acad Sci* 99:585–590.

39. Silvian LF, Wang J, Steitz TA (1999) Insights into editing from an ile-tRNA synthetase structure with tRNAile and mupirocin. *Science* 285:1074–1077.

40. Fukunaga R, Yokoyama S (2006) Structural basis for substrate recognition by the editing domain of isoleucyl-tRNA synthetase. *J Mol Biol* 359:901–912.

41. Igloi GL, von der Haar F, Cramer F (1977) Hydrolytic action of aminoacyl-tRNA synthetases from baker's yeast. "Chemical proofreading" of Thr-tRNA Val by valyl-tRNA synthetase studied with modified tRNA Val and amino acid analogues. *Biochemistry (Mosc)* 16:1696–1702.

42. Hopfield JJ (1974) Kinetic Proofreading: A New Mechanism for Reducing Errors in Biosynthetic Processes Requiring High Specificity. *Proc Natl Acad Sci U S A* 71:4135–4139.

43. Li L et al. (2013) Leucyl-tRNA synthetase editing domain functions as a molecular rheostat to control codon ambiguity in Mycoplasma pathogens. *Proc Natl Acad Sci U S A* 110:3817–3822.

44. Zhu B, Yao P, Tan M, Eriani G, Wang E-D (2009) tRNA-independent Pretransfer Editing by Class I Leucyl-tRNA Synthetase. *J Biol Chem* 284:3418–3424.

45. Splan KE, Ignatov ME, Musier-Forsyth K (2008) Transfer RNA Modulates the Editing Mechanism Used by Class II Prolyl-tRNA Synthetase. *J Biol Chem* 283:7128–7134.

46. Fersht AR (1977) *Enzyme structure and mechanism*.

47. Fersht AR, Kaethner MM (1976) Enzyme hyperspecificity. Rejection of threonine by the valyl-tRNA synthetase by misacylation and hydrolytic editing. *Biochemistry (Mosc)* 15:3342–3346.

48. Nureki O et al. (1998) Enzyme Structure with Two Catalytic Sites for Double-Sieve Selection of Substrate. *Science* 280:578–582.

49. Cusack S, Yaremchuk A, Tukalo M (2000) The 2 A crystal structure of leucyl-tRNA synthetase and its complex with a leucyl-adenylate analogue. *EMBO J* 19:2351–2361.

50. Fukai S et al. (2000) Structural basis for double-sieve discrimination of L-valine from L-isoleucine and L-threonine by the complex of tRNA(Val) and valyl-tRNA synthetase. *Cell* 103:793–803.

51. Mascarenhas AP, An S, Rosen AE, Martinis SA, Musier-Forsyth K (2009) in *Protein Engineering*, Nucleic Acids and Molecular Biology., eds Köhrer C, RajBhandary UL (Springer Berlin Heidelberg), pp 155–203.

52. Schmidt E, Schimmel P (1994) Mutational isolation of a sieve for editing in a transfer RNA synthetase. *Science* 264:265–267.

53. Sarkar J, Poruri K, Boniecki MT, McTavish KK, Martinis SA (2012) Yeast Mitochondrial Leucyl-tRNA Synthetase CP1 Domain Has Functionally Diverged to Accommodate RNA Splicing at Expense of Hydrolytic Editing. *J Biol Chem* 287:14772–14781.

54. Betha AK, Williams AM, Martinis SA (2007) Isolated CP1 Domain of Escherichia coli Leucyl-tRNA Synthetase Is Dependent on Flanking Hinge Motifs for Amino Acid Editing Activity†. *Biochemistry (Mosc)* 46:6258–6267.

55. Dock-Bregeon A et al. (2000) Transfer RNA-mediated editing in threonyl-tRNA synthetase. The class II solution to the double discrimination problem. *Cell* 103:877–884.

56. Dock-Bregeon A-C et al. (2004) Achieving Error-Free Translation. *Mol Cell* 16:375–386.

57. Beebe K, Ribas de Pouplana L, Schimmel P (2003) Elucidation of tRNA-dependent editing by a class II tRNA synthetase and significance for cell viability. *EMBO J* 22:668–675.

