Exploring the interplay of structure, stability, activity and localization in tRNA nucleotidyltransferase function

Matthew Leibovitch

A Thesis

In the Department

of

Chemistry and Biochemistry

Presented in Partial Fulfillment of the Requirements

For the degree of

Doctorate of Philosophy (Chemistry and Biochemistry) at

Concordia University

Montréal, Québec, Canada

April 2016

This is to certify that the thesis prepared

By: Matthew Leibovitch

Entitled: Exploring the interplay of structure, stability, activity and localization in tRNA nucleotidyltransferase function

and submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy Chemistry (Biochemistry)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee:

_____ Chair
Dr. Jack Kornblatt

_____ External Examiner
Dr. Alan Weiner

_____ External to Program
Dr. David Kwan

_____ Examiner
Dr. Joanne Turnbull

_____ Examiner
Dr. Peter Pawelek

_____ Thesis Supervisor
Dr. Paul Joyce

Approved by: _____
Dr. Heidi Muchall, Graduate Program Director

April 2016 _____
Dr. André Roy, Dean of Faculty

Abstract

Exploring the interplay of structure, stability, activity and localization in tRNA nucleotidyltransferase function

Matthew Leibovitch, PhD

Concordia University, 2016

The enzyme tRNA nucleotidyltransferase is of central importance in eukaryotic cells since it is required in multiple intracellular locations (nucleus, cytosol, mitochondria and plastids) for the production of mature tRNAs needed for protein synthesis. Therefore, tRNA nucleotidyltransferase serves as an excellent model protein to explore the complex roles that localization, activity, structure and stability play in defining protein function. I investigated how changes in the amino acid sequence of tRNA nucleotidyltransferase affect its structure, stability, activity and localization and may be linked, ultimately, to human disease.

Using *Arabidopsis thaliana* as a model organism, I was the first to show that amino acid substitutions located within a region of tRNA nucleotidyltransferase required for enzyme activity, but outside of the enzyme's classical amino-terminal organellar targeting sequence, affect localization to plastids and mitochondria. These studies suggest a direct link between intracellular distribution and how tightly the protein has folded. Moreover, these data indicate that the same seven amino acid changes that result in altered stability also reduce tRNA binding and enzyme activity. To more clearly define the role that these amino acids have in enzyme structure and function and to determine whether changes in protein stability, substrate binding, or structure bring about altered localization, I have fine-structure mapped the region encompassing residues 399-420 and shown that a single amino acid substitution (K418E) can dramatically alter tRNA binding and enzyme activity without detectable effects on enzyme

structure. Localization studies are now underway to determine whether the K418E substitution, or other single substitutions within this region, alters the intracellular localization of tRNA nucleotidyltransferase.

Using a second model organism, *Saccharomyces cerevisiae*, I have further explored the link between stability and activity in protein function. I showed that a single amino acid substitution (E189F) resulted both in a 5^{o}C decrease in thermal stability and a 25-fold decrease in enzyme activity and caused a temperature-sensitive phenotype. Moreover, converting arginine 64 to tryptophan in the variant enzyme restored enzyme activity and suppressed the temperature-sensitive phenotype, but did not restore thermal stability. These data suggest that the temperature-sensitive phenotype is defined by the reduction in activity and not thermal stability. Moreover, these data were the first to suggest a role for conserved motif C in active site organization and tRNA nucleotidyltransferase function.

Finally, I applied the techniques and insight acquired from my studies of the Arabidopsis and yeast enzymes to explore how seven distinct amino-acid substitutions in human tRNA nucleotidyltransferase, all previously identified in patients suffering from sideroblastic anemia with B cell immunodeficiency, periodic fevers and development delay (SIFD), affect enzyme structure, stability and function *in vitro*. I showed that each of these mutations affects some combination of tRNA nucleotidyltransferase structure, activity, or stability. As expected, substitutions (T154I, M158V, L166S, R190I and I223T) within the conserved catalytic amino terminal region of the protein altered stability and/or activity but more interestingly two substitutions (I326T and K416E) in the less well conserved carboxy-terminal region of the enzyme altered catalytic efficiency, tRNA binding and quaternary structure.

These studies have shown that a protein's function can be defined by a combination of some or all of its structure, stability, activity or localization.

Acknowledgments

I would first like to thank Dr. Paul Joyce for his support, patience and guidance for the duration of my undergraduate and PhD graduate degrees. The many years I have spent under his guidance have provided me with invaluable research and communication skills and a solid foundation for me to begin my career. The training that I received in communication, critical thinking and scientific writing has enhanced my skills in the areas in which I am weakest and will benefit me in the many years to come. It has been a pleasure working and learning under his supervision and I will apply these skills and lessons for the rest of my life. I want to acknowledge Dr. Pamela Hanic-Joyce for her valuable knowledge in molecular biology techniques and in life. Her expertise in molecular biology was used as a stepping stone in my research experience and has made me a better biochemist today. It was a pleasure working by her side for the past six years and she will be missed in future years. I would like to thank my committee members Dr. Joanne Turnbull and Dr. Peter Pawelek for guidance, helpful discussions and critical thinking at committee meetings. Dr. Joanne Turnbull's expertise in protein chemistry and kinetics and Dr. Peter Pawelek's expertise in protein interactions and crystallography were critical in acquiring my PhD. It was also a pleasure working with many talented faculty, colleagues and staff during my time at Concordia. I want to give a huge thank you to my family and friends for their love and support throughout the years. I want to specifically thank my mother, Elena and my father, Jay, my brother Aaron and my sister Rebecca for supporting my decisions and encouraging me to keep moving forward. Finally, I would like to give a special thanks to my true love, Jahsina Virgo for her love, patience and belief in my abilities throughout my degree. She has supported me and listened to numerous presentations even if she didn't understand anything. The years we spent together were precious and I hope to have many more. I am grateful for her continued caring, love and support.

Table of Contents

List of tables

List of Figures

List of abbreviations:

AMP	adenosine monophosphate
ATP	adenosine triphosphate
CCA	cytidine-cytidine-adenosine
cDNA	complementary deoxyribonucleic acid
CMP	cytidine monophosphate
CTP	cytidine triphosphate
dAMP	deoxyadenosine monophosphate
dATP	deoxyadenosine triphosphate
dCMP	deoxycytidine monophosphate
dCTP	deoxycytidine triphosphate
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphates
GFP	green fluorescent protein
GST	glutathione sulfotransferase
GTP	guanosine triphosphate
IPTG	Isopropyl β-D-1-thiogalactopyranoside
Mat-Mu	mutant of the mature domain
NTP	ribonucleotide triphosphate
OD	optical density
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PDB	protein data bank
RNA	ribonucleic acid
SDS	sodium dodecyl sulfate
TCEP	tris(2-carboxyethyl)phosphine
Tm	denaturation midpoint
tRNA	transfer ribonucleic acid
tRNA-NT	tRNA nucleotidyltransferase
UTP	uridine triphosphate
YT	yeast extract, tryptone

1.0 Introduction

1.1 The interplay in protein function of structure, stability, activity and localization

The central dogma of molecular biology "DNA makes RNA makes protein" (Crick, 1970) is a satisfying explanation for how the DNA contained in a cell can define that cell's fate. However, the story obviously is more complicated than this. A cell's fate is defined not only by the genes that it contains, but also by how transcription of those genes is regulated, how the resulting RNAs are processed, and how the final gene products are generated. Every protein that is produced plays an important role in the cell, either as an enzyme, perhaps playing a role in metabolism, *e.g.*, hexokinase; or as a structural component, *e.g.*, actin or tubulin in the cytoskeleton; or as a transport protein, *e.g.*, hemoglobin (which carries oxygen through the bloodstream) or the glucose transporter (which transports glucose across the plasma membrane); or as a signaling molecule, *e.g.*, insulin; or in protection, *e.g.*, antibodies, toxins. The roles of each of these proteins are defined by the interplay of factors including structure, stability, activity and localization that establish their functions. For example, the structure of an enzyme must be such that active site residues are arranged for the appropriate substrate to bind and catalysis to take place. A transport protein like hemoglobin must be stable enough to carry oxygen through the blood stream from the lungs to the tissue where it is needed. Even if a protein has the appropriate stability, structure and activity, if it is not directed to the appropriate destination it cannot function properly. For example, the glucose transporter must be directed to the plasma membrane and inserted in the correct orientation to play its part in glucose transport. Finally, a single protein may have more than one activity, for example, enolase which catalyzes the penultimate step in glycolysis is also involved in the import of tRNA into yeast mitochondria (Entelis *et al.*, 2006). Clearly, it is the combination of a protein's localization, activity, structure and stability that define its function. Consequently, changing any of these parameters may affect a protein's function and cause problems. For example, altered protein structure can lead to disease as is the case in Creutzfeldt-Jakob disease where a misfolded protein generates more

misfolded proteins and protein aggregation (Diack *et al.*, 2014). Other diseases such as cystic fibrosis and hereditary emphysema (caused by α-antitrypsin deficiency) (Ramirez-Alvarado, 2008) may result from protein misfolding caused by changes in structure leading to mislocation of a protein. Therefore, a better appreciation of how structure, stability, activity and localization define a protein's function is of more than fundamental interest and may aid in understanding human disease.

1.2 Regulation of protein function

Protein function is regulated through transcriptional, translational and post-translational processes. As described above and in the simplest sense, the amount of a protein produced in a cell is dependent on the regulation of transcription and translation to generate the appropriate gene product. These processes themselves often are tightly regulated. While some proteins are constitutively produced others are only synthesized when needed. The canonical example of this is the lac operon where the enzymes required in bacteria for lactose metabolism are only produced when lactose is present (Lewis, 2013). In addition, protein production can be regulated temporally and spatially, *e.g*., cell differentiation during embryogenesis (Sulston *et al*., 1983). The level of a protein in a cell also depends on the stability of that protein and proteins differ in their stabilities (Bachmair *et al.*, 1986) and may be targeted for rapid degradation (Bloom and Cross, 2007). Protein function also may be controlled post-translationally (Walsh *et al.*, 2005). Covalent modification is a classical way to control a protein's function. For example, phosphorylation or dephosphorylation of glycogen phosphorylase induces a conformational change near the active site that can regulate its activity (Agius, 2015). Post-translational modifications of proteins also can lead to altered cellular distribution, for example, nuclear *e.g.,* Swi5, Ace2 (Moll *et al.*, 1991, Sbia *et al.*, 2008) and plastid (Lamberti *et al.*, 2011) targeting may be regulated by phosphorylation at or near targeting signals. As these proteins need to be directed to their proper destinations to exert their activities, altering localization is another example of how function can be regulated.

1.3 ATP(CTP):tRNA nucleotidyltransferase as a model protein

ATP(CTP):tRNA nucleotidyltransferase represents an excellent protein for study to develop a deeper understanding of how protein localization, activity, structure and stability may play a role in defining protein function. This enzyme is responsible for the addition of the 3'-terminal cytidine-cytidine-adenosine (CCA) sequence to tRNAs that lack it (Deutscher, 1990). The chemistry behind this reaction has been the subject of much interest as it represents the addition of a specific nucleotide sequence without the use of a nucleic acid template (Cho *et al.*, 2002, Cho *et al.*, 2006, Cho *et al.*, 2007, Tomita and Weiner, 2001, Tomita *et al.*, 2004, Shi *et al.*, 1998, Pan *et al.*, 2010, Hoffmeier *et al.*, 2010, Hanic-Joyce and Joyce, 2002, Shan *et al.*, 2008). In addition, in the eukaryotic species studied to date, *i.e.*, *Saccharomyces cerevisiae* (Chen *et al.*, 1992), *Kluyveromyces lactis* (Deng *et al.*, 2000), human (Reichert *et al.*, 2001), Xenopus (Keady *et al.*, 2002), *Candida glabrata* (Hanic-Joyce and Joyce, 2002) and Arabidopsis (Schmidt von Braun *et al.*, 2007), a single gene codes for ATP(CTP):tRNA nucleotidyltransferase which then is targeted to multiple destinations, *e.g.*, nucleus, cytosol, mitochondria and plastids (when present). This combination of interesting enzymology and intracellular targeting makes this enzyme an excellent choice for the study of the multiple factors that regulate a protein's function.

1.3.1 Structure, stability and activity

Eukaryotic and bacterial tRNA nucleotidyltransferases belong to the Class II family of nucleotidyltransferases (Yue *et al.*, 1996). This class is defined by sequence conservation distributed throughout the 25 kDa amino terminal portion of the proteins (Yue *et al.*, 1996). Class II tRNA nucleotidyltransferase enzymes for which crystal structures have been solved, *i.e.*, human (Augustin *et al.*, 2003, Kuhn *et al.*, 2015), *Thermotoga maritima* (Toh *et al.*, 2009), *Aquifex aeolicus* (Tomita *et al.*, 2004), and *Bacillus stearothermophilus* (Li *et al.*, 2002) CCA-adding enzymes and the *Aquifex aeolicus* CC-

adding (Yamashita *et al.*, 2014) and A-adding (Tomita *et al.*, 2004) enzymes show a common higher order structure reminiscent of a sea horse (Li *et al.*, 2002) with four structurally conserved regions: head, neck, body and tail from N-terminus to C-terminus (Fig. 1.1).

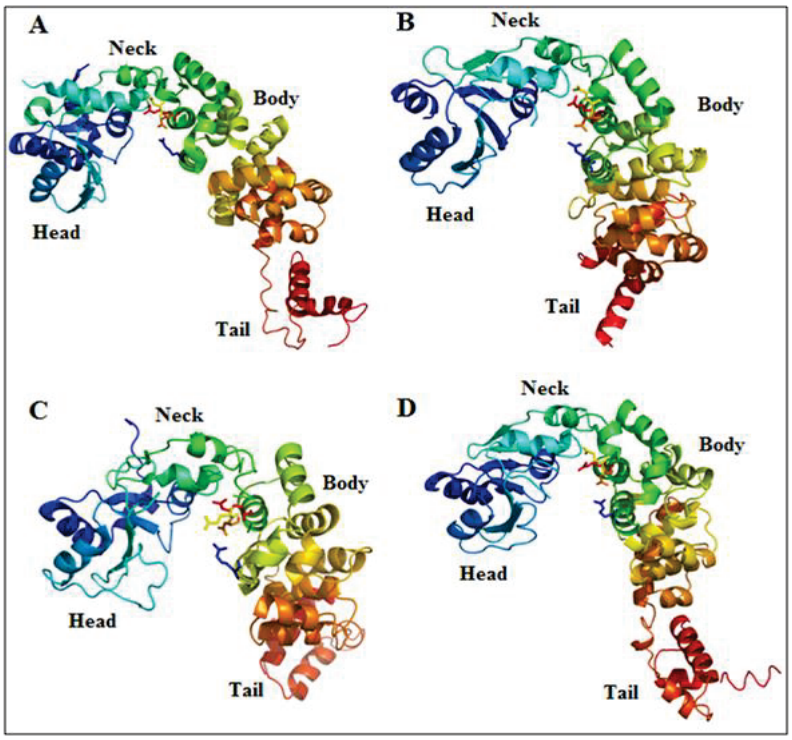


Figure 1.1: Crystal structures of Class II tRNA nucleotidyltransferases. Structures of tRNA nucleotidyltransferases from *Thermotoga maritima* (Toh *et al.*, 2009), *Aquifex aeolicus* (Tomita *et al.*, 2004), *Homo sapiens* (Kuhn *et al.*, 2015), and *Bacillus stearothermophilus* (Li *et al.*, 2002) are labeled A-D, respectively. Their PDB numbers are 3H37, 1VFG, 4X4W and 1MIV, respectively. The crystal structures were viewed with PyMol (Schrödinger).

1.3.1.1 Conserved motifs

Five conserved motifs (A-E) in the head and neck regions of Class II tRNA nucleotidyltransferases which help to define the active site were initially identified through sequence comparisons of the *Bacillus stearothermophilus* CCA-adding enzyme, *Escherichia coli* CCA-adding enzyme, *Aquifex aeolicus* A-adding enzyme, *Bacillus subtilis* CCA-adding enzyme, and *Saccharomyces cerevisiae* CCA-adding enzyme (Li *et al.*, 2002).

Motif A contains two conserved aspartic acid residues involved in coordinating the magnesium ions involved in catalysis and in binding the triphosphate moiety of each nucleotide triphosphate (Fig. 1.2) (Joyce and Steitz, 1994, Steitz, 1998).

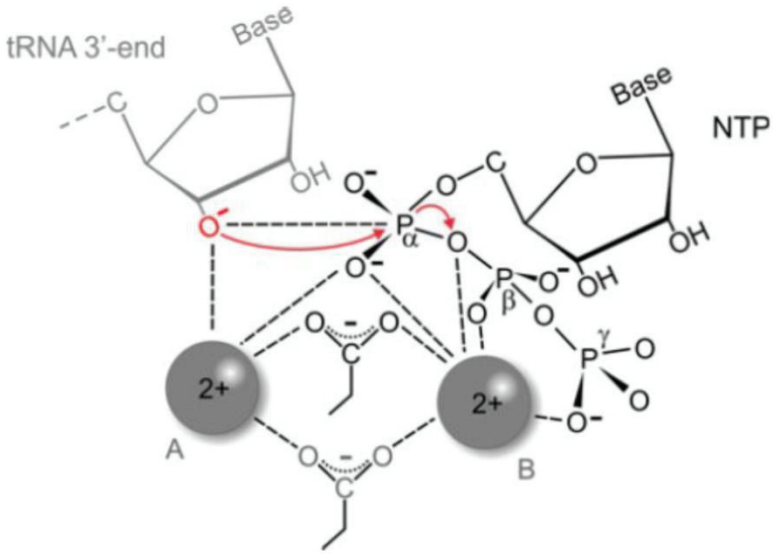


Figure 1.2: The role of Motif A (Betat *et al.*, 2010 modified from Steitz, 1998). The grey balls represent the two magnesium ions and are coordinated by the two carboxylates (aspartates). Magnesium ion "A" deprotonates the terminal 3' hydroxyl group of the tRNA and activates the 3' attack on the α-phosphate of the incoming nucleotide triphosphate (NTP). Magnesium ion "B" stabilizes the triphosphate moiety of the NTP in order to promote pyrophosphate release.

Motif B has a role in discriminating between deoxyribonucleotides and ribonucleotides by recognizing the 2' OH group found on ribose (Li *et al.*, 2002). A conserved arginine-arginine-aspartate sequence allows the central arginine to hydrogen bond with the 2' hydroxyl group found only on ribose (Fig. 1.3) allowing the enzyme to discriminate against deoxyribose (Cho *et al.*, 2007). A substitution of the central arginine by an isoleucine resulted in a functional protein that had lost the ability to discriminate between deoxyribonucleotides and ribonucleotides (Cho *et al.*, 2007).

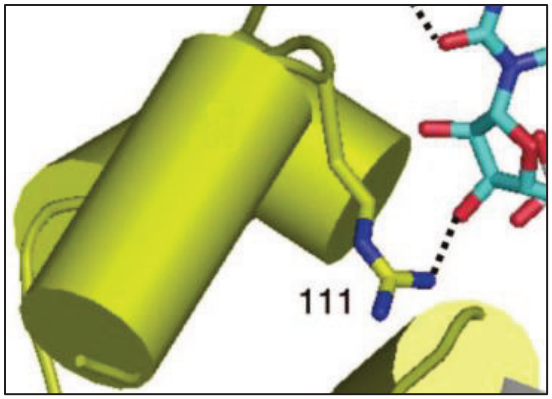


Figure 1.3: Nucleotide binding site of *Bacillus stearothermophilus* CCA-Adding enzyme (modified from Cho *et al.,* 2007). Stick model of catalytically important arginine 111 side chain interacting with 2' hydroxyl of ribose of substrate. Nitrogen is shown in blue; oxygen in red; carbon is yellow in the polypeptide and light blue in the nucleotide. Hydrogen bonds are shown as black dashed lines. Amino acid 111 corresponds to the central arginine in Motif B.

Motif C has low sequence conservation and initially was believed to act simply as a connection between the head and neck domains of the protein (Li *et al.*, 2002). However, we have shown (Shan *et al.*, 2008) that changing a single residue (glutamate 189) in motif C in the *Saccharomyces cerevisiae* tRNA nucleotidyltransferase can alter the structure, activity and thermostability of this enzyme such that a temperature-sensitive phenotype results suggesting a role for this domain. Recent data (Ernst *et al.*, 2015) implicate a highly conserved aspartic acid residue in AMP incorporation and in interdomain movements in the enzyme. This further supports a role for motif C suggesting that this region is used as a flexible spring element modulating the relative orientation of the head and body domains to help accommodate the growing 3' end of the tRNA substrate. This transition helps initiate the rearrangement of the templating amino acids in motif D to switch specificity from CTP to ATP during tRNA maturation (Ernst *et al.*, 2015).

Motif D forms a nucleotide binding pocket specific for CTP or ATP (Li *et al.*, 2002) (Fig. 1.4) and discriminates against UTP and GTP (Yue *et al.*, 1996, Shi *et al.*, 1998, Cho *et al.*, 2007). Three well conserved amino acids, glutamate, aspartate, and arginine form Watson-Crick-like hydrogen bonds with

incoming nucleotides (Cho *et al.*, 2007, Li *et al.*, 2002) to recognize specifically cytidine, cytidine and adenosine bases and discriminate against incorrect ribonucleotides in the binding pocket of the protein (Li *et al.*, 2002) (Fig. 1.4). After the addition of the first two CTPs, the binding pocket is reorganized and enlarged to change the specificity of the pocket from CTP recognition to ATP recognition (Li *et al.*, 2002). Site-directed mutagenesis of the appropriate amino acids resulted in the reengineering of the CCA-adding enzyme into a UUG- or dCdCdA-adding enzyme or a poly(C,A) or poly(U,G) polymerase (Cho *et al.*, 2007).

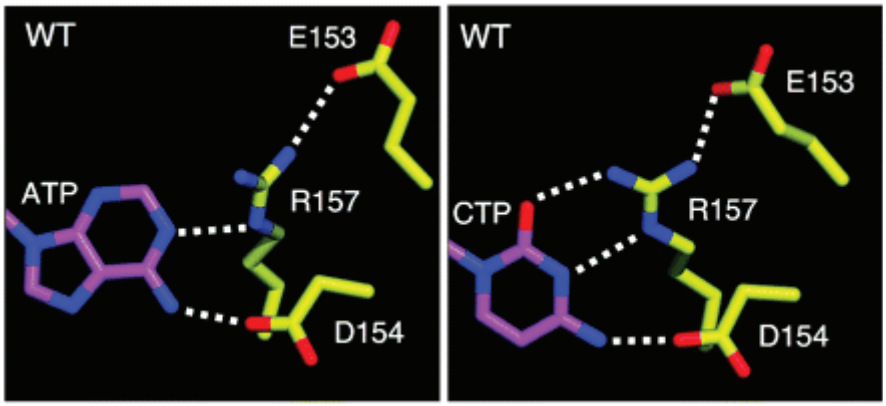


Figure 1.4: Watson-Crick-like base pairing of specific tRNA nucleotidyltransferase residues with ATP or CTP (Betat *et al.*, 2010, Li *et al.*, 2002). The nucleotide binding site is formed by the EDxxR conserved amino acids found in motif D (Li *et al.*, 2002). The incoming CTP or ATP forms three or two Watson-Crick-like hydrogen bonds, respectively.

The function of motif E is not well understood, however, it may help in stabilizing the helix-turn structure in motif D (Li *et al.*, 2002) and might interact with the growing 3'end of the tRNA substrate (Cho *et al.*, 2007).

While motifs A-E largely are conserved among CCA-adding, CC-adding and A-adding enzymes, the *A. aeolicus* CC-adding enzyme differs at motif D in that the conserved glutamic acid residue is replaced by an arginine which through repulsion with the conserved arginine constrains the NTP binding pocket to favour CTP binding by preventing the larger ATP from entering (Li *et al.*, 2002). Furthermore, a

flexible loop between motifs A and B is present in CCA-adding enzymes but not in CC-adding enzymes (Neuenfeldt *et al.*, 2008). A single glycine to aspartate substitution in this region of the *E. coli* CCA-adding enzyme altered AMP but not CMP incorporation (Zhu *et al.*, 1986). Interestingly, transplanting the conserved loop into a CC-adding enzyme restored complete CCA addition (Neuenfeldt *et al.*, 2008) further demonstrating the link between structure and function. For A-adding enzymes, the biochemical basis for their restricted activity is poorly understood (Betat *et al.*, 2010), however a number of replacements made between motifs A and E in the A-adding enzyme of *A. aeolicus* interfered with enzymatic activity (Tomita *et al.*, 2004).

1.3.1.2 The carboxy-terminal portion

Much less is known about the less well conserved C-terminal portion of the various tRNA nucleotidyltransferases, however, this region of the protein must play an important role in CCA addition as a deletion of the last two α-helices in the *E. coli* CCA-adding enzyme abolishes both AMP addition and CMP addition (Zhu and Deutscher, 1987). It was first proposed that the body and tail domains are involved in tRNA binding (Li *et al.*, 2002), and the crystal structure of the *A. aeolicus* A-adding enzyme with tRNA bound supports this (Tomita *et al.*, 2004). In addition, domain swapping experiments between bacterial poly(A) polymerase and the CCA-adding enzyme showed that the specificities of the two enzymes were dependent on their C-terminal regions with the C-terminus of the CCA-adding enzyme playing an anchoring role to limit polymerization to three nucleotides (Betat *et al.*, 2004). In addition, the C-terminal portion of the *E.coli* CCA-adding enzyme has a Ni^{2+}-dependent phosphatase activity, a 2'-nucleotidase phosphodiesterase activity, and a metal-independent 2',3'-cyclic activity which may act in concert with the CCA-adding activity to repair the 3’ ends of tRNAs (Yakunin *et al.*, 2004). This C-terminal HD domain is conserved among other bacterial CCA-adding enzymes (Yakunin *et al.*, 2004) but is not apparent in any of the eukaryotic CCA-adding enzymes. This does not exclude the possibility that the C-terminal regions of eukaryotic CCA-adding enzymes may have as yet uncharacterized

functions particularly since the CCA-adding enzymes in eukaryotes tend to be larger than those in bacteria. While the *E. coli* enzyme has 414 amino acids, the mature functional domains of the *Kluyveromyces lactis*, *Saccharomyces cerevisiae*, *Candida glabrata* and *Arabidopsis thaliana* proteins are 486, 529, 533 and 542 amino acids in length, respectively. In contrast, the mature human protein contains only 405 amino acids.

1.3.2 Localization, activity and stability

Protein synthesis in eukaryotic organisms is compartmentalized and takes place in the cytosol, mitochondria and plastids. Although prokaryotic-type ribosomes are used in mitochondrial and plastid protein synthesis and eukaryotic-type ribosomes are used in cytosolic protein synthesis, some features of translation are conserved in these multiple locations. One of these features is the need for a complete CCA sequence on tRNAs that can function in protein synthesis. This requires that tRNA nucleotidyltransferase be present in each of these intracellular locations as, in general, eukaryotic tRNA genes do not code for the 3'-terminal CCA sequence (Jühling *et al.*, 2009). Given that in the eukaryotes characterized to date (Chen *et al.*, 1992, Reichert *et al.*, 2001, Nagaike *et al.*, 2001, Keady *et al.*, 2002, Shanmugam *et al.*, 1996, Deng *et al.*, 2000, Hanic-Joyce and Joyce, 2002), there is a single gene coding for tRNA nucleotidyltransferase, this means that the gene generates products belonging to the class of proteins known as sorting enzymes (Zoladek *et al.*, 1995) where the gene products are distributed to more than one cellular destination. This distribution may be defined by the use of different transcription, *e.g.*, (Chen *et al.*, 1992) or translation (Wolfe *et al.*, 1994) start sites, post-transcriptional processing (Keady *et al.*, 2002), post-translational modifications (Zehorai *et al.*, 2010) or a combination of these (Martin and Hopper, 1994). In addition to classical amino-terminal targeting signals (Kunze and Berger, 2015), specific *cis* elements within the protein itself also have been shown to be involved in targeting, *e.g.*, nuclear import and cytoplasmic retention/nuclear export sequences in Mod5p (Tolerico *et al.*, 1999).

In addition to its role in tRNA maturation in the cytosol, mitochondria and plastids where protein synthesis takes place, CCA addition also may be required in the nucleus where addition of the CCA sequence and aminoacylation appear necessary for the export of some, but not all, tRNAs (Sarkar *et al.*, 1999, Steiner-Mosonyi and Mangroo, 2004). Also, crystallographic data demonstrate interactions between the 3′-CCA overhang of tRNAs and residues of the exportin Los1/Xpot component of the nuclear export machinery (Cook *et al.*, 2009). In spite of the importance of the CCA sequence in nucleocytoplasmic trafficking of tRNAs, the mechanism for sending tRNA nucleotidyltransferase into the nucleus remains unknown. Understanding this process is important as nucleocytoplasmic shuttling of tRNAs has been implicated in the regulation of processes as diverse as protein synthesis (Hopper, 2013), the cell cycle (Ghavidel *et al.*, 2007), and response to nutrient availability (Whitney *et al.*, 2007). Given that tRNAs may undergo retrograde transport into the nucleus under specific stress conditions (Murthi *et al.*, 2010, Hurto *et al.*, 2007, Shaheen and Hopper, 2005, Pierce *et al.*, 2010) and that tRNA nucleotidyltransferase functions both in the cytosol and the nucleus, we expect that tRNA nucleotidyltransferase localization also may be regulated. The importance of cytosolic tRNA nucleotidyltransferase was highlighted by studies in yeast showing that artificially increasing the nuclear pool of tRNA nucleotidyltransferase while decreasing the cytoplasmic pool resulted in cells that accumulated end-shortened tRNAs and grew slowly (Wolfe *et al.*, 1996). Moreover, even in *E. coli* where all of the tRNA genes encode a complete CCA sequence, cells grow less well in the absence of tRNA nucleotidyltransferase suggesting a repair function for this enzyme (Zhu and Deutscher, 1987) in this organism. Evidence also shows that tRNA nucleotidyltransferase has the ability to add a CCACCA sequence to the ends of unstable mature tRNAs (Wilusz *et al.*, 2011, Kuhn *et al.*, 2015). In doing so, it serves as a universal quality control mechanism for controlling tRNA levels and preventing errors in translation (Wilusz *et al.*, 2011). While the oxidative-stress activated nuclease angiogenin can remove the CCA sequence from tRNAs when cells are exposed to stress and reduce energy expenditure through

protein synthesis, tRNA nucleotidyltransferase can repair this damage and reactivate translation providing a mechanism to dynamically repress and reactivate translation at low metabolic costs (Czech *et al.*, 2013).

1.4 Outline and scope of the thesis

Given its central importance in cell viability (required for eukaryotic protein synthesis), its complicated and intriguing catalytic activity (it requires three different substrates and adds a specific nucleotide sequence without utilizing a nucleic acid template) and its intracellular localization in eukaryotic cells (distributed between the nucleus, cytosol, mitochondria and plastids), tRNA nucleotidyltransferase represents a rare example of a single protein that can be studied to explore the complex roles that localization, activity, structure and stability play on protein function in the cell.

This study was initiated to explore the interplay between protein structure, stability, activity and localization in the context of tRNA nucleotidyltransferase. Previously, the mature domain of the Arabidopsis enzyme (far from the amino-terminal targeting signal) had been shown to have a major impact on the distribution of tRNA nucleotidyltransferase within the cell (Schmidt von Braun *et al.*, 2007). Specifically adding the mature domain (in addition to the amino-terminal targeting signals) of tRNA nucleotidyltransferase to a reporter protein shifted targeting away from mitochondria and toward plastids. These data indicate that the regulation of dual localization involves not only the N-terminal signals, but also additional factors within the protein (folding competence, regions interacting with substrates or *trans*-acting targeting factors).

In Chapter 2 of the thesis, the enzymatic and physicochemical properties of tRNA nucleotidyltransferase and two variants are explored to correlate the properties of the enzyme with the intracellular distribution of the protein. Two sets of amino acid substitutions, Mat-Mu1 at amino acids 399-420 in the body domain and Mat-Mu2 at the extreme C-terminus of the protein were introduced

and the effects on protein structure, stability, activity and localization were monitored. The data show that changes in the body and tail domains of the Arabidopsis tRNA nucleotidyltransferase can alter the stability of the protein and reduce catalytic activity *in vitro*. In both cases, the distribution of the modified protein was shifted from mitochondrial targeting toward plastid targeting in protoplasts. Studies with a tightly folded Titin domain attached to tRNA nucleotidyltransferase also showed that plastid import was preferred over mitochondria import suggesting that the altered localization is dependent on the altered stability of tRNA nucleotidyltransferase and not on its reduced activity. This has been published as Leibovitch *et al.* (2013).

In Chapter 3, the specific amino acids in Mat-Mu1 responsible for the alterations in structure and enzyme activity were defined. Mat-Mu1 first was divided into two regions (Mat-Mu1A and Mat-Mu1B) and using the same combination of biophysical and biochemical techniques as used previously, the amino acid substitutions linked to reduced enzyme activity were mapped to Mat-Mu1B. Subsequently, single point mutations were introduced into this domain and the lysine at position 418 was shown to be most responsible for tRNA binding. If targeting is dependent on substrate binding then this substitution should be informative. In contrast, if changes in structure or interactions with another *trans*-acting factor either in the cytosol or at the membrane are important for targeting then the amino acid substitutions in Mat-Mu1A which affect structure but not activity should be informative. This work will be submitted for publication when the localization experiments have been completed in the laboratory of Dr. Enrico Schleiff, Goethe University.

In Chapter 4, a second model organism, *Saccharomyces cerevisiae*, is used to explore how a single amino acid substitution can alter enzyme structure, activity and stability. Previously, it was shown (Shan *et al.*, 2008) that changing a single residue (glutamate 189) in motif C in the *Saccharomyces cerevisiae* tRNA nucleotidyltransferase alters the activity and thermostability of this enzyme such that a temperature-sensitive (ts) phenotype results. Here, a combination of random and site-directed

mutagenesis, genetic, biophysical (*e.g.*, fluorescence and circular dichroism spectroscopy and thermal denaturation) and biochemical techniques (including enzyme assays) are used to explore further the role of this amino acid (in the context of motif C) in enzyme structure, stability and activity to explore the cause of the temperature-sensitive phenotype and to elucidate the function of motif C in tRNA nucleotidyltransferase structure and activity. A suppressor mutation changing arginine at position 64 to tryptophan did not increase thermostability but did restore enzyme activity such that cells grew at the restrictive temperature. This suggests that the temperature-sensitive phenotype is the result of a hypomorphic effect reflecting enzyme activity and not stability, *i.e.*, the ts phenotype arises not because tRNA nucleotidyltransferase is unstable at the restrictive temperature, but because its reduced activity (which is sufficient for growth at the permissive temperature) does not meet the needs of the cell at the elevated temperature. These data also demonstrate the critical role motif C residue 189 actually plays in the organization of the head and neck regions of the enzyme's active site. Furthermore, the suppressor mutation suggests a possible interaction between motifs A and C. This work has been published as Goring *et al*. (2013).

In Chapter 5, based on what was learned from the studies on the Arabidopsis and yeast enzymes, the effects of amino acid substitutions on the structure, stability and activity of the human tRNA nucleotidyltransferase were explored. In a recent study (Chakraborty *et al.*, 2014), patients with congenital sideroblastic anemia with B-cell immunodeficiency, periodic fevers, and developmental delay (SIFD) were shown to carry mutations leading to partial loss of function of TRNT1, the human tRNA nucleotidyltransferase. Five of the mutations were found in the well characterized motifs A-E of class II tRNA nucleotidyltransferases while two were found within the less well conserved C-terminal body and tail portions of the enzyme. Here, the same combination of biophysical and biochemical techniques used with the Arabidopsis and yeast enzymes was used to explore the variant human tRNA nucleotidyltransferases. Many of the variants showed the same general characteristics. For example, all

of the variants showed similar secondary and tertiary structures and all variants (except for K416E) showed reduced thermal stability. All of the variants showed levels of *in vitro* enzyme activity similar to those of the native enzyme except for the I326T and K416E variants. This is intriguing as these two changes are in the C-terminal portion of the protein while the others are in the more conserved catalytic domains. A more detailed kinetic analysis revealed that while the I326T variant showed a ten-fold drop in turnover number and no change in K_M for tRNA, the K416E variant showed no real change in turnover number but at least a seven-fold increase in K_M for tRNA. Finally, both the I326T and K416E variants showed altered quaternary structures as compared to the native enzyme with the I326T variant showing an increased ratio of monomer to dimer while the K416E variant showed an increased ratio of dimer to monomer as compared to the active enzyme. Taken together all of these data suggest that while point mutations leading to changes in the catalytic domains primarily alter stability which may lead to less functional protein and the disease phenotype, changes made at the C-terminus of the protein affect both activity and binding of the tRNA substrate which contribute to impaired function of the tRNA nucleotidyltransferase leading to the disease phenotype. Additionally, the altered quaternary structures seen with the I326T and K416E variants may result in mislocation of the enzyme. Altogether, these data suggest that the complex phenotype seen in SIFD patients may result from different defects in tRNA nucleotidyltransferase which affect its structure, stability, activity, location or various combinations of some or all of these. This work will be submitted for publication when the *in vivo* localization experiments in mammalian cell lines have been completed in the laboratories of our collaborators at the Children's Hospital of Eastern Ontario.

1.5 Objectives of the thesis

The overall key objectives of this thesis are:

(I) To characterize what properties of Arabidopsis tRNA nucleotidyltransferase influence its intracellular distribution.

(II) To define more precisely the role of specific amino acids in a region of the Arabidopsis tRNA nucleotidyltransferase where amino acid changes were shown to alter localization.

(III) To address the role of glutamate 189 in yeast tRNA nucleotidyltransferase structure and function and to investigate whether defects in thermal stability or activity or both are responsible for the temperature-sensitive phenotype.

(d) To explore the effects on tRNA nucleotidyltransferase structure, stability and activity of seven amino acid substitutions resulting from seven point mutations mapped to the human TRNT1 gene in patients with SIFD.

2.0 The folding capacity of the mature domain of the dual-targeted plant tRNA nucleotidyltransferase influences organelle selection

2.1 Preface

The work presented in Chapter 2 was publish in: Leibovitch M., Bublack D., Hanic-Joyce P.J., Tillmann B., Flinner N., Amsel D., Scharf D., Mirus O., Joyce P.B.M., and Schleiff E. (2013). **The folding capacity of the mature domain of the dual-targeted plant tRNA nucleotidyltransferase dictates the organelle selection**. Biochem J. 453(3):401–412. Paul Joyce and Enrico Schleiff designed the experiments and drafted the manuscript. Matthew Leibovitch performed the biophysical and biochemical characterization and Enrico Schleiff was involved in their analysis; Pamela Hanic-Joyce performed the yeast complementation experiments; Daniela Bublak, Bodo Tillmann and Dieter Scharf performed the *in vivo* import studies; Nadine Flinner, Daniel Amsel and Oliver Mirus performed the bioinformatic analysis. All authors contributed during manuscript preparation and approved the content.

2.2 Abstract of Manuscript

tRNA-NTs (tRNA nucleotidyltransferases) are required for the maturation or repair of tRNAs by ensuring that they have an intact cytidine-cytidine-adenosine sequence at their 3'-termini. Therefore this enzymatic activity is found in all cellular compartments, namely the nucleus, cytoplasm, plastids and mitochondria, in which tRNA synthesis or translation occurs. A single gene codes for tRNA-NT in plants, suggesting a complex targeting mechanism. Consistent with this, distinct signals have been proposed for plastidic, mitochondrial and nuclear targeting. Our previous research has shown that in addition to N-terminal targeting information, the mature domain of the protein itself modifies targeting to mitochondria and plastids. This suggests the existence of an as yet unknown determinate for the distribution of dual-targeted proteins between these two organelles. In the present study, we explore the enzymatic and physicochemical properties of tRNA-NT variants to correlate the properties of the

enzyme with the intracellular distribution of the protein. We show that alteration of tRNA-NT stability influences its intracellular distribution due to variations in organelle import capacities. Hence the fate of the protein is determined not only by the transit peptide sequence, but also by the physicochemical properties of the mature protein.

2.3 Introduction

Protein synthesis in plants is compartmentalized and takes place in the cytosol, plastids, and mitochondria. Although eukaryotic-type ribosomes are used in the cytosol and prokaryotic-type ribosomes are used in plastids and mitochondria, some features of the translational machinery are conserved. The requirement for tRNAs for translation represents one of these universal features. Whether the tRNAs are encoded by the nuclear genome, or by the plastidic or mitochondrial genomes (Jühling *et al.*, 2009), they must be processed and modified (Phizicky and Hopper, 2010) to function in protein synthesis. One essential modification is the addition of the 3'-terminal cytidine–cytidine–adenosine (CCA) sequence which is required for aminoacylation (Xiong and Steitz, 2006, Betat *et al.*, 2010). As no plant tRNA genes encode this sequence (Jühling *et al.*, 2009), it must be added post-transcriptionally. The CCA adding enzymes found in plants belong to the class II tRNA nucleotidyltransferase (tRNA-NT) family. They are composed of four regions, the N-terminal head domain, the neck domain, the body region and the C-terminal tail (Fig. 2.1A) (Toh *et al.*, 2009).

A tRNA-NT enzyme is required in the translationally active compartments, namely the cytosol, mitochondrion and, in plants, the plastids. Moreover, there also are nuclear pools of tRNA-NT thought to be involved in the maturation of cytosolic tRNAs in yeast and mammals (Phizicky and Hopper, 2010). This nuclear CCA addition may serve as a quality control point prior to the export of mature tRNAs to the cytosol *e.g*. Zasloff, (1983), Zasloff *et al.*, (1982). The nuclear localization of tRNA-NT also might reflect a function of this enzyme as a nucleus/cytosol-shuttling protein involved in tRNA transport (Feng and

Hopper, 2002). In the cytosol tRNA-NT is thought to catalyze the repair of damaged CCA sequences (Lizano *et al.*, 2008) but may have an additional role in quality control. Recent studies indicate that tRNA-NT may recognize tRNAs with flexible acceptor stems and guanosines at the first and second positions and add a CCACCA sequence to their 3'-termini thereby facilitating their entry into the rapid tRNA decay pathway (Wilusz *et al.*, 2011).