58. Naganuma M, Sekine S, Fukunaga R, Yokoyama S (2009) Unique protein architecture of alanyl-tRNA synthetase for aminoacylation, editing, and dimerization. *Proc Natl Acad Sci U S A* 106:8489–8494.

59. Sokabe M et al. (2009) The structure of alanyl-tRNA synthetase with editing domain. *Proc Natl Acad Sci* 106:11028–11033.

60. Roy H, Ling J, Irnov M, Ibba M (2004) Post-transfer editing in vitro and in vivo by the subunit of phenylalanyl-tRNA synthetase. *EMBO J* 23:4639–4648.

61. Sasaki HM et al. (2006) Structural and mutational studies of the amino acid-editing domain from archaeal/eukaryal phenylalanyl-tRNA synthetase. *Proc Natl Acad Sci U S A* 103:14744–14749.

62. Sanni A, Walter P, Boulanger Y, Ebel JP, Fasiolo F (1991) Evolution of aminoacyl-tRNA synthetase quaternary structure and activity: Saccharomyces cerevisiae mitochondrial phenylalanyl-tRNA synthetase. *Proc Natl Acad Sci U S A* 88:8387–8391.

63. Ling J, Yadavalli SS, Ibba M (2007) Phenylalanyl-tRNA synthetase editing defects result in efficient mistranslation of phenylalanine codons as tyrosine. *RNA* 13:1881–1886.

64. Beuning PJ, Musier-Forsyth K (2000) Hydrolytic editing by a class II aminoacyl-tRNA synthetase. *Proc Natl Acad Sci U S A* 97:8916–8920.

65. Kumar S, Das M, Hadad CM, Musier-Forsyth K (2012) Substrate and Enzyme Functional Groups Contribute to Translational Quality Control by Bacterial Prolyl-tRNA Synthetase. *J Phys Chem B* 116:6991–6999.

66. Wong F-C, Beuning PJ, Silvers C, Musier-Forsyth K (2003) An isolated class II aminoacyl-tRNA synthetase insertion domain is functional in amino acid editing. *J Biol Chem* 278:52857–52864.

67. Oki K, Sakamoto K, Kobayashi T, Sasaki HM, Yokoyama S (2008) Transplantation of a tyrosine editing domain into a tyrosyl-tRNA synthetase variant enhances its specificity for a tyrosine analog. *Proc Natl Acad Sci U S A* 105:13298–13303.

68. Calendar R, Berg P (1967) D-Tyrosyl RNA: formation, hydrolysis and utilization for protein synthesis. *J Mol Biol* 26:39–54.

69. Hussain T et al. (2006) Post-transfer editing mechanism of a D-aminoacyl-tRNA deacylase-like domain in threonyl-tRNA synthetase from archaea. *EMBO J* 25:4152–4162.

70. Rigden DJ (2004) Archaea recruited D-Tyr-tRNATyr deacylase for editing in Thr-tRNA synthetase. *RNA N Y N* 10:1845–1851.

71. Dwivedi S, Kruparani SP, Sankaranarayanan R (2005) A D-amino acid editing module coupled to the translational apparatus in archaea. *Nat Struct Mol Biol* 12:556–557.

72. Korencic D et al. (2004) A freestanding proofreading domain is required for protein synthesis quality control in Archaea. *Proc Natl Acad Sci U S A* 101:10260–10265.

73. Beebe K, Mock M, Merriman E, Schimmel P (2008) Distinct domains of tRNA synthetase recognize the same base pair. *Nature* 451:90–93.

74. Sokabe M, Okada A, Yao M, Nakashima T, Tanaka I (2005) Molecular basis of alanine discrimination in editing site. *Proc Natl Acad Sci U S A* 102:11669–11674.

75. Guo M, Schimmel P (2012) Structural analyses clarify the complex control of mistranslation by tRNA synthetases. *Curr Opin Struct Biol* 22:119–126.

76. Ishijima J et al. (2006) Crystal structure of alanyl-tRNA synthetase editing-domain homolog (PH0574) from a hyperthermophile, Pyrococcus horikoshii OT3 at 1.45 A resolution. *Proteins* 62:1133–1137.