To distribute tRNA-NT between organelles and the cytosol, multiple isoforms of the protein are synthesized from a single gene in fungi (Chen *et al.*, 1992, Deng *et al.*, 2000, Hanic-Joyce and Joyce, 2002), mammals (Reichert *et al.*, 2001, Nagaike *et al.*, 2001), and plants (Shanmugam *et al.*, 1996, Schmidt von Braun *et al.*, 2007). As such these tRNA-NTs are representative of a group of enzymes designated sorting isozymes (Gillman *et al.*, 1991), whose prevalence is greater than initially expected e.g. Carrie *et al.*, (2009). Other prominent examples of sorting isozymes include other nuclear-encoded tRNA modifying enzymes (Martin and Hopper, 1994) and the aminoacyl-tRNA synthetases (Duchêne *et al.*, 2005). Typically, these sorting isozymes are generated by alternative use of multiple transcription start sites and multiple in-frame start codons which give rise to long variants with amino-terminal mitochondrial and/or plastid transit peptides and shorter variants with cytosolic and/or nuclear functions.

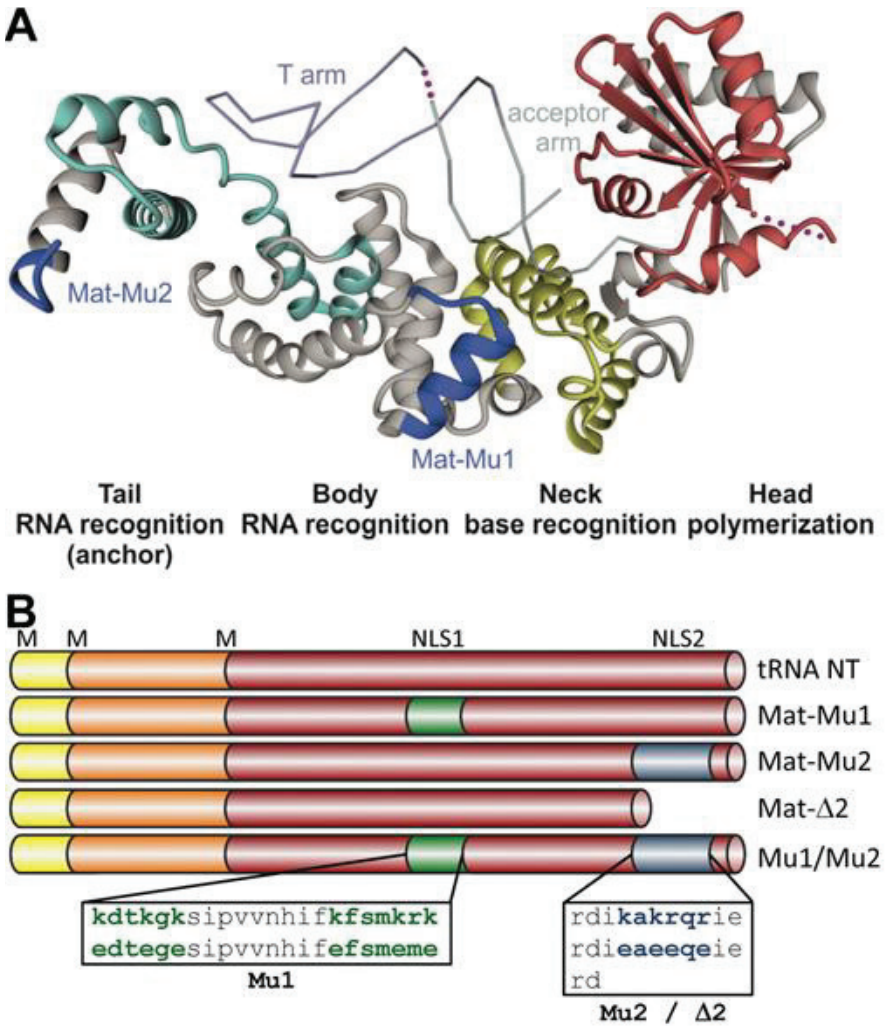


Figure 2.1 *A. thaliana* tRNA-NT and variants
(A) The X-ray structure of *Thermotoga maritima* CCA (PDB code 3H37 (Toh *et al.*, 2009) is shown in ribbon representation with a tRNA (from *Aquifex aeolicus* PDB code 1VFG (Tomita *et al.*, 2004) positioned based on the structural alignment of 1VFG with 3H37 (as shown in Supplementary Figure S2.2). The position of the T-arm and the acceptor arm are indicated. Regions in the tRNA and protein not solved in the corresponding crystal structures are indicated by magenta dots. In the structure the positions of Mat-Mu1 and Mat-Mu2 (dark blue) and the recognized parts of the PFAM domains are indicated (red, PolyA_pol; yellow, PolyA_polRNAbd; blue, tRNA_NucTran2_2/HD). The use of 3H37 for visualization of the *A. thaliana* protein motifs was justified by fold recognition (Supplementary Table S2.3). (B) The various versions of tRNA-NT used in the present study are depicted as coloured bars showing the mature domain translated from the third methionine residue in red, the additional region translated from the second methionine residue in orange and the additional region translated from the first methionine residue in yellow. Mat-Mu1 and Mat-Mu2 where amino acid substitutions were made are shown in green and blue, respectively. Coloured regions are not drawn to scale. The amino acid sequence of Mat-Mu1 and Mat-Mu2, as well as the amino acids changed, are indicated.

We previously demonstrated that different isoforms of *A. thaliana* tRNA-NT can be synthesized either with or without an N-terminal region containing the mitochondrial and plastid targeting information and that differential use of the first two in-frame ATGs could give rise to two sizes of transit peptides of differing intracellular distributions in onion cells (Schmidt von Braun *et al.*, 2007). While proteins with the longer N-terminal signal were targeted to both mitochondria and plastids, proteins with the shorter N-terminal signal were preferentially targeted to plastids. A similar pattern of localization had been observed for the dual-targeted glutathione reductase (Rudhe *et al.*, 2002). These observations, however, might reflect that the signal for efficient plastid translocation can be shorter than the minimal length required for mitochondrial import (Ruprecht *et al.*, 2010, Bionda *et al.*, 2010). Remarkably, in addition to a requirement for the N-terminal domain for mitochondrial and plastid import, the mature domain of Arabidopsis tRNA-NT also influenced the distribution of this protein. The presence of the mature domain in addition to the N-terminal region shifted the localization of the linked GFP reporter away from plastids and toward mitochondria (Schmidt von Braun *et al.*, 2007). Intriguingly, a variant of the mature domain bearing amino-acid substitutions within its C-terminal region lost this preference for mitochondrial targeting and was directed to plastids (Schmidt von Braun *et al.*, 2007). Given recent observations of unfolding and plastid import (Bionda *et al.*, 2010), the distinct translocation behavior of the protein might reflect a difference in stability between the native and variant mature domains.

To address how the physicochemical properties of the mature domain affect the targeting of Arabidopsis tRNA-NT to mitochondria and plastids we analyzed the structure, stability and tRNA-binding of the native and variant proteins in the context of intracellular localization. We show that for tRNA-NT, targeting is influenced by changes in protein stability. Further, by using the tRNA-NT N-terminal targeting signal fused to Titin-I27 or GFP, we show that the distribution of proteins in the cell is defined

in part by the portion of the protein following the targeting signal and not just only by the targeting signal.

2.4 Materials and Methods

2.4.1 Plasmid construction

For the generation of GFP fusion proteins with the C-terminal KAKRQRIE sequence of the *Arabidopsis* tRNA-NT converted into EAEEQEIE, PCR amplification was carried out using primer 1 (Table S2.1) and the appropriate second primer and template as described previously (Sikorski and Hieter, 1989). The resulting product was digested with *Bcl*I and inserted into the *BamH*I site of pBIN35S35SEmGFP. For protein production in *E. coli*, the appropriate clones (Sikorski and Hieter, 1989) were amplified by PCR using combinations of primers 2–6 (Table S2.1): primer 2 (ATG1) or primer 3 (ATG3) and primer 4 (native C-terminus) or primer 5 (changes KAKRQRIE to EAEEQEIE) or primer 6 (removes IKAKRQRIE). The resulting products were digested with *BclI* and *SalI* and cloned into a modified pGEX-2T vector (Mumberg *et al.*, 1995) digested with *Bam*HI and *Sal*I. Expression clones containing the EDTEGE and EFSMEME sequences were generated by moving a *Bpu1102*I fragment containing these mutations (Sikorski and Hieter, 1989) into *Bpu1102*I-digested pGEX-2T containing the appropriate *Arabidopsis* insert. For expression in yeast, tRNA-NT–GFP fusion sequences were excised as *Xba*I fragments from the appropriate pBIN35SEmGFP derivatives and inserted into *Xba*I-digested pG131-2, a pRS313 (Sikorski and Hieter, 1989) derivative in which the *Kpn*I–*Sac*I multiple cloning region was replaced by the *Sac*I–*Kpn*I fragment of p426GPD (Mumberg *et al.*, 1995) containing the *TDH3* promoter and *CYC1* terminator. Constructs were confirmed by sequence analysis (BioS&T or McGill University and Génome Québec Innovation Centre).

2.4.2 Plasmid shuffling

Plasmids expressing various tRNA-NT–GFP fusion proteins were transformed (Schiestl and Gietz, 1989) into strain SCDT8 (*MATα leu2-3,112trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 cca1::TRP1*) bearing a derivative of pRS316 (Sikorski and Hieter, 1989) containing the wild-type *CCA1* gene of *S. cerevisiae* with selection on SC medium minus uracil and histidine (Sherman, 2002). For plasmid shuffling, patches of two representative transformants were replica-plated from SC medium minus uracil and histidine, to YPD medium (Sherman, 2002), incubated 1 day at 23^{o}C, replica-plated to 5-FOA medium (Boeke *et al.*, 1984) supplemented with 30 mg/l leucine, lysine and tyrosine and 20 mg/l adenine, histidine and tryptophan, with or without FOA and incubated for 3 days at 23^{o}C. FOA-resistant cells then were replica-plated to YPD or YPG plates and incubated at 23^{o}C for 3 or 4 days respectively.

2.4.3 Expression and purification of tRNA nucleotidyltransferase

Proteins were expressed in *E. coli* BL21(DE3) (Shan *et al.*, 2008) except that cells were induced with 0.02% w/v D-lactose and 0.5 mM IPTG at 19^{o}C for 16-20 hours. Cells were harvested by centrifugation, stored at -80^{o}C for a minimum of 1 hour, frozen pellets were thawed on ice and resuspended in cold lysis buffer (PBS plus 1 mM EDTA) at 2 ml/g of cell pellet prior to lysis in a French Pressure Cell press (Thermospectonic). Cleared lysate was cycled through glutathione Sepharose fast flow 4B resin (GE Healthcare or Gold Biotechnology USA) overnight (4^{o}C, 1ml/min). The column was washed and the protein eluted with 15 mM reduced glutathione, 50 mM Tris-HCl (pH: 8.3), 140 mM NaCl and 2.5 mM $CaCl_2$. Fractions of interest were digested (5 units/mg) with thrombin (2 188 NIH units/mg, GE Healthcare) while being dialyzed overnight (8 kDa dialysis bag) at 4^{o}C in 50 mM Tris-HCl (pH 8.3), 140 mM NaCl and 2.5 mM $CaCl_2$ to eliminate glutathione. Protein was passed 3x through a glutathione Sepharose fast flow 4B column to remove the GST tag and dialysed against PBS (50 kDa dialysis bag) to remove remaining thrombin and to exchange buffer.

2.4.4 Biophysical characterization of tRNA nucleotidyltransferase

2.4.4.1 UV-visible spectroscopy

The UV-visible spectra (200-400 nm) were collected using the Varian Cary 100 Bio UV-Visible Spectrophotometer using a 1.0 cm quartz cuvette. The approximate amount of tRNA in any sample was calculated using the extinction coefficient of 2.22 x 10^{-2} ml·µg^{-1}cm^{-1} at OD_{260} for tRNA. To remove nucleic acid that co-purified with the proteins of interest, RNaseA (>60 Kunitz units/mg protein, BioShop Canada Inc. Burlington, ON) was added (30 ng RNaseA/ml protein solution). The RNaseA was removed by extensive dialysis against PBS as described above and the efficient removal was confirmed by the stability of an added radiolabeled tRNA template.

2.4.4.2 CD-spectroscopy

Far UV CD spectra (185-280 nm) were recorded on a Jasco-815 CD- spectropolarimeter under a constant nitrogen flow in a 0.1 or 0.2 cm cell, using a protein concentration of 0.1-0.3 mg/ml in PBS (pH 7.4), accumulated over five scans and buffer corrected for those signals. A bandwidth of 1 nm, a response time of 0.25 seconds, a data pitch of 0.2 nm, and a scanning speed of 20 nm/min was used. Prior to measurements, samples were centrifuged (5 min. 4^{o}C). The data obtained were smoothed using the Spectra Analysis program (Jasco). The CD spectrum for tRNA was recorded as for the protein.

2.4.4.3 Thermal denaturation

Thermal denaturation was monitored by CD in a 0.2 cm cell using protein concentrations of 0.1-0.3 mg/ml in phosphate buffer (pH 7.4). The Variable Temperature program (Jasco) was used (222 nm, standard sensitivity) with standard settings and a data pitch of 0.2°C. The rate was 25°C/h with start and end temperatures of 25°C and 90°C, respectively.

2.4.4.4 Fluorescence measurements

Intrinsic fluorescence of the proteins (1.2 μM or 0.8 μM) was measured in PBS (pH 7.4) and 10% glycerol using the Varian Cary Eclipse Fluorescence Spectrophotometer with a 1 cm path length. After excitation at 280 nm or 295 nm, emission spectra were recorded in the range of 310-400 nm with a scan rate of 10 nm/s and excitation and emission bandwidths of 5 nm. Emission spectra at room temperature were averaged over 10 scans with a 1.0 nm data sampling interval and voltage was set to 600 volts. Protein concentrations were determined by OD_{280} and samples were centrifuged for five minutes at $4^{o}C$ before use. All spectra were baseline-corrected and the data obtained were smoothed using the moving average smoothing function (Varian). Baker's yeast tRNA (Roche), extensively purified by phenol extraction and ethanol precipitation was used in fluorescence quenching experiments. Protein (concentrations ranging between 1.0-2.6 μM) was placed in a 1 cm cell and titrated with increasing amounts of tRNA at room temperature. All measurements were taken at 295 nm as described above. UV-visible spectral scans (200-400 nm) were taken after each tRNA addition to take into account any inner filter effect (Shan *et al.*, 2008).

2.4.5 Enzyme activity assays

Plasmids G73 and pmBsDCCA were kindly provided by Dr. Alan Weiner (University of Washington) and digested with restriction enzymes to generate the tRNA-N, tRNA-NC, and tRNA-NCC templates by run-off transcription. Each transcription reaction (100 μl) contained 1X transcription buffer (Fermentas), 500 μM ATP, CTP and UTP, 50 μM GTP, approximately 50 μCi [α-^{32}P] GTP (10 μCi/μl, 3000 Ci/mmol, Perkin Elmer), 13 μg of linearized DNA template, 60 units of T7 RNA polymerase (Fermentas), and nuclease-free water. Each reaction was incubated at 37°C for three hours, phenol extracted, ethanol precipitated and resuspended in 10 μl of nuclease-free dH_2O. To these samples, 10 μl Peattie's loading buffer (Peattie, 1979) was added and samples electrophoresed for 90 minutes on a 7 M urea, 20% acrylamide gel.

Radioactive bands were identified by autoradiography, purified (Rubin, 1973) and resuspended in a final volume of 100-150 µl of nuclease free dH_2O. Standard enzyme assays (Cudny *et al.*, 1978) were carried out in a final volume of 10 µl enzyme reaction mixture which contained 100 mM glycine buffer (pH 9), 10 mM $MgCl_2$, 1 mM ATP, 0.4 mM CTP, 1 µl of the appropriate [α-^{32}P] GTP transcribed tRNA after gel extraction, and the amount of protein as indicated. Assays were performed for the times indicated and stopped by adding 10 µl of loading buffer containing 7M urea (Peattie, 1979). Enzyme kinetic assays were performed in 50 µl containing a fixed amount of [α-^{32}P] GTP transcribed tRNA-NCC template and increasing amounts (0.1 µM to 42 µM) of purified baker's yeast tRNA (Roche). After addition of the appropriate amount of enzyme (5-10 ng), 10 µl aliquots of the reaction mixture were removed at 0.5, 1, 2, 5 and 10 minutes and placed in 10 µl of loading buffer (Peattie, 1979) to stop the reaction. All samples were stored on ice and heated to 70^{o}C for 10 minutes before being loaded onto 7M urea/20% polyacrylamide gels and separated for 12-16 hours at 1950 V. Transcripts were detected using the Typhoon™ TRIO Variable Mode Imager (GE Healthcare) and band intensities measured using the ImageQuantTL 1D version 7.0 (GE Healthcare).

2.4.6 Analysis of *in vivo* localization of the proteins.

The isolation of *Solanum lycopersicum* var. Moneymaker leaf mesophyll protoplasts as well as the polyethylene glycol (PEG) mediated transformation was carried out as described (Mishra *et al.*, 2002). For each transformation 20 µg plasmid DNA was used for 10^5 protoplasts. If necessary the final amount of DNA was adjusted by addition of the vector pRT-Neo which encodes neomycin phosphotransferase only. Incubation was performed in light for 15h at 26°C. The intracellular GFP localization was analyzed using a Leica SP5 confocal laser scanning microscope (Leica, Germany). GFP and chlorophyll were excited at 488 nm and their fluorescence emission recorded at 495-535 and 660-710 nm, respectively.

2.4.7 Bioinformatic analysis and assignment of *A. thaliana* to the PDB structures

Annotated bacterial and eukaryotic tRNA-NT sequences were downloaded from the refSeq database and aligned with MAFFT (Katoh *et al.*, 2005, Katoh and Toh, 2008, Katoh and Toh, 2010). The alignment was used to build pHMMs with hmmer (http://hmmer.janelia.org/) for the neck and the body/tail domains of bacterial and eukaryotic proteins, resulting in four pHMMs. The nr-database (http://www.ncbi.nlm.nih.gov/) was searched with the four Pfams (Punta *et al.*, 2012). pHMMs [PF01743 (head), PF12627 (neck), PF13735 (body/tail) and PF01966 (body/tail)] for all available tRNA-NTs and sequences, which matched at least two pHMMs for different domains, were accepted. Poly(A) sequences were eliminated by CLANS analysis (Figure S2.1). Organisms for which the complete genome sequence was not available (http://www.ncbi. nlm.nih.gov/RefSeq) were excluded. Proteins derived from the same gene locus are treated as a single sequence.

The tRNA-NT sequences were collected from the nr-database using Blast (Altschul *et al.*, 1997) resulting in a dataset with 971 sequences, which was subsequently reduced with cd-hit (Li and Godzik, 2006) to a maximal sequence identity of 70% leaving 409 sequences in the dataset. These sequences were aligned with MAFFT v6.811b (Katoh *et al.*, 2005, Katoh and Toh, 2008, Katoh and Toh, 2010). This MSA (multiple sequence alignment) was compared with the alignment generated by the fold recognition server HHpred (Soding *et al.*, 2005) in order to validate the results and mark reliable regions. The server recommended the structure of *Thermotoga maritima* as the best hit, the structure of *Geobacillus stearothermophilus* as second best hit and the structure of the A-adding enzyme with bound tRNA of *Aquifex aeolicus* as the third hit as possible templates. This result was confirmed by another fold recognition server (Table S2.3). The predicted position for Mat-Mu1 and -Mu2 is the same for all three templates and especially the internal position of Mat-Mu1 is also confirmed by the MSA.

2.5 Results

2.5.1 The structural properties of native and variant tRNA nucleotidyltransferase

Our previous studies of a tRNA–NT-Mat-Mu1 variant, bearing modifications within the body domain of tRNA-NT (Fig. 2.1, Fig. S2.2), demonstrated that properties of this domain influence the intracellular distribution of this protein. To analyze the effect of changing another region within the mature domain we generated a variant, Mat-Mu2, bearing alterations at the extreme C-terminus of the tail domain (Fig. 2.1B, Fig. S2.3). Both Mat-Mu1 and Mat-Mu2 variants were expressed in *E. coli,* purified and the GST tag was removed yielding proteins of high purity (Fig. S2.3).

To address the hypothesis that protein stability influences intracellular distribution we analyzed the structural integrity of both variants by CD spectroscopy. For the native and Mat-Mu2 variant we observed comparable CD spectra, while for the Mat-Mu1 variant the CD spectrum suggested a significant reduction of secondary structure (Fig. 2.2A). However, the UV-visible absorption spectra of the native and Mat-Mu1 variant, but not of the Mat-Mu2 variant showed a significant absorption at 260 nm (Fig. S2.4). Given that nucleic acids and nucleotides absorb strongly at this wavelength and that tRNAs have a positive CD-signal with peak intensity around 220 nm (Fig. S2.4), we hypothesized that the native and Mat-Mu1 variants had retained *E. coli* RNAs and specifically tRNAs. Consistent with tRNA association with these proteins, dialysis using tubing with a 50 000 Dalton cut off, but not with an 8000 Dalton pore size reduced the UV-visible absorbance of the Mat-Mu1 variant.

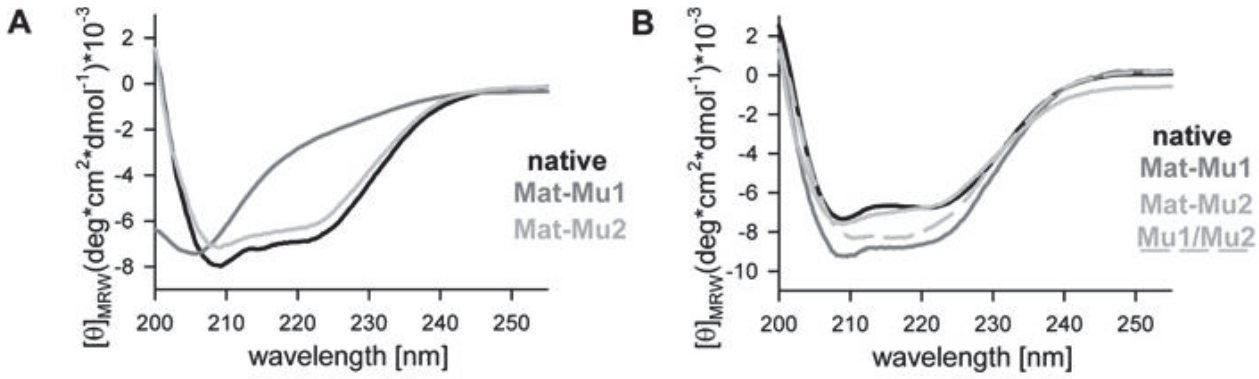


Figure 2.2 tRNA is co-purified with the tRNA-NT
(A) The CD spectra of native protein (4.24 μM, black), and Mat-Mu1 (3.04 μM, dark grey) and Mat-Mu2 (4.01 μM, light grey) variant proteins as purified from *E. coli* (Supplementary Figure S2.3). (B) The CD spectra after removal of tRNA by extensive dialysis with RNaseA treatment for native protein (2.5 μM, black), or without RNAseA treatment for Mat-Mu1 (1.19 μM, dark grey), Mat-Mu2 (4.01 μM, light grey) and Mu1/Mu2 (2.72 μM, light grey broken line, Figure 2.1B) variant proteins.

Furthermore, treatment with RNaseA followed by additional dialyses reduced the UV-visible absorption peak at 260 nm for both proteins (Fig. S2.4, Table S2.2). While the RNaseA treatment yielded an identical CD-spectrum for the native enzyme both before and after treatment (Fig. 2.2B), the CD-spectrum for the Mat-Mu1 variant was substantially different from the CD spectrum before dialysis. After dialysis the Mat-Mu1 variant showed a CD-spectrum similar to the one of the native protein and the Mat-Mu2 variant. From the analysis of the CD-spectra we conclude that the variants are folded and have a primarily α-helical content as expected from the available crystal structures (Fig. 2.1A, supp. Fig. S2.2), although slight variations in the CD-spectra suggest that the folding is not identical. Thus, the mutations introduced at Mat-Mu1 and Mat-Mu2 did not alter the overall secondary structure of the proteins to a large extent.

Unlike Mat-Mu1 and the native enzyme, the Mat-Mu2 variant did not show a major UV-visible absorption peak at 260 nm (Fig. S2.4) which suggests that regardless of what accounts for the difference between the native enzyme and the Mat-Mu1 variant, much less tRNA remains associated with the Mat-Mu2 variant throughout the purification protocol. To determine whether the Mat-Mu1 and Mat-Mu2

modifications would be additive, synergistic or antagonistic we generated a protein containing both modifications. This Mu1/Mu2 variant showed a CD-spectrum that was intermediate to the other analyzed variants (Fig. 2.2B).

Fluorescence spectra (Fig. S2.5A, Table 2.1) on the untreated samples revealed little red or blue shifting for either the Mat-Mu1 or Mat-Mu2 variant (1-4 nm), but did show a major change in fluorescence intensity (40-50% reduction) when compared to the native enzyme. After samples had been treated to remove RNA a slightly larger and more consistent blue shift (4-7 nm) but a lower reduction in fluorescence intensity (10-15%) was observed for both the Mat-Mu1 and Mat-Mu2 variants. Interestingly, the Mu1/Mu2 variant shows emission maxima nearly equal to either the Mat-Mu1 or Mat-Mu2 variant at both wavelengths and shows a small decrease in fluorescence intensity. Fluorescence intensity and wavelength maxima are defined primarily at 295 nm excitation by tryptophan and at 280 nm by both tryptophan and tyrosine. As seven of the 11 tryptophan residues contained in the Arabidopsis tRNA-NT are between Mat-Mu1 and Mat-Mu2 (Fig. S2.2) these changes in fluorescence intensity may highlight changes in the structure of the tRNA-NT in this region. Moreover, there are clearly differences in these signals before and after removal of RNA (compare Fig. S2.5A and S2.5B).

2.5.2 Activity of tRNA-NT and its variants

To determine if differences in the levels of tRNA associated with the native and variant enzymes reflected differences in enzyme activity we assessed the ability of the purified enzymes to support CCA addition *in vitro*. As expected, the native enzyme added a complete CCA sequence to a model tRNA substrate under standard assay conditions (Fig. 2.3, 100 ng protein). Although the affinity for tRNA of the Mat-Mu2 variant appeared to be reduced (Fig. S2.4, S2.5) we observed that this enzyme also adds a complete CCA sequence (Fig. 2.3, 100 ng). In contrast, the Mat-Mu1 variant and the Mat-Mu1/Mu2 variant showed a complete loss of activity (Fig. 2.3, 100 ng). This was the reverse of what might have

been expected based on the absorbance peaks at 260 nm (Fig. S2.4) which suggested that tRNAs were associated with the Mat-Mu1 variant, but not with the Mat-Mu2 variant. However, the activity of the Mat-Mu2 variant is reduced when compared to the native enzyme, which becomes more obvious at reduced enzyme concentrations (Fig. 2.3, 20 ng protein).

Table 2.1 Physiochemical parameters of tRNA-NT and its variants

K_d is taken from Supplementary Figure S6 (at http://www.biochemj.org/bj/453/bj4530401add.htm). F (with excitation at 295 nm and 280 nm) is taken from Supplementary Figure S5 (at http://www.biochemj.org/bj/453/bj4530401add.htm). T_m, ΔH and $\Delta\Delta G_{25°C}$ are taken from Figure 4. $\Delta\Delta G_{25°C}$ is the difference of $\Delta\Delta G_{25°C}$ (native) − $\Delta G_{25°C}$ (variant).

			F (excitation at 295 nm)		*F* (excitation at 280 nm)				
tRNA-NT	K_d (μM/r^2)	K_m/k_{cat} (μM · s^{-1})	F_{max} (nm)	F (I_{MT}/I_{WT})	F_{max} (nm)	F (I_{MT}/I_{WT})	T_m (K)*	ΔH (kJ/mol)	$\Delta\Delta G_{25°C}$ (kJ/mol)
Native	0.66/0.991	11 ± 4/4.0 ± 0.4	341/345†	–	341/344	–	316.3‡	1001.7‡	–
						–	355.2§	391.2§	–
Mu1	0.56/0.990	–	341/338	0.55/0.91	337/340	0.64/0.93	306.8	615.8	39.7
Mu2	1.82/0.989	12 ± 3/2.3 ± 0.2	340/338	0.64/0.83	342/339	0.61/0.84	319.5	147.9	47.9
Δ2	0.20/0.969	16 ± 5/1.6 ± 0.2	–	–	–	–	–	–	–
Mu1/2	–	–	343/338	0.97/0.85	337/340	0.89/0.84	311.6	118.5	52.1

*Values for samples after treatment with RNAseA and re-purification.
†Sample before/after treatment with RNAseA and re-purification.
‡State 1.
§State 2.

To further explore the effects on tRNA binding of altering the enzyme at Mat-Mu1 and Mat-Mu2 we determined tRNA binding constants for the native and variant enzymes by analyzing the quenching of the intrinsic tryptophan fluorescence (Fig. S2.6) as we had done previously for yeast tRNA nucleotidyltransferase (Shan *et al.*, 2008). We found a dissociation constant of about 0.6 µM for both the native and Mat-Mu1 variant proteins indicating that the alteration of amino acids in Mat-Mu1 does not change the ability of the protein to bind tRNA (Table 2.1). This dissociation constant is in good agreement with the values (0.6 to 16 µM for various tRNA substrates) reported previously for class II (Evans and Deutscher, 1978) and class I tRNA nucleotidyltransferases (Shi *et al.*, 1998). The dissociation constant (1.8 µM) of the Mat-Mu2 variant is also well within this range but about three-fold higher than that of the native and Mat-Mu1 variant. Although small, this increased dissociation constant also reflects the reduced association of tRNA with the Mat-Mu2 variant suggested by the association of tRNA with protein during the purification protocol (Fig. S2.4, S2.5). Clearly though, the tRNA association with the

enzyme does not directly define enzyme activity as the Mat-Mu1 variant shows a similar association constant to that of the native enzyme but no activity, while the Mat-Mu2 variant shows an increased association constant and retains activity (Fig. 2.3).

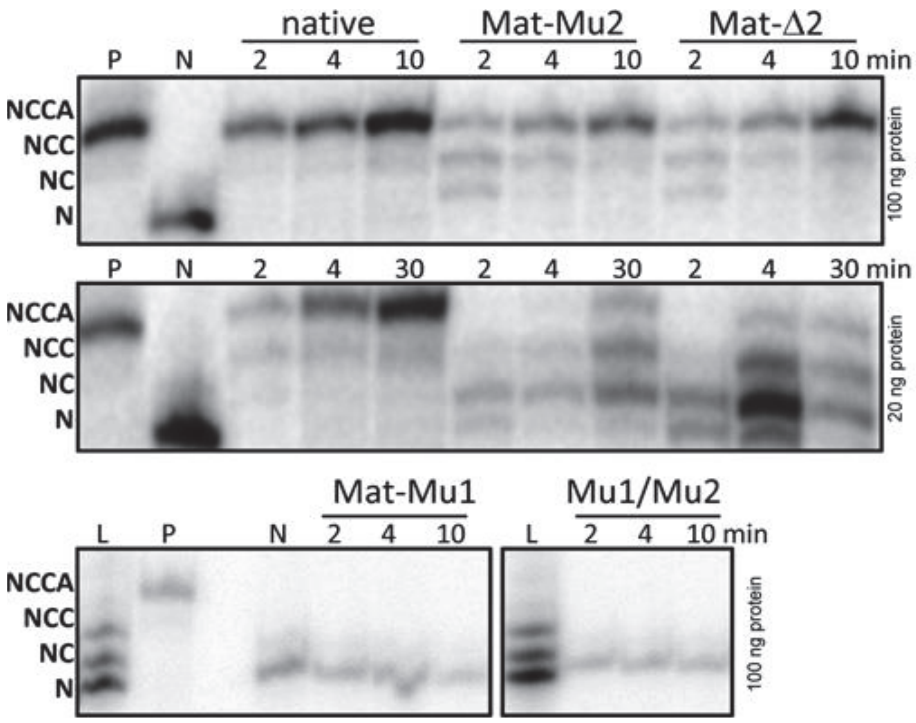


Figure 2.3 The enzymatic activity of native and variant tRNA-NTs
The proteins indicated were incubated as described with a tRNA substrate for the times indicated. The amounts of enzyme used in the assays are indicated to the right of each assay. The reaction products were separated by electrophoresis on 7 M urea/20% polyacrylamide gels and visualized using the Typhoon™ TRIO Variable Mode Imager. The migration of different modification states is shown on the left. P is the active enzyme positive control (tRNA with a complete CCA sequence), N is the boiled enzyme negative control (original tRNA template lacking the CCA sequence) and L represents a ladder containing tRNA products ending at N, NC, NCC and NCCA.

To more precisely link tRNA binding to enzyme activity we used a mixture of a radiolabelled tRNA template lacking the terminal AMP residue and the same yeast tRNAs used for fluorescence quenching as a substrate to define kinetic parameters for *Arabidopsis* tRNA nucleotidyltransferase. In agreement with the activity assays (Fig. 2.3), the Mat-Mu1 variant showed no activity under our assay conditions. In contrast, the Mat-Mu2 variant and the native enzyme showed almost identical apparent K_M values of approximately 12 µM (Table 2.1). It is important to note that although both the association

constants and the kinetic parameters were calculated using a mixture of tRNAs (ending at the discriminator base, the first CMP, the second CMP or the terminal AMP), the apparent kinetic parameters were based on addition of AMP to a template lacking only the terminal AMP, so differences observed for the determination of the K_D and K_M may simply reflect this difference in template. As noted above (Fig. 2.3), the Mat-Mu2 variant was somewhat less active than the native enzyme. This is consistent with a two-fold lower apparent k_{cat} for the Mat-Mu2 variant as compared to the native enzyme although the apparent K_M values of these two enzymes were identical (Table 2.1). Taken together, all of these data indicate that tRNA binding (as measured by K_D or co-purification with enzyme) is not strictly correlated with enzyme activity.

To further address the role of Mat-Mu2 in enzyme activity, we generated and purified a variant of the tRNA-NT lacking the last nine amino acids (Mat-Δ2, Fig. 2.1B). This variant showed a reduced rate of CCA addition *in vitro* when compared to the native enzyme or the Mat-Mu2 variant (Fig. 2.3, 20 ng) with apparent k_{cat} values reduced to 71% and 32% (Table 2.1), respectively. As with the Mat-Mu2 variant, we saw a three-fold change in association constant as compared to the native enzyme, but in this case the change reflected a reduced K_D for the variant enzyme (Fig. S2.6, Table 2.1) such that the difference in association constants between the Mat-Mu2 variant and the Mat-Δ2 variant is approximately 10 fold (Table 2.1). Intriguingly, this difference in association constants is not reflected in differences in apparent K_M's as the apparent K_M's for all three enzymes are within experimental error.

To assess whether *in vitro* tRNA nucleotidyltransferase activities reflect *in vivo* activities, we analyzed whether tRNA-NT-GFP gene constructs could complement a null mutation in *Saccharomyces cerevisiae.* Native and mutant gene fusions, encoding tRNA-NT proteins with all three methionines (M(1,2,3)) or only the third methionine (M(3)) and a C-terminal fused GFP, inserted downstream of the *TDH3* promoter in a high-copy *HIS3*-bearing plasmid, were introduced into a haploid strain (*cca1::TRP1*) carrying plasmid-borne yeast *CCA1* and *URA3* genes. Following selection of Ura^+ His^+ transformants, 5-

FOA was used to select for loss of the *URA3-CCA1*-bearing plasmid and to assess whether variant alleles supported growth. As shown previously for the native enzyme (Schmidt von Braun *et al.*, 2007), the M(3)-tRNA-NT-GFP construct allowed growth on glucose (Table 2.2, Fig. S2.7), confirming the functionality of the fusion protein *in vivo*. For M(1,2,3)-tRNA-NT-GFP we observed growth on fermentable (glucose) and non-fermentable (glycerol) carbon sources, indicating that the fusion protein was successfully imported into yeast mitochondria to facilitate mitochondrial protein synthesis consistent with earlier results for other plant proteins (Chaumont *et al.*, 1990, Bowler *et al.*, 1989). The Mat-Mu2 variants of M(3)-tRNA-NT-GFP and M(1,2,3)-tRNA-NT-GFP behaved like the native enzyme, while the Mat-Mu1 variants, were unable to grow on either glucose or glycerol supporting the lack of enzyme activity observed *in vitro*.

Table 2.2 Ability of native and variant Arabidopsis tRNA-NTs to replace yeast tRNA-NT

Start	Variant	Glucose	Glycerol
(M1,2,3)	Native	+	+
(M3)	Native	+	−
(M1,2,3)	Mat-Mu1	−	−
(M3)	Mat-Mu1	−	−
(M1,2,3)	Mat-Mu2	+	+
(M3)	Mat-Mu2	+	−

2.5.3 The structural stability of native and variant tRNA nucleotidyltransferases

Next we explored whether the enzyme activity or the ability to associate with tRNA is related to protein stability. Following the CD-signal at 222 nm as an indication of α-helicity we observed a three-state equilibrium for the native enzyme with transition temperatures at 316 K and 355 K, respectively. The first transition represents a small change of the secondary structure content, while the second transition reflects a full loss of secondary structure. From the unfolding trace we can conclude that the

intermediate state is rather stable (Fig. 2.4). The value found for the first transition is consistent with the one reported for the yeast tRNA nucleotidyltransferase (Shan *et al.*, 2008).

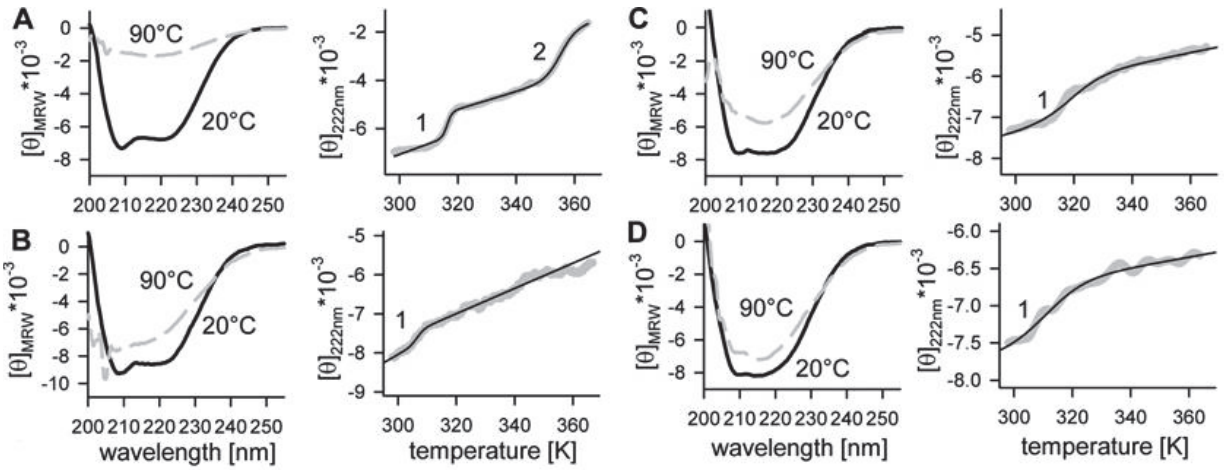


Figure 2.4 Thermal unfolding of the native and variant tRNA-NT
The CD spectra at 20 and 90^{o}C of (A) native tRNA-NT, (B) Mat-Mu1 variant, (C) Mat-Mu2 variant and (D) Mu1/Mu2 variant. The right-hand panels show secondary structure content in relation to the temperature by measuring the CD signal at 222 nm for each protein shown in the left-hand panel. The black lines represent the least-square fit result for a thermal transition using a two-state equation as described in Zoldák *et al.* (2004). For native enzyme, the behaviour was described with two coupled two-state equations. The results are listed in Table 2.1.

For the variants we observed a two-state transition only (Fig. 2.4 B-D). The transition temperature found is in the range of the first transition observed for the native protein (Table 2.1). Remarkably, the transition temperature for the Mat-Mu1 variant is shifted by about minus 10 K and for the Mu1/Mu2 variant by about minus 4 K. In contrast, for the Mat-Mu2 variant the transition temperature is slightly increased (plus 3 K). However, for all variants the energy required for the first unfolding is lower than that of the native protein and as expected the unfolding process requires less energy than observed for the complete unfolding of the native enzyme (Table 2.1).

2.5.4 The cellular distribution of native and variant tRNA nucleotidyltransferases

To see if the differences in stability affected protein distribution, we assessed the intracellular localization of C-terminal-GFP fusions of proteins containing all three methionines (M(1,2,3)), only the

first and the third methionine (M(1,3)) or the second and third methionine (M(2,3)). As a control, we produced the mature protein with only the third methionine (M(3)). We transformed these constructs into tomato protoplasts as described in Materials and Methods. Distribution was quantified in at least three independent experiments for each protein. Chloroplast localization was identified by overlay with chlorophyll fluorescence, mitochondrial localization was confirmed by Mitotracker™ staining and nuclear localization was confirmed by DAPI staining (exemplified in Fig. 2.5). When the GFP fluorescence did not overlay discretely with any of the known cellular structures the localization was annotated as cytosolic. With this system we can determine the distribution of the proteins *in vivo* and express the frequency of occurrence for a certain localization for each fusion protein in percentages of protoplasts in which this case was observed (Table 2.3, Schmidt von Braun *et al.*, 2007). To ensure that we were studying the localization of the intact fusion proteins, protein from each batch of protoplasts was assayed with anti-GFP antibodies in order to prove protein stability and visualize potential breakdown products of the GFP fusion proteins (Fig. S2.8).