77. Fukunaga R, Yokoyama S (2007) Structure of the AlaX-M trans-editing enzyme from Pyrococcus horikoshii. *Acta Crystallogr D Biol Crystallogr* 63:390–400.

78. Chong YE, Yang X-L, Schimmel P (2008) Natural Homolog of tRNA Synthetase Editing Domain Rescues Conditional Lethality Caused by Mistranslation. *J Biol Chem* 283:30073–30078.

79. An S, Musier-Forsyth K (2004) Trans-editing of Cys-tRNAPro by Haemophilus influenzae YbaK Protein. *J Biol Chem* 279:42359–42362.

80. Murayama K et al. (2005) Structure of a putative trans-editing enzyme for prolyl-tRNA synthetase from Aeropyrum pernix K1 at 1.7 A resolution. *Acta Crystallograph Sect F Struct Biol Cryst Commun* 61:26–29.

81. An S, Musier-Forsyth K (2005) Cys-tRNAPro Editing by Haemophilus influenzae YbaK via a Novel Synthetase·YbaK·tRNA Ternary Complex. *J Biol Chem* 280:34465–34472.

82. Bell EA (2003) Nonprotein amino acids of plants: significance in medicine, nutrition, and agriculture. *J Agric Food Chem* 51:2854–2865.

83. Nunn PB, Bell EA, Watson AA, Nash RJ (2010) Toxicity of non-protein amino acids to humans and domestic animals. *Nat Prod Commun* 5:485–504.

84. Fowden L, Lewis D, Tristram H (1967) Toxic amino acids: their action as antimetabolites. *Adv Enzymol Relat Areas Mol Biol* 29:89–163.

85. Bertin C et al. (2007) Grass roots chemistry: meta-tyrosine, an herbicidal nonprotein amino acid. *Proc Natl Acad Sci U S A* 104:16964–16969.

86. Rodgers KJ, Wang H, Fu S, Dean RT (2002) Biosynthetic incorporation of oxidized amino acids into proteins and their cellular proteolysis. *Free Radic Biol Med* 32:766–775.

87. Peterson PJ, Fowden L (1965) Purification, properties and comparative specificities of the enzyme prolyl-transfer ribonucleic acid synthetase from Phaseolus aureus and Polygonatum multiflorum. *Biochem J* 97:112–124.

88. Rubenstein E et al. (2009) Azetidine-2-carboxylic acid in the food chain. *Phytochemistry* 70:100–104.

89. Thomas DA, Rosenthal GA (1987) Metabolism of L-[guanidinooxy-14C]canavanine in the rat. *Toxicol Appl Pharmacol* 91:406–414.

90. Thomas DA, Rosenthal GA (1987) Toxicity and pharmacokinetics of the nonprotein amino acid L-canavanine in the rat. *Toxicol Appl Pharmacol* 91:395–405.

91. Warren RP, Hunt GE (1971) The biosynthesis of canavanine from (14)CO 2 and its asymmetric labeling in isolated pericarp, tissue of Canavalia ensiformis. *Planta* 100:258–261.

92. Rubenstein E (2000) Biologic effects of and clinical disorders caused by nonprotein amino acids. *Medicine (Baltimore)* 79:80–89.

93. Bradley WG, Mash DC (2009) Beyond Guam: the cyanobacteria/BMAA hypothesis of the cause of ALS and other neurodegenerative diseases. *Amyotroph Lateral Scler Off Publ World Fed Neurol Res Group Mot Neuron Dis* 10 Suppl 2:7–20.

94. Duncan MW (1992) β-Methylamino-L-Alanine (BMAA) and Amyotrophic Lateral Sclerosis–Parkinsonism Dementia of the Western Pacific. *Ann N Y Acad Sci* 648:161–168.

95. Vega A, Bell EA, Nunn PB (1968) The preparation of l- and d-α-amino-β-methylaminopropionic acids and the identification of the compound isolated from Cycas circinalis as the l-isomer. *Phytochemistry* 7:1885–1887.