When the protein lacking the first two methionines (M(3)) was used, the majority of the GFP fluorescence was found in the nucleus co-localizing with the DAPI fluorescence (Fig. 2.5, top row). The M(2,3)-tRNA-NT-Mat-Mu1-GFP and M(1,2,3)-tRNA-NT-Mat-Mu1-GFP constructs showed co-localization with chlorophyll fluorescence indicating that these proteins were directed to plastids (2nd and 3rd rows, respectively). In addition, the M(1,2,3)-tRNA-NT-Mat-Mu1-GFP also is targeted to mitochondria as confirmed by Mitotracker™ staining (3rd row). The M(2,3)-tRNA-NT-Mat-Mu2-GFP constructs showed a clear localization in the cytosol, but also a co-localization with chlorophyll fluorescence indicating that these proteins were directed to plastids (bottom row).

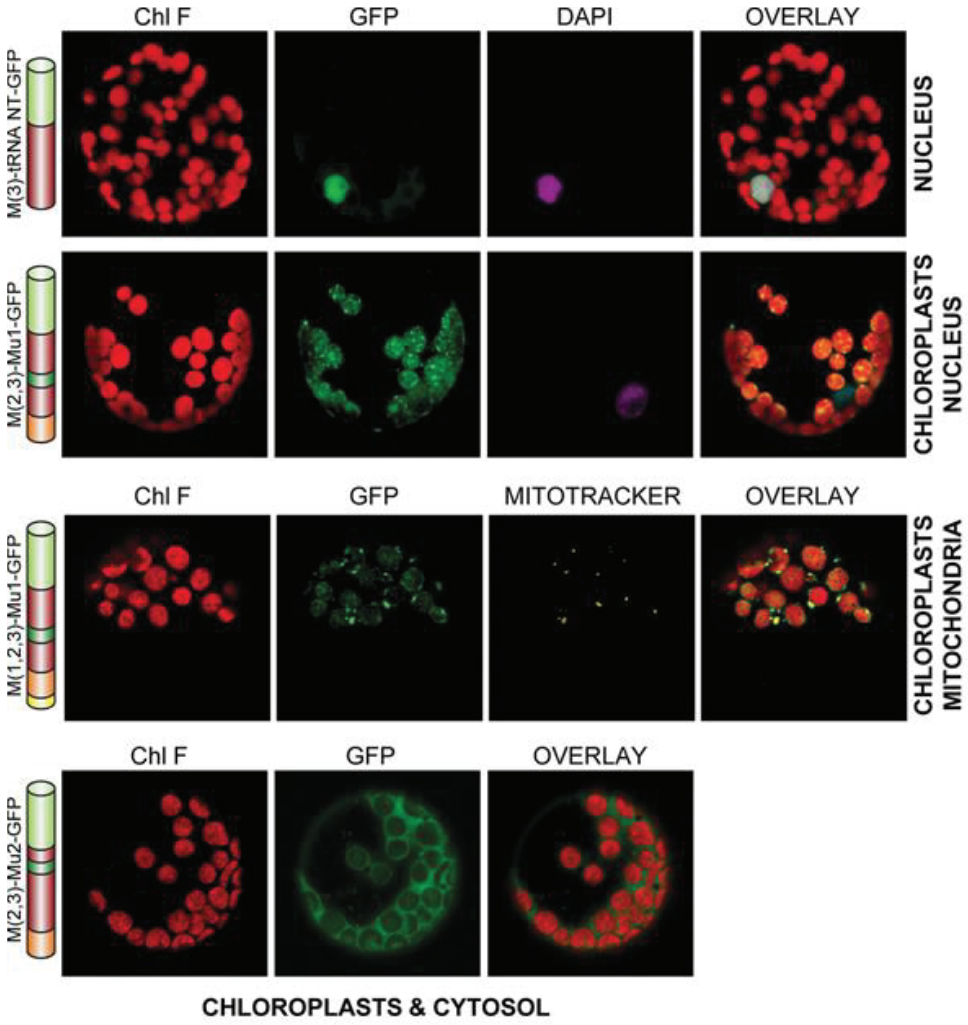


Figure 2.5 *In vivo* localization of variants of tRNA-NT

Different constructs were generated for expression in tomato protoplasts (Table 2.3). Shown are examples to document the different localizations of the GFP fluorescence and the confirmation of the compartments targeted. Shown is the fluorescence of chlorophyll (Chl F), the GFP signal (GFP), the DAPI staining to visualize the nucleus (DAPI), the Mitotracker™ staining to visualize mitochondria (MITOTRACKER) and the overlay of all signals for protoplasts transformed with M(3)-tRNA-NT–GFP, M(2,3)-tRNA-NT-Mat-Mu1–GFP, M(1,2,3)-tRNA-NT-Mat-Mu1–GFP and M(2,3)-tRNA-NT-Mat-Mu2–GFP. The interpretation as scored in Table 2.3 is given on the right. The expression control is shown in Supplementary Figure S2.8.

Figure 2.3: Localization of tRNA-NT in protoplasts

The numbers give the percentage of protoplasts where a given localization was observed. The values have been calculated from at least 80 different protoplasts from at least three independent experiments. The quantification was performed before assigning the results to the samples to avoid any bias.

Start	Variant	Chloroplast (%)	Mitochondrion (%)	Nucleus (%)	Cytosol (%)
M(1,2,3)	Native	31	100	3	2
M(1,3)	Native	16	100	11	4
M(2,3)	Native	85	91	9	1
M(1,2,3)	Mat-Mu1	86	45	7	11
M(1,3)	Mat-Mu1	100	21	5	6
M(2,3)	Mat-Mu1	100	39	31	19
M(1,2,3)	Mat-Mu2	86	15	9	9
M(1,3)	Mat-Mu2	100	8	22	17
M(2,3)	Mat-Mu2	55	3	4	100

Any native protein containing its amino-terminal extension, regardless of whether the first, the second or both the first and second methionines are present, showed a GFP signal consistent with mitochondrial localization in almost every case (Table 2.3). We also noted, that consistent with earlier results in non-green tissue (Schmidt von Braun *et al.*, 2007), the number of protoplasts showing chloroplast localization increased when methionine one was not present in the construct. Furthermore, for the native protein containing its amino-terminal extension we did not observe a significant population of protoplasts with nuclear or cytosolic GFP fluorescence.

Analysis of either variant of the mature domain revealed a shift towards chloroplast localization such that most of the protoplasts showed GFP fluorescence co-localizing with chlorophyll fluorescence. In addition, we noted that few protoplasts showed GFP fluorescence co-localizing with Mitotracker™ fluorescence. This again parallels with the results for Mat-Mu1 in non-green tissue (Schmidt von Braun *et al.*, 2007). In contrast to the earlier results for M(2,3)-tRNA-Mat-Mu1-GFP, we saw a significant portion of the GFP fluorescence in the nucleus (Fig. 2.5, Table 2.3). When the mature domain contained the Mat-Mu2 mutation, we observed chloroplast localization for each construct, but almost no mitochondrial or nuclear localization (Table 2.3). In contrast, when the first methionine was missing we observed a cytosolic distribution of the protein in all cells with a small number of protoplasts with GFP localization in chloroplasts.

2.5.5 Guidance of different marker proteins by the signal of the tRNA nucleotidyltransferase

The difference in the distribution of the precursor proteins resulting from the changes in the mature domain might suggest that the stability of the precursor protein modifies its intracellular localization. To test this we used two defined cargo proteins, namely GFP and Titin-I27 (Titin hereafter). Titin has been shown to be folded in the context of a precursor protein and modifies the transport efficiency of short signals, while GFP does not affect protein translocation (Ruprecht *et al.*, 2010, Bionda *et al.*, 2010). When three different amino-terminal targeting signals – M(1), M(1,2), or M(2) - were fused to GFP, we always observed a co-localization of GFP with chloroplasts (Table 2.4), which is consistent with the earlier results (Schmidt von Braun *et al.*, 2007). However, we also found for all constructs a large portion of protoplasts with mitochondrial localization of GFP (Table 2.4), which was not seen for M(2)-GFP in onion bulb scales (Schmidt von Braun *et al.*, 2007). Similarly, we observed a localization of M(1)-GFP and M(2)-GFP to the nucleus (30% of all protoplasts) and a cytosolic localization of M(2)-GFP (45% of all protoplasts), which was not observed before (Schmidt von Braun *et al.*, 2007).

Table 2.4: tRNA-NT signal targeting in protoplasts

The numbers give the percentage of protoplasts where a given localization was observed. The values have been calculated from at least 80 different protoplasts from at least three independent experiments. The quantification was performed before assigning the results to the samples to avoid any bias.

Start	Marker protein	Chloroplast (%)	Mitochondrion (%)	Nucleus (%)	Cytosol (%)
M(1,2)	GFP	98	88	2	5
M(1)	GFP	95	100	35	8
M(2)	GFP	77	49	26	45
M(1,2)	Titin–GFP	100	5	3	8
M(1)	Titin–GFP	100	25	65	18
M(2)	Titin–GFP	100	5	35	100

The import of the precursor proteins with Titin-GFP instead of the mature domain (Fig. 2.6, Table 2.5) revealed a remarkable difference to the results obtained with GFP alone as cargo protein. For all constructs, a chloroplast, but not a mitochondrial localization of the GFP fluorescence signal was observed. Only for the M(1)-Titin-GFP construct did we observe any appreciable (25% of all protoplasts) mitochondrial localization. Remarkably, the number of protoplasts showing nuclear localization for the

M(1) or the M(2) guided artificial precursor was increased and the cytosolic localization in the case of the M(2)-containing precursor was greatly increased.

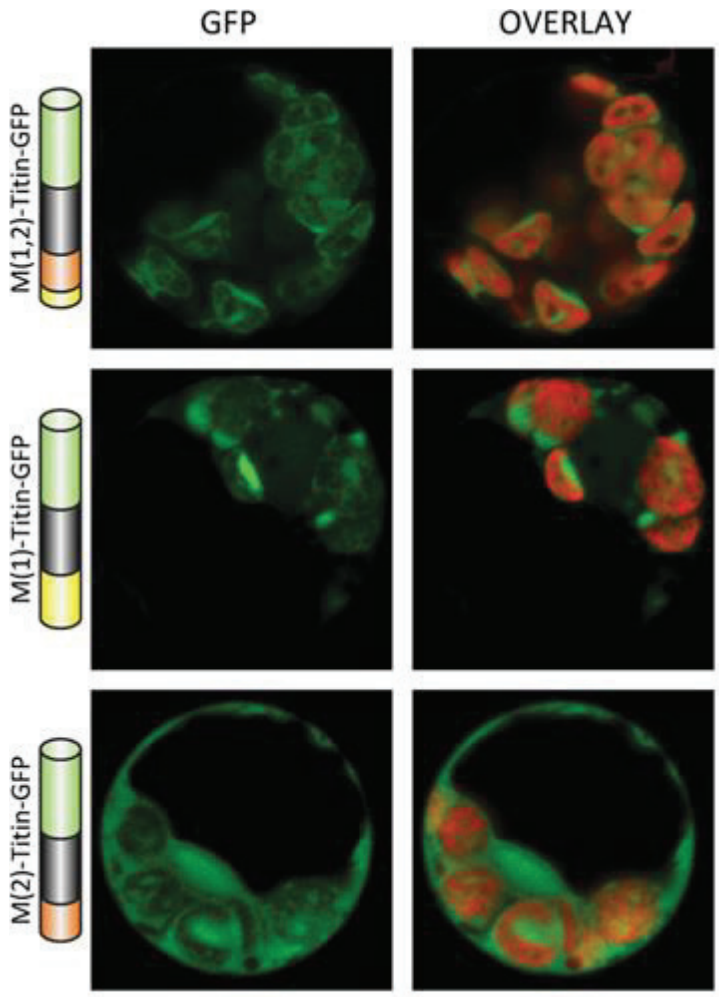


Figure 2.6 *In vivo* localization of Titin–GFP defined by the tRNA-NT signal
Different constructs were generated for expression in tomato protoplasts (Table 2.4). Shown are examples of the localization of the GFP fluorescence (left) and the overlay with the fluorescence of the chlorophyll (right) of protoplasts transformed with M(1,2)–Titin–GFP, M(1)–Titin–GFP and M(2)–Titin–GFP. The expression control is shown in Supplementary Figure S2.8.

Table 2.5: Number of tRNA-NT paralogues in individual species in eukaryotic phyla

Phyla	One paralogue	Two paralogues	Genomes
Fungi	73	5	78
Metazoa	43	1	44
Amoebozoa	–	1	1
Viridiplantae	17	1	18
Stramenopiles	–	2	2
Alveolata	10	–	10
Diplomonadia	3	–	3
Euglenozoa	7	–	7

2.6 Discussion

2.6.1 Functional relevance of the BODY- and TAIL-domain

The sequence of the head and neck region of class II tRNA-NTs shows high sequence similarity and five conserved motifs (Fig. 2.1, Fig. S2.2), which contain catalytically important residues. In contrast, the body and tail regions show limited sequence similarity, but the available crystal structures suggest a conserved fold of α-helical secondary structure (Betat *et al.*, 2010). The modifications introduced into the body (Mat-Mu1) and tail (Mat-Mu2) domains are well separated from the described catalytic residues. While the body domain interacts with the TΨC loop and acceptor stems of tRNAs (Tomita *et al.*, 2004), the tail domain is involved in interacting with the TΨC loop of tRNAs (Tomita *et al.*, 2004, Cho *et al.*, 2006). Because tRNA-NTs recognize diverse tRNAs these interactions are primarily through the sugar-phosphate backbone of the tRNA rather than through base-specific interactions (Tomita *et al.*, 2004, Cho *et al.*, 2006). Once a tRNA has been bound at the active site, these interactions may function as an anchor and prevent the RNA from dislodging from the enzyme surface during CCA addition. We were drawn to the Mat-Mu1 and Mat-Mu2 sequences for mutagenesis as both sequence stretches are rich in positively charged amino acids, (i) a feature of many nuclear targeting signals (Christophe *et al.*, 2000) and (ii) likely targets for interaction with the sugar phosphate backbone of the tRNA during tRNA binding (Tomita *et al.*, 2004). By converting the positively charged residues in these regions of the

protein to negatively charged residues we hoped to abolish any electrostatic interactions important to the function of these regions of the protein while maintaining the hydrogen bonding potential which would allow these regions of the protein to remain surface exposed.

Indeed, the two mutations altered the tRNA binding behavior. The Mat-Mu1 and the native protein were isolated from *E. coli* associated with RNA while Mat-Mu2 did not show an association suggesting a higher release rate for this variant (Fig. 2.2, Fig. S2.4). Further, RNA removal resulted in a massive CD-signal alteration for Mat-Mu1, but not for the native enzyme (Fig. 2.2). Indeed, based on the ratio of absorbance at 260 nm to absorbance at 280 nm (Table S2.2), the Mat-Mu1 variant and the native protein have approximately the same amount of tRNA associated. RNA removal resulted in a massive CD signal alteration for Mat-Mu1, but not for the native enzyme (Fig. 2.2). The CD signal of the Mat-Mu1 variant shows an approximately 3-fold increase of θ_{MRW} at 222 nm. This suggests that the interaction of the Mat-Mu1 variant with tRNA either alters the structure of the protein or follows a different mode of interaction or that the association of RNA molecules differs between the two proteins.

A different mode of interaction would be consistent both with the observation that the RNA molecule can be removed by dialysis from Mat-Mu1, but not from the native enzyme, and that there is a more dramatic alteration of the fluorescence of the native tRNA-NT by removal of the RNA molecules (Table 2.1). This latter observation supports the idea that fluorescence emission spectra generally reflect changes of the local environments of tryptophan and tyrosine residues in a protein (Lakowicz, 2006). It further suggests that the Mat-Mu1 mutations change the organization of the protein in regions where the tryptophan residues are found, such that RNA molecules are more exposed on the surface of the Mat-Mu1 variant. Mat-Mu1 is positioned in a predicted helix-loop region located in space close to the tRNA substrate (dark blue area in Fig. 2.1) and near the beginning of the body domain of the protein (Fig. 2.1A, Fig. S2.2) which is thought to be involved primarily in tRNA binding and anchoring (Tomita *et al.*, 2004, Cho *et al.*, 2006). Importantly, the portion of the *A. thaliana* tRNA-NT that is altered to

generate Mat-Mu1 appears to form an additional loop in the *A. thaliana* protein which is not present in any structurally-characterized tRNA-NT. Therefore, it is difficult to directly correlate changes in this region with any of the known structures, however, it is possible that the Mat-Mu1 changes could extend the predicted helix in this region toward the tRNA substrate either changing the orientation of regions of the protein responsible for tRNA binding or excluding tRNA from the active site. Either of these possibilities could result in a change of the position of the bound tRNA such that it is not tightly bound at the active site and remains accessible to be dialyzed away. This latter observation is consistent with the loss of function of the Mat-Mu1 variant both *in vivo* and *in vitro* (Table 2.2, Fig. 2.3, Fig. S2.7). That tRNA can still be bound to the Mat-Mu1 variant at the body and tail domains is supported by the copurification of tRNA with the Mat-Mu1 variant and with the similar dissociation constants of the native and Mat-Mu1 variant proteins (Table 2.2).

Unlike the Mat-Mu1 variant, the Mat-Mu2 variant shows a CD-spectrum similar to native tRNA-NT and a more reduced fluorescence emission maximum compared to native enzyme after RNaseA treatment (Fig. 2.2, Fig. S2.5B, Table 2.1). This suggests that Mat-Mu2 also changes the structural composition of the protein, but in a different way than Mat-Mu1. Given the low degree of sequence similarity in this region of the protein and the fact that in most available crystal structures *e.g*. (Augustin *et al.*, 2003) this region, although present in the analyzed construct, has not been solved it is difficult to precisely define a specific role for Mat-Mu2.

Even though there is no solved tail domain with tRNA complexed to the protein there is other experimental evidence that can be used to interpret our observations. For example, (i) *in vitro* studies have shown that the specificities of bacterial poly(A) polymerase and tRNA-NT can be altered by switching their C-terminal regions (Betat *et al.*, 2010). (ii) The C-terminal sequences can block entry of tRNAs lacking a CC sequence into an enzyme that catalyzes only A-addition (Tretbar *et al.*, 2011). (iii) The *Drosophila melanogaster* tRNA-NT lacking its last 124 amino acids still retains some activity (Yang, 2008)

suggesting that the tail portion of the protein carries out its functions without interacting directly with another region of the protein. (iv) Here, we show that altering the C-terminus of tRNA-NT by adding GFP does not inactivate the enzyme *in vivo* (Table 2.2, Fig. S2.7). Taken together these data suggest that the C-terminal region appears to define specificity, but plays no direct role in catalysis. However, our results from introducing mutations to modify or delete the C-terminal nine amino acids show that modifying this region of the protein does reduce enzyme activity to some degree (Fig. 2.3, Table 2.1). This indicates that the tail domain may play some role in enzyme function (Table 2.1). If we consider that these C-terminal nine amino acids are exposed on the surface of the protein they may represent one of perhaps many early interaction sites for tRNA. After binding to this site the tRNA then may be subsequently moved into the body region of the protein such that it is ultimately displayed appropriately at the active site. This hypothesis would explain why there are fewer tRNAs associated with the Mat-Mu2 or Mat-Δ2 variants during purification and would be consistent with the increased dissociation constant of Mat-Mu2 for tRNA (Table 2.1, Fig. S2.6, Table S2.2). Replacing a positively charged region of the protein with a negatively charged region would reduce the ability of that portion of the protein to bind to tRNAs. While it might be argued that the change in K_D is not biologically relevant as the K_M's for the native protein and the Mat-Mu2 variant remain the same (Table 2.1), which holds true for Mat-Δ2 as well (Table 2.1, Fig S2.6), it is interesting to note that the K_D values for the Mat-Mu2 and Mat-Δ2 variants differ by about 10-fold (Fig. S2.6, Table 2.1). This suggests that the mutation and deletion of this region of the tRNA-NT affects tRNA binding differently. In addition, while our *in vitro* (Fig. 2.3) and *in vivo* studies (Fig. S2.7, Table 2.2) suggest that this region of the protein can be modified such that the enzyme retains activity, they also indicate that altering the C-terminal nine amino acids does reduce activity (Table 2.1). So while this region may represent an initial interaction site it also appears to affect the enzymatic process.

The proposed structural variations are further consistent with the thermostabilities of the tRNA-NT variants (Fig. 2.4, Table 2.1). For the native enzyme we observed a two step unfolding at 43°C and 82°C, respectively, suggesting that the first event is an unfolding to a molten globule state with larger β-structure content (Greenfield and Fasman, 1969) while the second represents the complete unfolding of the protein. For the mutants, while the transition to the molten globule state was observed we never witnessed the fully unfolded state. The transition of the Mat-Mu1 variant to the molten globule state occurs at a lower temperature than seen for the native enzyme, while it was observed at a slightly higher temperature for the Mat-Mu2 variant. However, the alteration of the molten globule state characterized by helical content is consistent with the positioning of Mat-Mu1 and Mat-Mu2 in the C-terminal region of the protein with primarily α-helical structure (Fig. 2.1A).