96. Spencer PS et al. (2007) in *Ciba Foundation Symposium 126 - Selective Neuronal Death*, eds Organizer GB, O'Connor eve (John Wiley & Sons, Ltd.), pp 221–238.

97. Spencer PS et al. (1987) Guam amyotrophic lateral sclerosis-parkinsonism-dementia linked to a plant excitant neurotoxin. *Science* 237:517–522.

98. Cox PA, Banack SA, Murch SJ (2003) Biomagnification of cyanobacterial neurotoxins and neurodegenerative disease among the Chamorro people of Guam. *Proc Natl Acad Sci U S A* 100:13380–13383.

99. Dunlop RA, Dean RT, Rodgers KJ (2008) The impact of specific oxidized amino acids on protein turnover in J774 cells. *Biochem J* 410:131–140.

100. Rodgers KJ, Hume PM, Dunlop RA, Dean RT (2004) Biosynthesis and turnover of DOPA-containing proteins by human cells. *Free Radic Biol Med* 37:1756–1764.

101. Chan SW, Dunlop RA, Rowe A, Double KL, Rodgers KJ (2012) L-DOPA is incorporated into brain proteins of patients treated for Parkinson's disease, inducing toxicity in human neuroblastoma cells in vitro. *Exp Neurol* 238:29–37.

102. Rubenstein E (2008) Misincorporation of the proline analog azetidine-2-carboxylic acid in the pathogenesis of multiple sclerosis: a hypothesis. *J Neuropathol Exp Neurol* 67:1035–1040.

103. Rubenstein E, Zhou H, Krasinska KM, Chien A, Becker CH (2006) Azetidine-2-carboxylic acid in garden beets (Beta vulgaris). *Phytochemistry* 67:898–903.

104. Rosenthal GA, Dahlman DL, Janzen DH (1976) A novel means for dealing with L-canavanine, a toxic metabolite. *Science* 192:256–258.

105. Rosenthal GA, Hughes CG, Janzen DH (1982) L-Canavanine, a Dietary Nitrogen Source for the Seed Predator Caryedes brasiliensis (Bruchidae). *Science* 217:353–355.

106. Igloi GL, Schiefermayr E (2009) Amino acid discrimination by arginyl-tRNA synthetases as revealed by an examination of natural specificity variants. *FEBS J* 276:1307–1318.

107. Lee JW et al. (2006) Editing-defective tRNA synthetase causes protein misfolding and neurodegeneration. *Nature* 443:50–55.

108. Bacher JM, Schimmel P (2007) An editing-defective aminoacyl-tRNA synthetase is mutagenic in aging bacteria via the SOS response. *Proc Natl Acad Sci U S A* 104:1907–1912.

109. Ruan B et al. (2008) Quality control despite mistranslation caused by an ambiguous genetic code. *Proc Natl Acad Sci U S A* 105:16502–16507.

110. Pezo V et al. (2004) Artificially ambiguous genetic code confers growth yield advantage. *Proc Natl Acad Sci U S A* 101:8593–8597.

111. Bacher JM, Waas WF, Metzgar D, de Crécy-Lagard V, Schimmel P (2007) Genetic code ambiguity confers a selective advantage on Acinetobacter baylyi. *J Bacteriol* 189:6494–6496.

112. Li L et al. (2011) Naturally occurring aminoacyl-tRNA synthetases editing-domain mutations that cause mistranslation in Mycoplasma parasites. *Proc Natl Acad Sci U S A* 108:9378–9383.

113. Burke B, Lipman RSA, Shiba K, Musier-Forsyth K, Hou Y-M (2001) Divergent Adaptation of tRNA Recognition by Methanococcus jannaschii Prolyl-tRNA Synthetase. *J Biol Chem* 276:20286–20291.

114. Lee J, Shin M-K, Ryu D-K, Kim S, Ryu W-S (2010) in *In Vitro Mutagenesis Protocols*, Methods in Molecular Biology., ed Braman J (Humana Press), pp 137–146.

115. Nishihara K, Kanemori M, Kitagawa M, Yanagi H, Yura T (1998) Chaperone Coexpression Plasmids: Differential and Synergistic Roles of DnaK-DnaJ-GrpE and GroEL-GroES in Assisting Folding of an Allergen of Japanese Cedar Pollen, Cryj2, in Escherichia coli. *Appl Environ Microbiol* 64:1694–1699.

116. Nishihara K, Kanemori M, Yanagi H, Yura T (2000) Overexpression of Trigger Factor Prevents Aggregation of Recombinant Proteins in Escherichia coli. *Appl Environ Microbiol* 66:884–889.

117. Pauling L (1958) in *Festschrift Arthur Stoll Siebzigsten Geburtstag*, pp 597–602.

118. Fersht AR, Dingwall C (1979) Establishing the misacylation/deacylation of the tRNA pathway for the editing mechanism of prokaryotic and eukaryotic valyl-tRNA synthetases. *Biochemistry (Mosc)* 18:1238–1245.

119. Ruan B, Söll D (2005) The Bacterial YbaK Protein Is a Cys-tRNAPro and Cys-tRNACys Deacylase. *J Biol Chem* 280:25887–25891.

120. Stehlin C et al. (1998) Species-specific differences in the operational RNA code for aminoacylation of tRNAPro. *Biochemistry (Mosc)* 37:8605–8613.

121. Nordin BE, Schimmel P (2002) Plasticity of Recognition of the 3′-End of Mischarged tRNA by Class I Aminoacyl-tRNA Synthetases. *J Biol Chem* 277:20510–20517.

122. Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72:248–254.

123. Fersht AR et al. (1975) Active site titration and aminoacyl adenylate binding stoichiometry of aminoacyl-tRNA synthetases. *Biochemistry (Mosc)* 14:1–4.

124. Ledoux S, Uhlenbeck OC (2008) [3'-32P]-labeling tRNA with nucleotidyltransferase for assaying aminoacylation and peptide bond formation. *Methods San Diego Calif* 44:74–80.

125. Splan KE, Musier-Forsyth K, Boniecki MT, Martinis SA (2008) In vitro assays for the determination of aminoacyl-tRNA synthetase editing activity. *Methods* 44:119–128.

126. D.A. Case, T.A. Darden, T.E. Cheatham, III, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, R.C. Walker, W. Zhang, K.M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A.W. Goetz, I. Kolossváry, K.F. Wong, F. Paesani, J. Vanicek, R.M. Wolf, J. Liu, X. Wu, S.R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D.R. Roe, D.H. Mathews, M.G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, and P.A. Kollman (2012) AMBER 12.

127. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79:926–935.

128. Cornell WD et al. (1995) A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J Am Chem Soc* 117:5179–5197.

129. Cheatham TE 3rd, Kollman PA (2000) Molecular dynamics simulation of nucleic acids. *Annu Rev Phys Chem* 51:435–471.

130. York DM, Darden TA, Pedersen LG (1993) The effect of long-range electrostatic interactions in simulations of macromolecular crystals: A comparison of the Ewald and truncated list methods. *J Chem Phys* 99:8345–8348.

131. Ryckaert J-P, Ciccotti G, Berendsen HJC (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comput Phys* 23:327–341.

132. Sindhikara DJ, Kim S, Voter AF, Roitberg AE (2009) Bad Seeds Sprout Perilous Dynamics: Stochastic Thermostat Induced Trajectory Synchronization in Biomolecules. *J Chem Theory Comput* 5:1624–1631.

133. Uberuaga BP, Anghel M, Voter AF (2004) Synchronization of trajectories in canonical molecular-dynamics simulations: Observation, explanation, and exploitation. *J Chem Phys* 120:6363–6374.

134. Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. *J Mol Graph* 14:33–38, 27–28.

135. DeLano WL (2009) PyMOL molecular viewer: Updates and refinements. *Abstr Pap Am Chem Soc* 238.

136. DeLano WL (2005) PyMOL: A communications tool for computational models. *Abstr Pap Am Chem Soc* 230:U1371–U1372.

137. Hanwell MD et al. (2012) Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *J Cheminformatics* 4:17.

138. Morris GM et al. (2009) AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem* 30:2785–2791.

139. Trott O, Olson AJ (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* 31:455–461.

140. Ahel I et al. (2002) Cysteine Activation Is an Inherent in Vitro Property of Prolyl-tRNA Synthetases. *J Biol Chem* 277:34743–34748.