Interestingly, the structural effects of the two mutations are not additive. On the one hand, the wavelength of fluorescence maximum is similar to the one observed for Mat-Mu2 variant (Table 2.1) and after treatment the Mat-Mu1/Mu2 variant shows a CD-spectrum comparable to that of the other variants (Fig. 2.2). On the other hand, the reduction of the fluorescence intensity of the Mat-Mu1/Mu2 variant by RNaseA treatment is comparable to that of the native enzyme. This restoration of fluorescence may represent a more “native” folding pattern in the presence of RNA. Furthermore, the Mat-Mu1/Mu2 variant showed an intermediate melting temperature for the molten globule state suggesting an additive affect from the individual changes. How the two mutations restore this fluorescence intensity in the presence of tRNA is unclear.

2.6.2 The import model for the tRNA nucleotidyltransferase

Here we analyzed the localization of tRNA-NT variants in a tomato leaf mesophyll protoplast system complementing the results presented in earlier studies using onion bulb scales (Schmidt von Braun *et al.*, 2007). Such complementary analysis is important because i) chloroplast-containing tissues

have the advantage that plastid localization of GFP fusion-proteins can be unambiguously shown employing the chlorophyll autoflourescence as a plastidic marker. This alleviates any problems associated with misleading interpretation of the localization due to problems in distinguishing between mitochondrial and proplastid localization in the onion bulb system (Jaedicke *et al.*, 2011). Furthermore, recent results indicate that dual localization of proteins can in some cases be tissue dependent (Kriechbaumer *et al.*, 2012). The *in vivo* localization experiments reported here further document that the cellular distribution of the tRNA nucleotidyltransferase is dependent on the N-terminal signal sequence as well as on the mature domain (Fig. 2.5, Schmidt von Braun *et al.*, 2007) regardless of the tissue type used. The first five amino acids of the N-terminal signal sequence appear to be required for efficient translocation into mitochondria since variants lacking the first start codon show less mitochondrial import (Table 2.3). This effect is again reflected by the fusions between the signal sequences and GFP or Titin-GFP (compare M(1,2) and M(1) to M(2) in Table 2.4). This observation is consistent with the data for the dual-targeted zinc-metalloprotease, Zn-MP, and the dual-targeted mitochondrial processing peptidase subunit alpha (MPP2), for which the domain at the very N-terminus of the signal sequence was shown to be required for mitochondrial targeting (Bhushan *et al.*, 2003). Thus, it appears to be a general scheme that the N-terminus of the signal of dual-targeted proteins is a determinant for mitochondrial localization.

Interestingly, either mutation in the mature domain reduces the localization to mitochondria and in turn enhances the translocation into chloroplasts (Table 2.3). Only the M(2)-tRNA-NT-Mat-Mu2-GFP construct showed a reduced organellar localization in general accompanied by a shift towards a cytosolic localization (Table 2.3). Correlating these results to the data from the thermal unfolding experiments suggests that the folding state of the mature protein has a regulatory function in the distribution of the tRNA nucleotidyltransferase. In contrast to the native domain, the variant domains unfold to the molten globule state such that the β-strand portion of the amino-terminus plays a more

significant role in maintaining protein structure. This suggests a connection between the stability of the N-terminal portion of the protein and the reduced mitochondrial import of these variants. Consistent with this notion, the signals fused to GFP are able to drive translocation into mitochondria, while the fusion with the more tightly folded Titin shifts distribution to chloroplasts (Table 2.4).

How can this observation be translated into a functional mode of preprotein translocation? The molten globule state has a significant enrichment in α-sheet content (Fig. 2.4) and β-strands are located in the N-terminal region of the tRNA-NT (Fig. 2.1). Thus, it is likely that the mutations at the C-terminus stabilize the N-terminal domain by an as yet unknown mechanism. This stabilization interferes with mitochondrial translocation, which is comparable to the impaired mitochondrial import of the dual targeted yeast fumarase after cytosolic folding, for which a co-translational mode was concluded (Sass *et al.*, 2003). Our results suggest a similar model in which tRNA nucleotidyltransferase is imported into mitochondria prior to folding either in a co-translational mechanism or post-translationally in association with targeting factors that keep the protein unfolded. Folding in the cytosol blocks mitochondrial import and redirects the protein to the nucleus or chloroplasts. Nuclear import is generally achieved in a folded state (Grunwald and Singer, 2012), while the translocon of the chloroplast was shown to provide a strong mechanical pulling force sufficient to translocate the tightly folded Titin domain I27 (Bionda *et al.*, 2010). The signal sequence of the tRNA nucleotidyltransferase is most likely exposed even in the correctly folded state of the mature domain (Fig. 2.1) which provides another hint that the protein can be imported into chloroplasts after cytosolic folding. This hypothesis would also explain the necessity of the very N-terminus of the signal sequence for efficient translocation into mitochondria. In this scheme, the mitochondrial targeting factors have to be recruited to the nascent precursor chain at an early state of translation in order to target the translating ribosome to mitochondria and/or to prevent folding in the cytosol.

This hypothesis is consistent with the mitochondrial import of the GFP marker protein and the rejection of the fast and tightly folding Titin-GFP domain (Fig. 2.6, Table 2.4). In contrast, import into chloroplasts was not affected by the nature of the mature domain (Table 2.4). This shows that the signal sequence alone is not sufficient for import into mitochondria and that the folding state of the domain following the signal drastically influences the localization of the protein. The signal for nuclear localization is conferred by characteristics of the mature domain since the fusion construct between GFP and the mature domain that lacked an N-terminal targeting signal was mainly found in the nucleus (Fig. 2.5, M(3)-tRNA-NT-GFP). Increased nuclear localization of the fusion constructs between the signal sequences and GFP or Titin-GFP (Table 2.4) is most likely the result of passive diffusion into the nucleus as the size of both of these fusion proteins is below the selection cut off for nuclear entry. Even strictly plastidic transit peptides fused to Titin-GFP are found in the nucleus when chloroplast import is impaired (Schmidt von Braun *et al.*, 2007, Grebenok *et al.*, 1997).

2. 7 Conclusions

We have shown that Mat-Mu1 and Mat-Mu2 alter the structure and activity of Arabidopsis tRNA-NT although in different ways. Mat-Mu1 eliminates enzyme activity and alters how the tRNA is associated with the protein. In contrast, Mat-Mu2 retains enzyme activity and tRNA binding. In addition, our results suggest that the translocation mode between mitochondria and chloroplasts might be different. Here, the translocation into mitochondria requires largely unfolded proteins, while the preprotein can be folded before translocation into chloroplasts. This mechanism of dual targeting of tRNA nucleotidyltransferase appears to be conserved in plants because for all of the plant genomes analyzed only one gene coding for tRNA nucleotidyltransferase was identified (Table 2.5). The second isoform found in *Arabidopsis lyrata* (gi|297813035) is lacking a major portion of its C-terminal region and thus most likely does not represent a functional equivalent of the tRNA-NT. The concept proposed here for translocation might even hold true for most of the other phylogenetic clades as multiple genes

are only found in stramenopiles containing complex plastids. These plastids have a translocation path remarkably different from those of normal plastids (Bolte *et al.*, 2009). Thus, we propose that tRNA nucleotidyltransferase, like fumarase, may represent a protein which requires coupling between translation and translocation for efficient import into mitochondria.

2.8 Supplementary materials

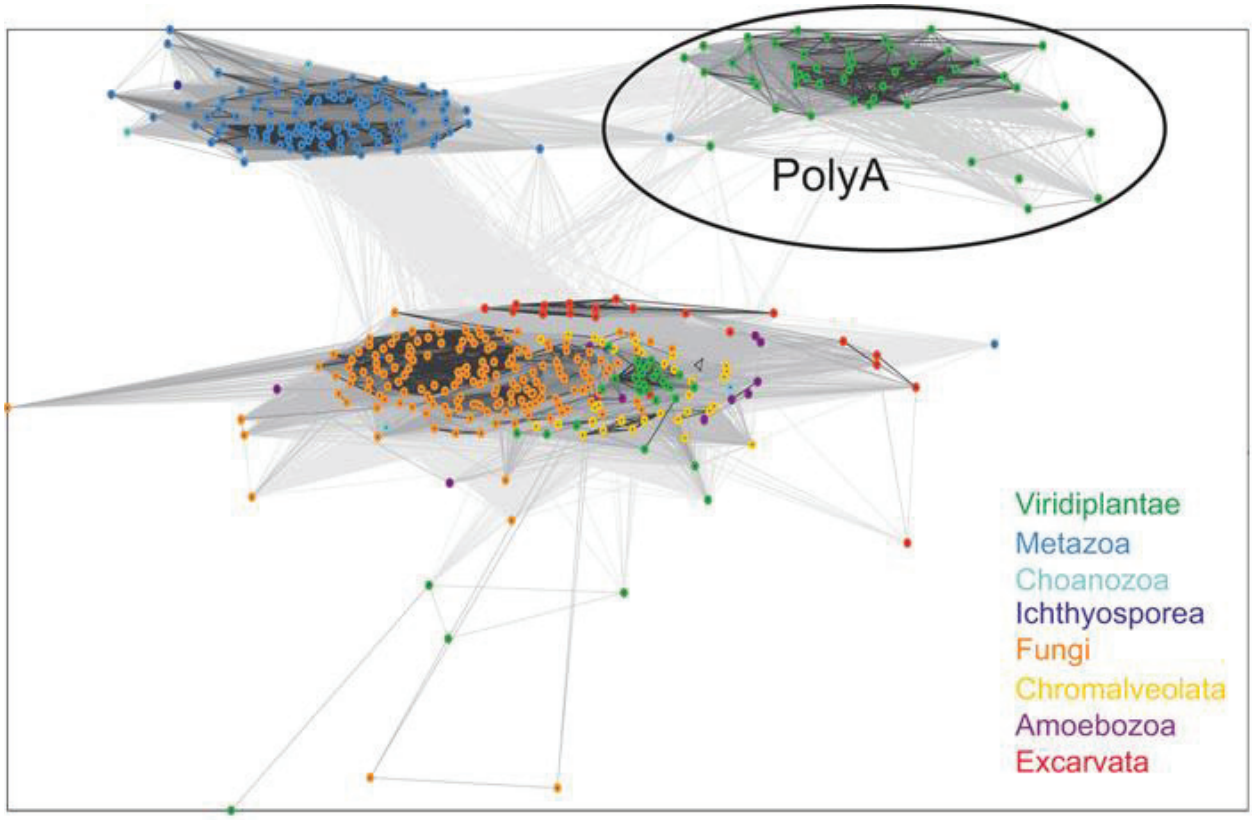


Figure S2.1 CLANS analysis of the identified tRNA-NT
The data to date have suggested that in fungi, mammals or plants there is a single gene coding for tRNA-NT (Chen *et al.*, 1992, Reichert *et al.*, 2001, Shanmugam *et al.*, 1996). To confirm this notion, the massive data set of available genome sequences was examined. A database with all eukaryotic class II tRNA-NTs was generated from the nr-database, which includes tRNA-NTs and poly(A) polymerases from bacteria (Yue *et al.*, 1996). In addition to existing pHMMs from the PFAM database, pHMMs based on alignments of the neck and the body/tail domains were created to enhance the sensitivity for eukaryotic sequences. Using these pHMMs 435 eukaryotic sequences, which match at least two profiles of different domains, were detected. With a clustering based on sequence similarity, 49 sequences, which belong to one cluster and share a high-sequence similarity with an experimentally proven poly(A) polymerase of *A. thaliana* (Zimmer *et al.*, 2009), were excluded from the dataset, leaving 386 eukaryotic tRNA-NTs, which are shown as a CLANS representation. Subsequently, all sequences obtained from organisms whose entire genome sequences are not deposited in the NCBI genome database were removed to be certain that all paralogues of a species were included. In most species of Viridiplantae, metazoa and fungi only one paralogue was found (Table 2.5 of the main text). For exceptional cases, more detailed analysis was performed. Remarkably, among fungi five different organisms have two paralogues, namely *Schizosaccharomyces japonicas*, *Schizosaccharomyces pombe*, *Penicillium marneffei*, *Fusarium oxysporum* and *Aspergillus nidulans*. The same is true for the Amoebozoa *Dictyostelium discoideum* and the Metazoa *Heterocephalus glaber* . Interestingly, both stramenopiles *Phaeodactylum tricornutum* and *Thalassiosira pseudonana* , contain two different predicted tRNA-NTs. However, to the best of our knowledge no activity has been demonstrated for any of these paralogues.

Table S2.1: Oligonucleotides used in the present study

Primer	Primer sequence (5′→3′)
1	CTGCAGCTGATCAGCCTCTATCTCTTGCTCTTCAGCTTCGATATCTCTCATCCACTC
2	GAGCTCGAGTGATCAATGATACTAAAAACCATG
3	GAGCTCGAGTGATCAATGACGAATGTTGGAGAGG
4	GAATTCGTCGACTCACTCTATCCTTTGTCG
5	GAATTCGTCGACTCACTCTATCTCTTGCTCTTCAGCTTCGATATCTCTCATCCACTC
6	CATATGTCGACTCAATCTCTCATCCACTCCTTG

Table S2.2: Absorption ratio (A_{280}/A_{260})

WT, wild-type.

Construct	Before RNAse treatment	After RNAse treatment
WT	1.50 ± 0.20	0.66 ± 0.03
Mu1	1.50 ± 0.30	0.88 ± 0.04
Mu2	0.91 ± 0.04	0.81 ± 0.07
Δ2	1.12 ± 0.09	0.88 ± 0.05

Table S3.3: Results of three different fold recognition servers

For the first five hits the following details are given: PDB code of the structure, the organism the enzyme was originated from, the rank and the corresponding scores.

PDB code	Organism	Rank in			E-value HHpred	Score ffas	Confidence Phyre2 (%)
		HHpred [12]	ffas [23]	Phyre2 [24]			
3H37/3H38 [18]	*Thermotoga maritima*	1	1	2	1.2×10^{-67}	−77.2	100
1MIV/1MIW/1MIY [19]	*Geobacillus stearothermophilus*	2	2	1	1.0×10^{-62}	−75.8	100
1OU5 [22]	*Homo sapiens*	4	4	5	1.1×10^{-64}	−72.3	100
1VFG [20]	*Aquifex aeolicus*	3	5	3	5.6×10^{-62}	−71.3	100
3AQI/3AQN	*Escherichia coli* DH1	5	3	4	1.7×10^{-54}	−73.6	100

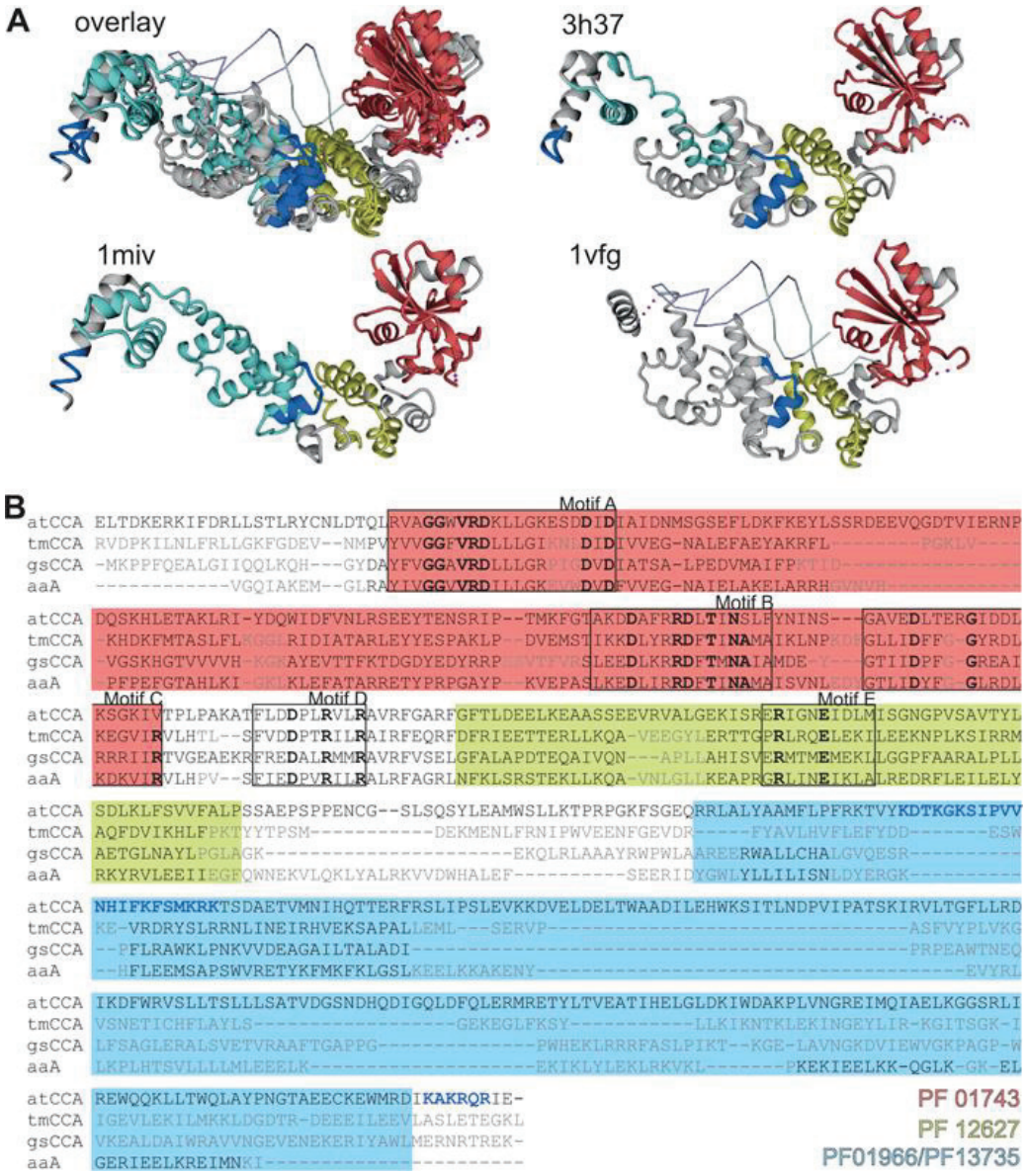


Figure S2.2 Secondary structure prediction of the Arabidopsis tRNA-NT
(A) The structures of *Thermotoga maritime* [18], *Geobacillus stearothermophilus* (Li *et al.*, 2002) and *Aquifex aeolicus* (Tomita *et al.*, 2004) are shown and the sequence parts which align to the Mat-Mu1 and Mat-Mu2 sequence of *A. thaliana* are indicated (dark blue). The regions corresponding to the Pfam pHMMs are shown in colours [red, PolyA_pol (PF01743); yellow, polyA_polRNAbd (PF12627); and blue, RNA_NucTran2_2 (PF01966)/HD (PF13735)]. Additionally the structural alignment generated with Sheba (Jung and Lee, 2000) is shown (upper right), visualizing that the conformational space of the body and tail domain is larger in the crystal than for the head and neck domain. The three structures were chosen based on the fold recognition results presented in Table S2.3. (B) Pairwise alignment generated by the HHpred server for atCCA and the corresponding hits. Positions which do not occur in our MSA are coloured grey, and positions supported by our MSA are coloured black. Motifs A–E (Toh *et al.*, 2009) are boxed and labelled; absolutely conserved residues are printed in bold and the colour code used is as in (A).

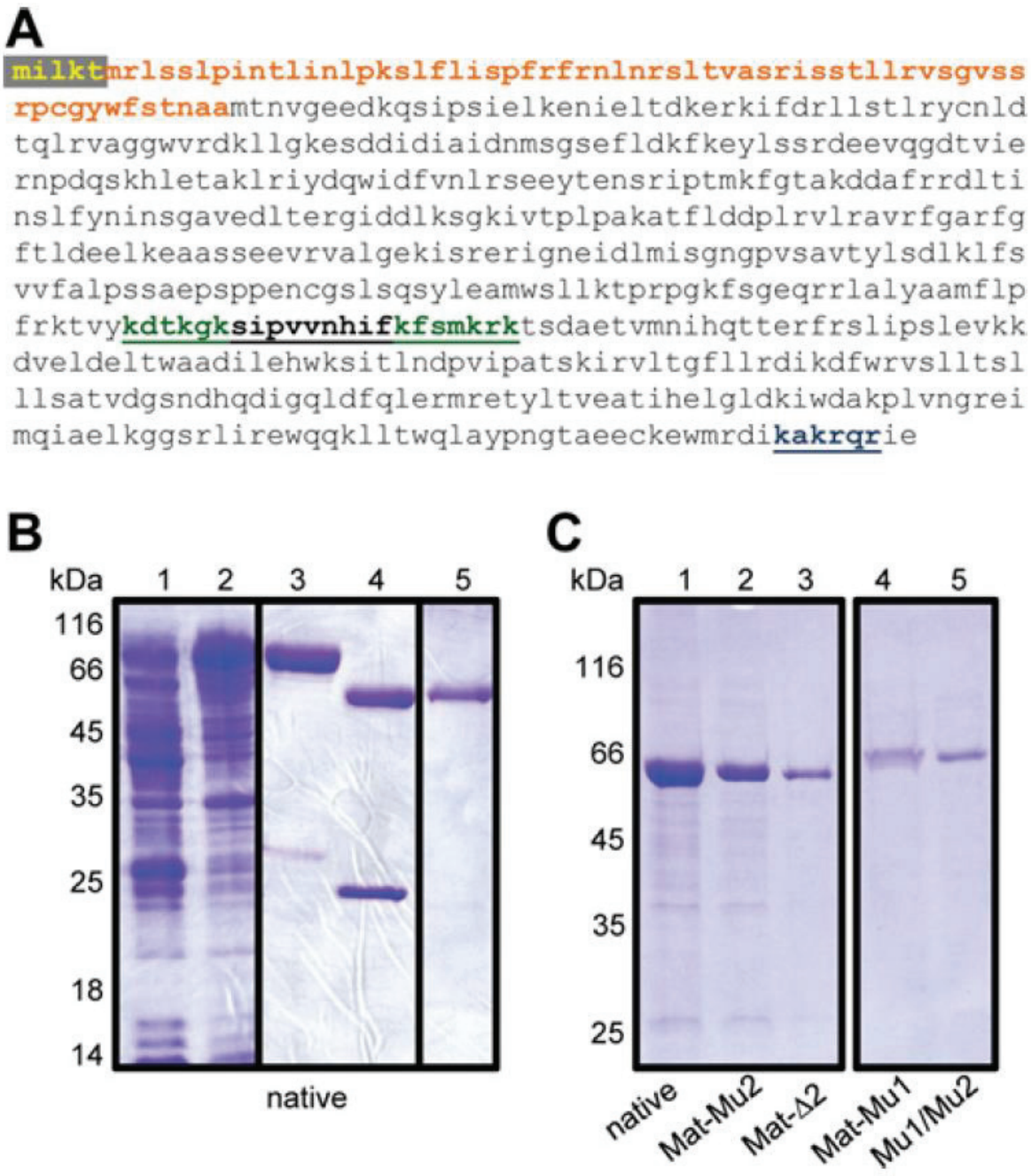


Figure S2.3 *Arabidopsis thaliana* tRNA-NT and its heterologous expression
(A) The amino acid sequence of the tRNA-NT from *A. thaliana* is given. (B) Coomassie-Blue-stained SDS/PAGE of supernatant (lane 1) and pellet (lane 2) after lysis of *E. coli* cells overexpressing tRNA-NT, as well as the GST column-purified protein before (lane 3) and after (lane 4) thrombin cleavage and repurification (lane 5). (C) Coomassie-Blue-stained SDS/PAGE of purified native or variant tRNA-NTs.

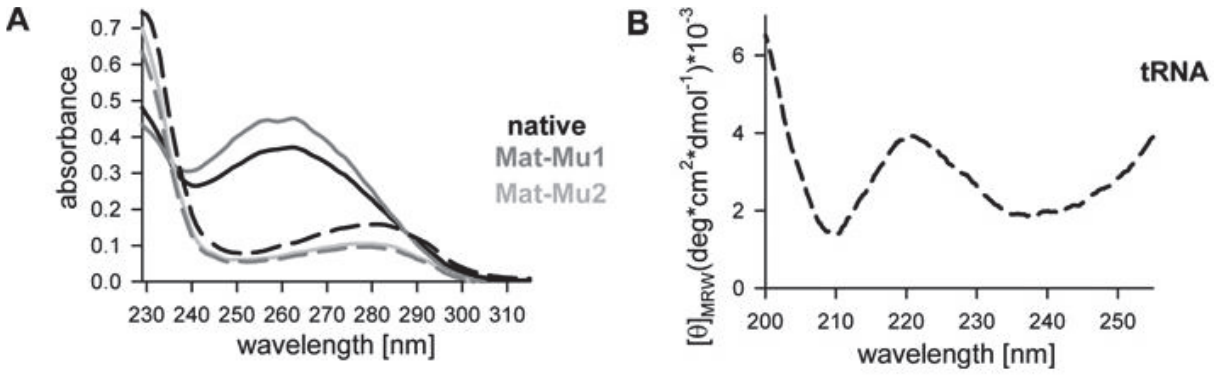


Figure S2.4 tRNA is co-purified with the tRNA-NT
(A) The absorption spectra of the proteins used in Figure 2.2 (A) of the main text. Spectra were recorded before (solid lines) and after (broken lines) RNaseA treatment. The untreated native sample was diluted 2-fold before measuring absorbance. (B) The CD signal of 1 μM tRNA is shown.

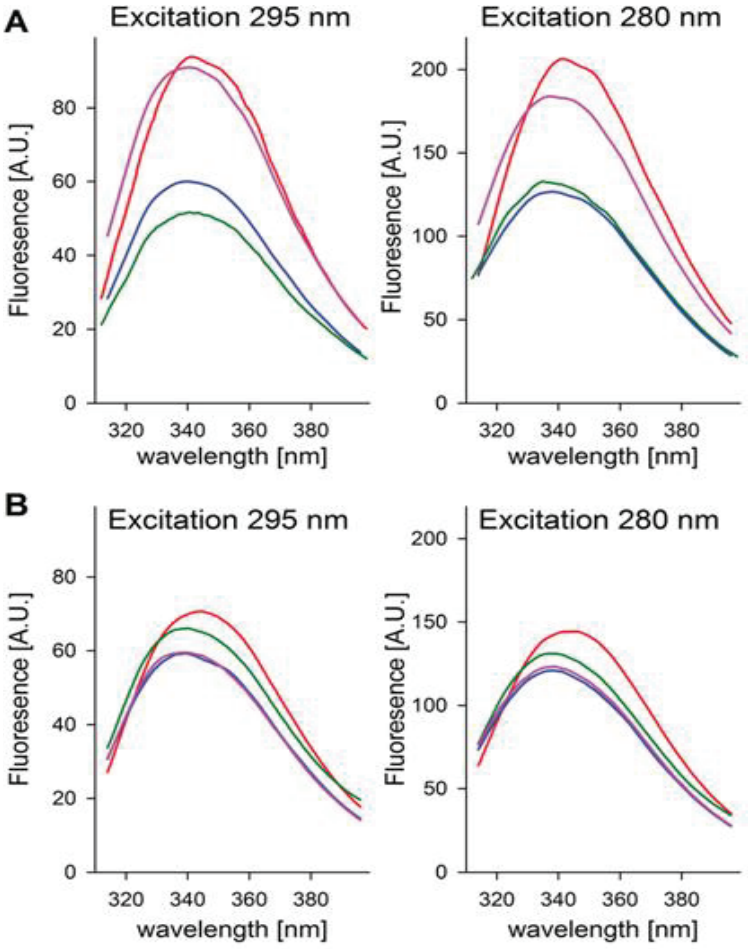


Figure S2.5 Fluorescence emission spectra for the native and variant tRNA-NTs
Proteins were excited at the wavelengths indicated and the fluorescence emission was measured from 310 to 400 nm. The lines represent native (red), Mat-Mu1 (green), Mat-Mu2 (blue) and Mu1/Mu2 (purple) samples. Samples are shown before (A) and after (B) tRNA removal.

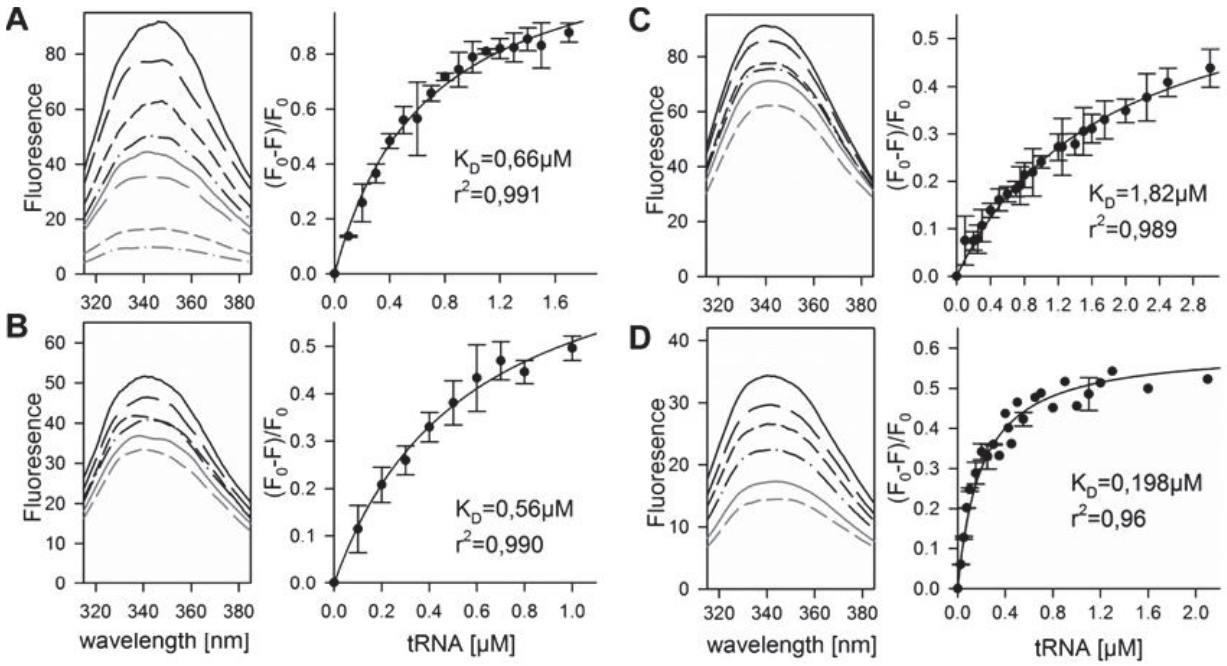


Figure S2.6 The binding of tRNA to native and variant tRNA-NTs

(A) The quenching of native fluorescence by tRNA (solid line) was analysed by the addition of 0.1, 0.2, 0.3, 0.4, 0.5, 1.0 and 1.7 μM tRNA (left). The fluorescence intensity at 346 nm was determined and (F 0 −F)/F 0 plotted against the tRNA concentration (right). The data were analysed by least-square fit analysis to (F 0 −F)/F 0 = MAX×[tRNA]/(K d +[tRNA]), and the determined K_D as well as the r^2 value of the fit is indicated. (B) The quenching of the Mat-Mu1 variant fluorescence (solid line) by tRNA was determined by the addition of 0.1, 0.2, 0.3, 0.4, 0.5 and 0.8 μM tRNA (left). The fluorescence intensity at 346 nm was determined and (F 0 −F)/F 0 plotted against the tRNA concentration (right). The data were analysed as in (A). (C) The quenching of the Mat-Mu2 variant fluorescence (solid line) by tRNA was determined by the addition of 0.1, 0.3, 0.5, 0.8 and 1.6 μM tRNA (left). The fluorescence intensity at 346 nm was determined and (F 0 −F)/F 0 plotted against the tRNA concentration (right). The data were analysed as in (A). (D) The quenching of the Mat-_2 variant fluorescence (solid line) by tRNA was determined by the addition of 0.05, 0.15, 0.35, 0.8 and 1.3 μMtRNA (left). The fluorescence intensity at 346 nm was determined and (F 0 −F)/F 0 plotted against the tRNA concentration (right). The data were analysed as in (A).

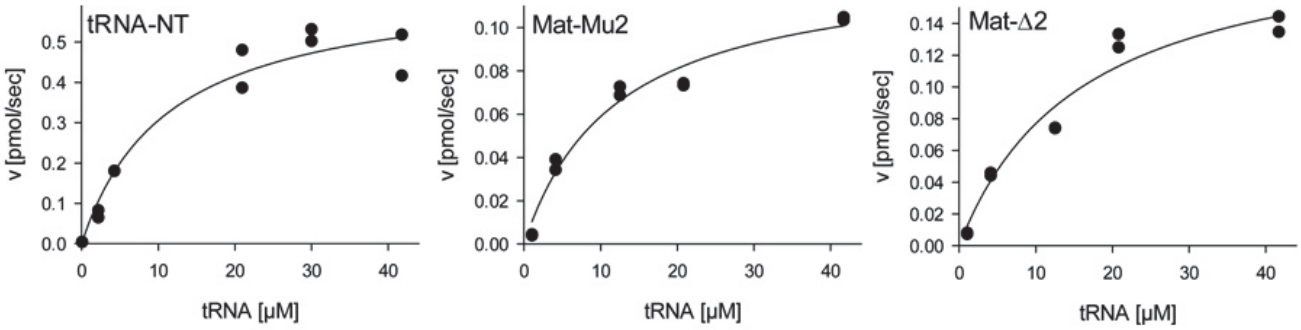


Figure S2.7 Enzyme kinetics of the native version and the mutants of the C-terminus
The rate of addition of adenosine monophosphate to a fixed amount of tRNA-NCC template (v [pmol/sec]) by the native tRNA-NT (top), Mat-Mu2 (middle) or Mat-_2 (bottom) variant was determined at different tRNA concentrations (tRNA [μM]; results shown as dots). The data were analysed by a least-squares method using the equation for Michaelis–Menten kinetics, v = V max [S]/(K m +[S]), with the results shown as the line.

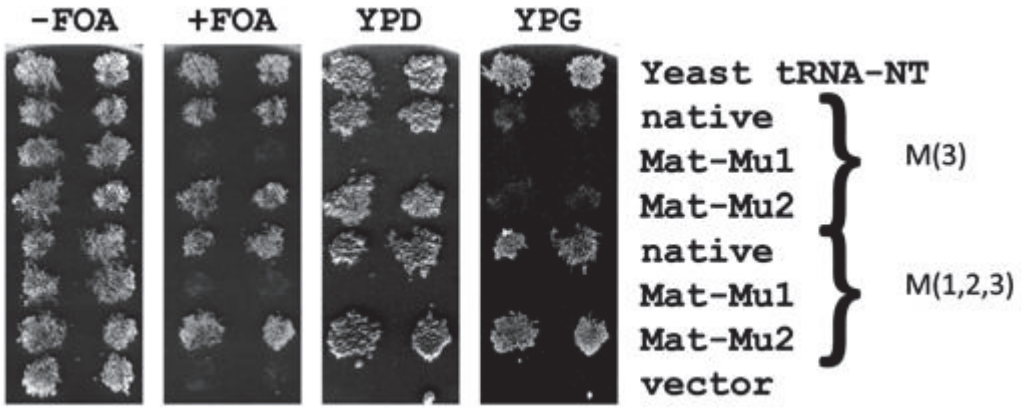


Figure S2.8 Ability of Arabidopsis tRNA-NT–GFP gene fusions to complement a *CCA1* null mutation in yeast
An *S. cerevisiae* strain bearing a null mutation in the chromosomal *CCA1* gene and plasmid-borne *CCA1* and *URA3* genes was transformed with the *HIS3* -bearing G131-2 plasmid alone or G131-2 derivatives expressing Arabidopsis tRNA-NT–GFP native or variant fusions under the control of the *TDH3* promoter. To accomplish loss of the *URA3-CCA1* bearing plasmid, two independent His+ Ura+ transformants were patched on SC-ura-his medium, replica-plated to YPD (glucose) medium, incubated for 24 h, and replica-plated to medium with or without FOA. The ability of FOA-resistant yeast to grow on different carbon sources was tested by subsequently replica-plating to YPD (glucose) or YPG (glycerol) medium.

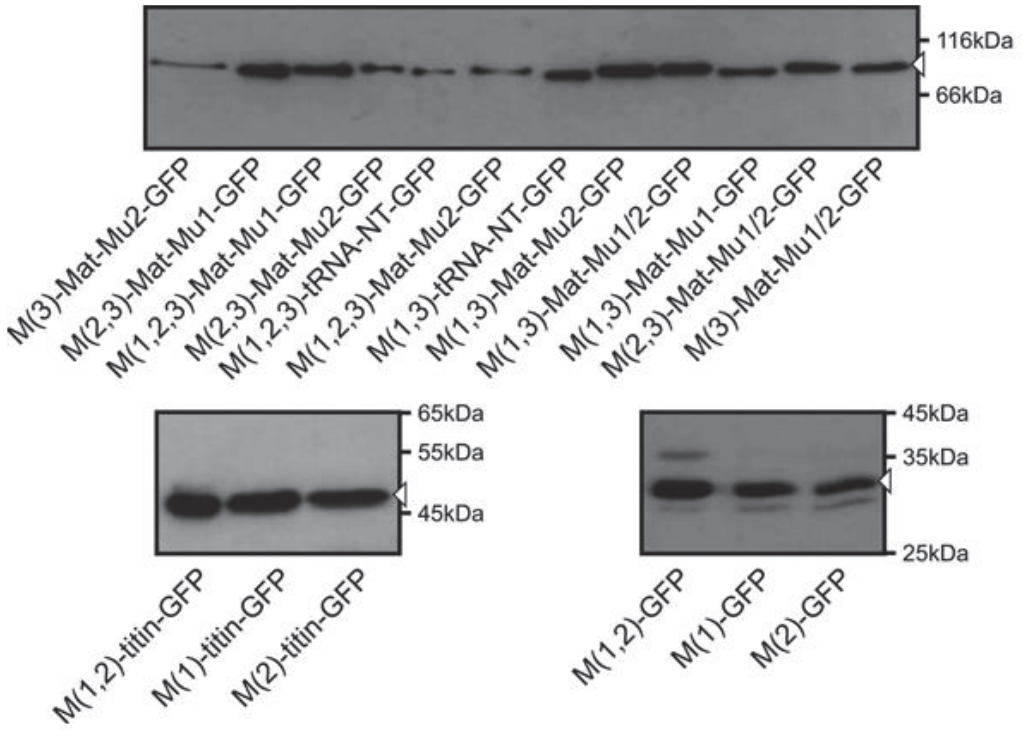


Figure S2.9 Expression of the tRNA-NT fusion constructs in protoplasts
Protoplasts were transformed with the indicated constructs as shown in Figures 5 and 6 of the main text. The protoplasts were harvested, subjected to SDS/PAGE and immunodecorated with antibodies against GFP. The white arrowhead shows the migration of the mature domain.

Chapter 3: Exploring the interplay of localization, activity, structure and stability in Arabidopsis tRNA nucleotidyltransferase function.

3.1 Preface

The work presented in Chapter 3 corresponds to the following manuscript in preparation: Leibovitch M., Hanic-Joyce P.J., and Joyce P.B.M. (2016). **Exploring the interplay of localization, activity, structure and stability in Arabidopsis tRNA nucleotidyltransferase function**. Target journal: Biochemical Journal. Paul Joyce and Matthew Leibovitch designed the experiments and drafted the manuscript. Matthew Leibovitch carried out the biophysical and biochemical characterization. Pamela Hanic-Joyce carried out the yeast complementation assays shown in Figure 3.3. Localization studies in plants will be performed in the lab of Enrico Schleiff, Goethe Universität, Frankfurt.

3.2 Abstract of manuscript

The function of any protein is defined by a combination of its structure, stability, activity and localization. Here, the Arabidopsis tRNA nucleotidyltransferase enzyme which functions in the cytosol, mitochondrion, nucleus and plastid is studied to explore the interplay of all of these factors. Previously, a bipartite motif (Mat-Mu1) was identified and when changed, it altered the activity, stability and intracellular distribution of this enzyme (Leibovitch *et al.*, 2013). Here, the combination of site-directed mutagenesis, spectroscopic techniques and enzyme activity assays were used to explore the roles of specific amino acids within this motif. The data show that keeping either half of this bipartite motif is sufficient to maintain enzyme structure and stability at near native levels. While one half of the motif seems to play a role in structure the other appears more important in maintaining enzyme activity. Specific amino acids playing a role in activity have been identified. With this more defined analysis of Mat-Mu1 in hand, the interplay between activity or structure and localization can be explored further.

3.3 Introduction

A cell is a complicated place. Growth and viability are defined at a number of levels ranging from the genes that any cell contains, to how those genes are expressed, to how the resulting transcriptional products (RNAs) are utilized and processed and ultimately to which proteins may be produced. However, the story does not end there. The specific function of any protein in a cell is defined by a combination of factors including its stability, structure, activity and localization (Whisstock and Lesk, 2003). For example, hemoglobin can perform its function as an oxygen transporter because it has a long half-life (Thom *et al.*, 2013). In contrast, other proteins such as cyclins play a role in defining the cell cycle because they exist only transiently (Darzynkiewicz *et al.*, 2015). Since the time of Anfinsen, it has been obvious that structure defines function and proteins may be activated or inactivated by changes in their structure, *e.g.*, proteolytic activation of proteases such as trypsin or chymotrypsin (Nemoda and Sahin-Toth, 2006). Furthermore, altered protein structure as exemplified by protein misfolding has been implicated in a number of diseases, *e.g.*, cystic fibrosis, Creutzfeldt-Jakob disease, Parkinson's disease, Huntington's disease, *etc*. (Ramirez-Alvarado, 2008). Moreover, proteins must be directed to their appropriate destinations to carry out their functions. For example, when mutations in the gene coding for α1-antitrypsin (which normally protects against the action of proteases) cause an alteration of its structure such that it is not released from cells, obstructive pulmonary disease and liver disease result (Wang *et al.*, 2014). Finally, the function of certain proteins may be defined by multiple activities and these proteins may be directed to specific destinations to carry out these different activities. For example, Mod5p in yeast functions to modify tRNAs in the cytosol but acts as a transcriptional regulator in the nucleus (Smaldino *et al.*, 2015) and enolase not only catalyzes one of the reactions of glycolysis in the cytosol but also is involved in the import of tRNA into yeast mitochondria (Entelis *et al.*, 2006).