141. Loftfield RB, Vanderjagt D (1972) The frequency of errors in protein biosynthesis. *Biochem J* 128:1353–1356.

142. Zhang K, Li H, Cho KM, Liao JC (2010) Expanding metabolism for total biosynthesis of the nonnatural amino acid L-homoalanine. *Proc Natl Acad Sci* 107:6234–6239.

143. Zhang W et al. (2012) Efficient bioconversion of l-threonine to 2-oxobutyrate using whole cells of *Pseudomonas stutzeri* SDM. *Bioresour Technol* 110:719–722.

144. Zhu L et al. (2011) Removal of l-alanine from the production of l-2-aminobutyric acid by introduction of alanine racemase and d-amino acid oxidase. *Appl Microbiol Biotechnol* 90:903–910.

145. Carroll WR, Stacy GW, Vigneaud V du (1949) α-Ketobutyric Acid as a Product in the Enzymatic Cleavage of Cystathionine. *J Biol Chem* 180:375–382.

146. Larimer FW et al. (2004) Complete genome sequence of the metabolically versatile photosynthetic bacterium Rhodopseudomonas palustris. *Nat Biotechnol* 22:55–61.

147. VerBerkmoes NC et al. (2006) Determination and Comparison of the Baseline Proteomes of the Versatile Microbe Rhodopseudomonas palustris under Its Major Metabolic States. *J Proteome Res* 5:287–298.

148. Praetorius-Ibba M, Hausmann CD, Paras M, Rogers TE, Ibba M (2007) Functional association between three archaeal aminoacyl-tRNA synthetases. *J Biol Chem* 282:3680–3687.

149. Schuck P (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys J* 78:1606–1619.

150. LaRiviere FJ, Wolfson AD, Uhlenbeck OC (2001) Uniform Binding of Aminoacyl-tRNAs to Elongation Factor Tu by Thermodynamic Compensation. *Science* 294:165–168.

151. Murayama K et al. (2007) Crystal structure of ProX-AlaSA complex from T. thermophilus. *Be Publ*:null–null.

152. Kim M-K, Harwood CS (1991) Regulation of benzoate-CoA ligase in Rhodopseudomonas palustris. *FEMS Microbiol Lett* 83:199–203.

153. Walker DJ (1958) The purification and properties of the L-threonine deaminase of the rumen micro-organism LC1.

154. Fotheringham IG, Grinter N, Pantaleone DP, Senkpeil RF, Taylor PP (1999) Engineering of a novel biochemical pathway for the biosynthesis of l-2-aminobutyric acid in Escherichia coli K12. *Bioorg Med Chem* 7:2209–2213.

155. Umbarger HE (1978) Amino Acid Biosynthesis and its Regulation. *Annu Rev Biochem* 47:533–606.

156. Asao M, Alber BE (2013) Acrylyl-Coenzyme A Reductase, an Enzyme Involved in the Assimilation of 3-Hydroxypropionate by Rhodobacter sphaeroides. *J Bacteriol* 195:4716–4725.

157. Crutchfield CA, Lu W, Melamud E, Rabinowitz JD (2010) in *Methods in Enzymology*, ed Jonathan Weissman; Christine Guthrie and Gerald R. Fink (Academic Press), pp 393–426.

158. Graber D et al. (2012) Deoxyribozyme-based, semisynthetic access to stable peptidyl-tRNAs exemplified by tRNAVal carrying a macrolide antibiotic resistance peptide. *Methods Mol Biol Clifton NJ* 848:201–213.

159. Karkhanis VA, Mascarenhas AP, Martinis SA (2007) Amino Acid Toxicities of Escherichia coli That Are Prevented by Leucyl-tRNA Synthetase Amino Acid Editing. *J Bacteriol* 189:8765–8768.