Here, we study tRNA nucleotidyltransferase as a model protein to understand how structure, stability, activity and localization may interact in defining function. This enzyme post-transcriptionally

adds the 3'-terminal cytidine-cytidine-adenosine sequence (Deutscher and Evans, 1977) required for aminoacylation to tRNAs encoded in the nuclear, mitochondrial and plastid genomes (Abe *et al.*, 2014). In addition, it plays a repair function in the cytosol (Rosset and Monier, 1965), may be involved in the stress response targeting tRNAs for rapid decay (Wilusz *et al.*, 2011, Kuhn *et al.*, 2015), may serve as a quality-control checkpoint for nuclear export of tRNAs (Lizano *et al.*, 2008) and may be involved in nucleocytosolic shuttling of tRNAs (Feng and Hopper, 2002). A single gene coding for this enzyme exists in most eukaryotes (Leibovitch *et al.*, 2013) such that the product(s) of this gene must function in multiple cellular locations. For example, in Arabidopsis this protein functions in the nucleus, cytosol, mitochondria and plastids (Schmidt von Braun *et al.*, 2007).

A traditional understanding of protein localization relies on the interplay of targeting signals contained on the protein of interest, *trans*-acting factors (which may or may not be specific components of the transport machinery) that recognize these signals and a transport complex at the appropriate membrane. Targeting signals may be primarily amino-terminal, *e.g.*, endoplasmic reticulum, mitochondrion or plastid (Kunze and Berger, 2015), carboxy-terminal, *e.g.*, peroxisomal (Emmanouilidis *et al.*, 2015) or anywhere within the protein sequence, *e.g.*, nuclear targeting signals (Christie *et al.*, 2015). These targeting signals may define specific primary sequences, *e.g.*, stretches of basic amino acids (Kunze and Berger, 2015), secondary structures, *e.g.*, amino terminal α-helices(Dudek *et al.*, 2013), or higher order structures, *e.g.*, nuclear localization signals (Cokol *et al.*, 2000) that are recognized by the transport machinery. Proteins may be delivered directly to the membrane (co-translational transport) or may be targeted post-translationally (Kunze and Berger, 2015) with the assistance of chaperones, *e.g.*, Hsp70 (Deshaies *et al.*, 1988) or accessory factors, *e.g.*, mitochondrial import stimulatory factor (Hachiya *et al.*, 1993), or importins (Christie *et al.*, 2015). Post-translational modifications, *e.g.*, phosphorylation of a targeting signal (Bauer *et al.*, 2015), or proteolytic processing of a transit peptide or protein may alter targeting (Bauer *et al.*, 2015). Also, the presence or absence of a

targeting signal may be defined at the level of transcription or translation (Yogev *et al.*, 2011, Schmidt von Braun *et al.*, 2007). In addition, some proteins may contain more than one sorting signal such that there is a delicate balance to ensure that the appropriate amount of protein is targeting to the appropriate destination at the appropriate time (Kunze and Berger, 2015).

We previously showed that multiple transcripts are generated from the single Arabidopsis gene coding for tRNA nucleotidyltransferase indicative of protein isoforms beginning from methionines one, six, or 69 of the tRNA nucleotidyltransferase open reading frame (Schmidt von Braun *et al.*, 2007). We showed that the first 68 amino acids were not required for enzyme activity but were both necessary and sufficient for mitochondrial and plastid targeting. Proteins beginning from amino acid one were targeted more efficiently to mitochondria than were proteins beginning from amino acid six. Intriguingly, the presence of the mature domain of tRNA nucleotidyltransferase in addition to the amino-terminal targeting sequences shifted the localization of the reporter protein GFP toward mitochondria suggesting that some portion of the mature domain of tRNA nucleotidyltransferase also could affect targeting (Schmidt von Braun *et al.*, 2007). To address this possibility, mutations were made in the mature domain of the protein and shown to shift distribution from mitochondria to plastids (Schmidt von Braun *et al.*, 2007). To address the role of the mature domain in protein targeting and to define the role of the amino acids that were changed, we explored the enzymatic and physicochemical properties of tRNA nucleotidyltransferase variants to correlate the properties of the enzyme with the intracellular distribution of the protein. Our results suggested that the folding capacity of the mature domain may influence organelle selection and that mitochondrial targeting is impeded with a more tightly folded protein (Leibovitch *et al.*, 2013). These data suggest that tRNA nucleotidyltransferase is imported most efficiently into mitochondria prior to folding, either through a co-translational mechanism or post-translationally in association with targeting factors that keep the protein unfolded. The fact that there was no apparent change in secondary or tertiary structure as measured by circular dichroism and

fluorescence spectroscopy, respectively, suggests that the variant folds similarly to the native enzyme *in vitro*. In addition, the variant enzyme showed a reduced transition temperature (10°C lower than native) and a loss of activity below the limits of detection with our standard enzyme assay (Leibovitch *et al.*, 2013). This suggests that there may be a combination of structure, stability and activity involved in defining the targeting of this protein.

The purpose of this study is to explore in greater detail the role of the mature domain of Arabidopsis tRNA nucleotidyltransferase in protein targeting. This will be done by defining the role that specific amino acids (Fig. 3.1) may play in protein structure, stability and activity to dissect the contributions of these processes in localization. One could envision that amino acid substitutions would alter a region of primary sequence (such as a basic stretch of amino acids characteristic of a nuclear localization signal) that would disrupt interaction with a membrane receptor or accessory transport protein (Christie *et al.*, 2015). Altering these amino acids also could change the higher order structure of the protein such that it is folded into a less import competent configuration, or it is no longer able to interact with a receptor or accessory protein. Finally, one could envision that protein transport is mediated through an enzyme/substrate (*e.g.,*tRNA-NT/tRNA in this case) complex and that these amino acid changes disrupt that complex.

To address all of these possibilities, the structure, stability and activity of the native enzyme (399-***KDTKGK***SIPPVVNHIF***KFSKRK***-420), the Mat-Mu1 variant (399-***eDTeGe***SIPVVNHIF***eFSMeme***-420) (Leibovitch *et al*., 2013) and seven new variants: Mat-Mu1A (***eDTeGe***), Mat-Mu1B (***eFSMeme***), K414EK418E (***eFSMeRK***), K414E (***eFSMKRK***), K418E (***KFSMeRK***), K418Q (***KFSMqRK***) and K420E (***KFSMKRe***) will be analyzed.

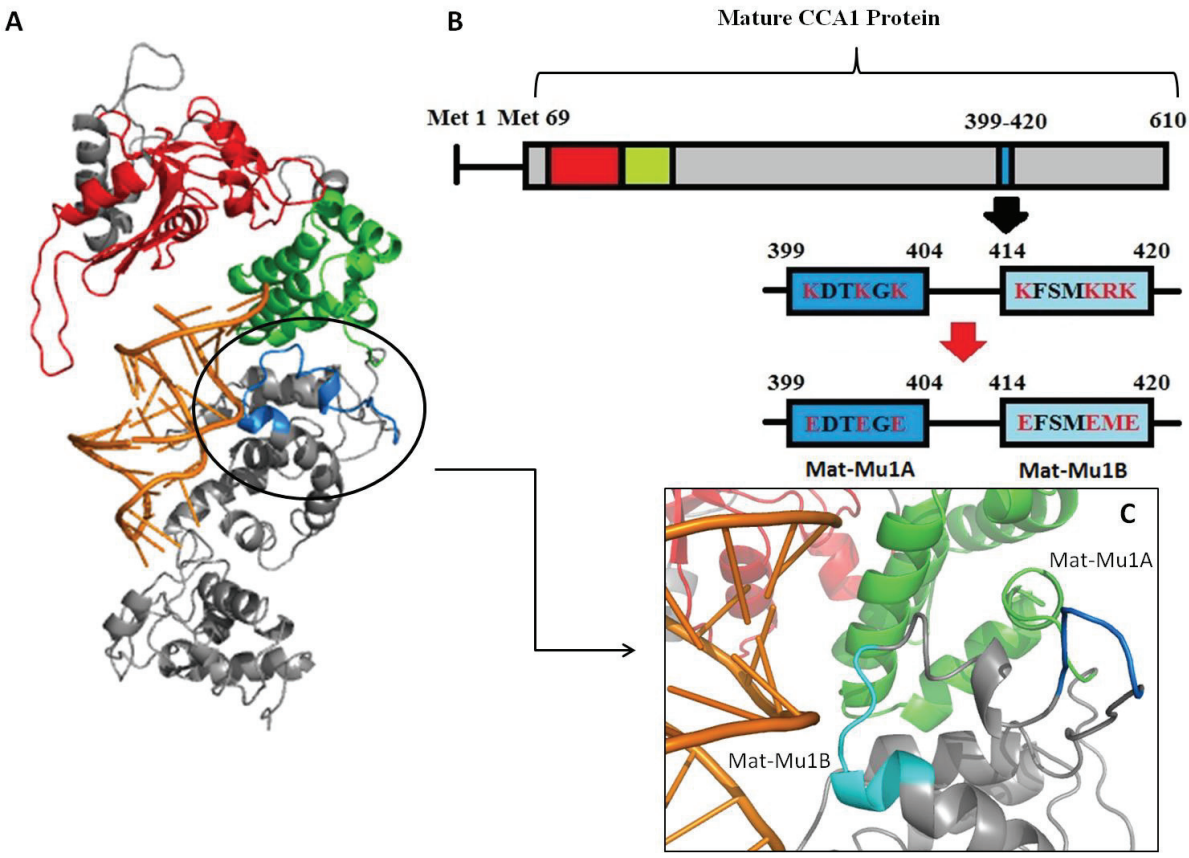


Figure 3.1: Model of *A. thaliana* tRNA nucleotidyltransferase
(A) The homology model of *Arabidopsis thaliana* tRNA nucleotidyltransferase is show in ribbon representation with a tRNA substrate (from *Aquifex aeolicus* (1VFG,(Tomita *et al.*, 2004)). The head (red) and neck (green) regions are conserved among Class II tRNA nucleotidyltransferases and contain conserved motifs A-E. Regions that are less well conserved are grey and include the carboxy-terminal portion of the enzyme which includes the body and tail. (B) The mature protein used in this study starts from the third in frame methionine (Met69) and the head and neck regions are shown in red and green, respectively. The region changed to produce Mat-Mu1, between amino acids 399 and 420, is depicted by the thin blue and the native sequence is shown below. After the amino acids have been changed, Mat-Mu1A (light blue) and Mat-Mu1B (dark blue) are shown. (C) The same colour scheme is used to show Mat-Mu1A and Mat-Mu1B (blue) on the homology model.

Neither Mat-Mu1A nor Mat-Mu1B alone show any effect on thermal stability, however, Mat-Mu1B does affect the catalytic efficiency of the enzyme. The K418E substitution in Mat-Mu1B increases Km by 20 fold for tRNA suggesting that Mat-Mu1B interacts with the tRNA substrate and is required for catalytic function. This is particularly interesting as little is known about the function of this region in the body domain of the enzyme as it shows a low level of sequence conservation (Li *et al.*, 2002, Tomita *et*

al., 2004, Kuhn *et al.*, 2015, Toh *et al.*, 2009, Xiong and Steitz, 2004). In contrast, no major changes in structure, stability or activity were associated with Mat-Mu1A. However, a protein containing modifications in Mat-Mu1A showed fewer proteolytic fragments than the native enzyme suggesting that modification of this region resulted in a conformational change that altered access to proteolytic processing sites. Perhaps it is this conformational change that affects mitochondrial import. These data suggest that while changes in Mat-Mu1A may affect mitochondrial import either by altering the folding of tRNA nucleotidyltransferase or by disrupting interactions with other proteins in the cell, any role for Mat-Mu1B in protein targeting is more likely related to its role in substrate binding.

3.4 Materials and Methods

3.4.1 Plasmid construction

A cDNA coding for the mature domain (starting at amino acid 69) of native *Arabidopsis* tRNA nucleotidyltransferase was cloned into a modified pGEX-2T expression plasmid (Shan *et al*., 2007) for expression in *E. coli* or into pG131-2 for expression in yeast. Quikchange™ mutagenesis (Stratagene) was performed using templates and primers as previously described (Schmidt von Braun *et al.*, 2007) and modified DNA was moved as a *Bpu*1102I fragment between vectors. The primers used are listed in Table 3.1.

Table 3.1: Oligonucleotides used in present study

Oligonucleotide	Primer sequence (5' ----> 3')
K414E F	GTCAACCACATTTTCGAATTCTCCATGAAGAGG
K414E R	CCTCTTCATGGAGAATTCGAAAATGTGGTTGAC
K418E F	TTTAAATTCTCCATGGAGAGGAAGACC
K418E R	GGTCTTCCTCTCCATGGAGAATTTAAA
K418Q F	CCACATTTTTAAATTCAGCATGCAGAGGAAGACCAGT
K418Q R	ACTGGTCTTCCTCTGCATGCTGAATTTAAAAATGTGG
K420E F	CCATGAAGAGGGAGACTAGTGACGCTGAAAC
K420E R	GTTTCAGCGTCACTAGTCTCCCTCTTCATGG

3.4.2 Plasmid shuffling

Plasmids carrying the yeast *HIS3* gene and expressing various tRNA nucleotidyltransferase GFP fusion proteins were transformed (Schiestl and Gietz, 1989) into strain SCDT8 (*MATα leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15* cca1::TRP1) bearing a derivative of pRS316 (Sikorski and Hieter, 1989) containing the wild-type *CCA1* gene of *Saccharomyces cerevisiae*. Cells carrying both plasmids were selected on SC medium minus uracil and histidine (Sherman, 2002). For plasmid shuffling, patches of two representative transformants were replica-plated from synthetic complete medium minus uracil and histidine, to YPD medium (Sherman, 2002), incubated 1 day at 23°C, and then replica plated to 5-FOA medium (Boeke *et al.*, 1984) supplemented with 30 mg/l leucine, lysine and tyrosine and 20 mg/l adenine, histidine and tryptophan, with or without 5-FOA with incubation for 3 days at 23°C. FOA-resistant cells then were replica-plated to YPD or YPG plates and incubated at 23°C for 3 or 4 days, respectively.

3.4.3 Expression and purification of tRNA nucleotidyltransferase

A 5 ml overnight culture of *E.coli* BL21 (DE3) carrying the plasmid of interest was diluted in 1.3 l of YT + ampicillin (100 µg/ml) and placed on a shaker at 37^{o}C and 225 rpm for 2-3 hours until an OD_{600} of 0.4-0.6 was achieved. The cells then were induced with a final concentration of 0.02% w/v of D-lactose + 0.5 mM IPTG and returned to the shaker at 19^{o}C and 193 rpm for 16-20 hours. Cells from 2.6 l were pelleted by centrifugation at 4^{o}C in a Beckman JA-10 rotor for 15 minutes at 6 000 rpm. Cell pellets were stored at -80^{o}C and frozen a minimum of 1 hour before use. Cell pellets were thawed on ice and resuspended in 2 ml/g cold lysis buffer (PBS + 1 mM EDTA) by vortexing. The cell suspension was passed through a French Pressure cell press (Thermospectonic) up to four times at 1 000 psi. The resulting cell lysate was centrifuged at 18 000 rpm in a Beckman JA-20 rotor for 30 min at 4°C. The supernatant was kept and centrifuged again as above. After centrifugation, the cleared lysate was cycled through a 10 cm Bio-Rad column containing 1-2 ml of glutathione Sepharose fast flow 4B resin (GE Healthcare or GoldBio) overnight at 4^{o}C at a flow rate of 1 ml/min. The column subsequently was washed with 300-500 ml of PBS (pH 7.4) to eliminate non-specific binding and the protein of interest was eluted with 20 ml of elution buffer (15 mM reduced glutathione, 50 mM Tris-HCl (pH 8.3), 140 mM NaCl and 2.5 mM $CaCl_2$). Fractions (1 ml) were collected and analyzed by SDS-PAGE and fractions containing the protein of interest were saved. The fractions were placed in an 8 kDa dialysis bag (SpectroPor) with 50 units of thrombin (GE Healthcare) with dialysis against 5 l of dialysis buffer (10 mM Tris, pH 8, 140 mM NaCl, 2.5 mM $CaCl_2$) overnight at 4^{o}C to cleave the GST tag and eliminate glutathione. After dialysis the protein was passed through a freshly regenerated glutathione Sepharose column to remove GST and any protein retaining the GST tag. The eluted protein was dialyzed overnight against 5 l of PBS at 4^{o}C in a 50 kDa dialysis bag to reduce the level of thrombin. SDS-PAGE analysis was used to ensure that the GST was removed from the protein solution. The protein solution then was passed once more through the glutathione Sepharose column if needed to ensure purity. Glycerol was added to the protein solution to

a final concentration of 10% and the samples stored at -80^{o}C. Routinely, protein concentrations were determined based on the predicted molecular mass (61 943 g mol^{-1}) and extinction coefficient at 280 nm (71390 M^{-1}cm^{-1}) for Arabidopsis tRNA nucleotidyltransferase (Gasteiger, 2006). To ensure that all samples were free of RNA, RNaseA (Bioshop), was added to a concentration of 30 ng/ml and dialysis was repeated as above in PBS (pH 7.4).

3.4.4 Biophysical characterization of tRNA nucleotidyltransferase

3.4.4.1 Circular dichroism Spectroscopy

Secondary structure determination was done using circular dichroism spectroscopy. A Jasco-815 circular dichroism spectrophotometer along with the "Spectrum Measurement" program (Jasco) was used. The instrument was set to scan over the region from 200 nm to 280 nm using a 0.1 or 0.2 cm cell. At standard sensitivity, a bandwidth of 1 nm was used, along with a response time of 0.25 seconds, data pitch of 0.2 nm, and a scanning speed of 20 nm/min. Data were accumulated five times at 20^{o}C using a Pelletier water bath accessory while nitrogen was set to flow through the instrument at a constant rate of 3 l/min. Protein (0.1-0.4 mg/ml) from freezer stocks was used for the spectrum measurements. All protein samples were RNaseA treated as described previously. Prior to taking measurements, samples were centrifuged at 14 000 g for 5 minutes at 4^{o}C. The data obtained were smoothed using the smoothing function in the Spectra Analysis (Jasco) program.

3.4.4.2 Thermal denaturation monitored by circular dichroism

Thermal denaturation was monitored by circular dichroism spectroscopy using 0.1-0.5 mg/ml protein. The program used was "Variable Temperature" where the signal was monitored at 222 nm at standard sensitivity using a bandwidth of 1 nm, response time of 0.25 seconds, and a data pitch of 0.2^{o}C. A temperature slope of 25^{o}C/hour was used while the start and end temperatures were 25^{o}C and 90^{o}C. A 0.2 cm cell was used and was sealed to limit evaporation. Prior to taking measurements, samples were

centrifuged and prepared as described above (3.4.4.1). Data obtained were smoothed using the smoothing function in the Spectra Analysis (Jasco) program.

3.4.4.3 Tryptophan fluorescence spectroscopy

Tertiary structure determination was done by using fluorescence spectroscopy. The instrument used was a Varian Cary Eclipse Fluorescence Spectrophotometer along with a Varian single cell Pelletier Cary accessory (water bath). The application "Scan" was used to obtain emission spectra at room temperature averaged over 10 scans between 310 nm to 400 nm with an excitation wavelength of 280 nm or 295 nm. Scan speed was set to medium (600 nm/min) with a 1.0 nm data sampling interval, 5 nm excitation and emission slit widths, and voltage was set to medium (600 volts). A 1.0 cm cell was used with aliquots of 1 ml at 1.2 µM of protein determined by OD_{280} for the spectrum measurements. Prior to taking measurements, samples were RNaseA treated, dialyzed, and centrifuged as described above. The data obtained were smoothed using the moving average smoothing function.

3.4.4.4 Fluorescence quenching with baker's yeast tRNA

Fluorescence quenching experiments were done to determine apparent dissociation constants for proteins using tRNA as a quenching agent. The instrument used was a Varian Cary Eclipse Fluorescence Spectrophotometer along with a Varian single cell Pelletier Cary accessory (water bath). The application "Scan" was used to obtain emission spectra at room temperature averaged over 10 scans between 310 nm to 400 nm with an excitation wavelength of 295 nm. Scan speed was set to medium (600 nm/min) with a 1.0 nm data sampling interval, 5 nm excitation and emission slit widths, and voltage was set to medium (600 volts). Baker's yeast tRNA (Roche) was extensively purified by phenol extraction and ethanol precipitation. A volume of 1 ml of either 2.6 µM or 1.2 µM of protein was placed in a 1 cm cell and titrated by adding 0.1 µM or 0.2 µM of tRNA at room temperature. Ultraviolet visible spectral scans were taken between titrations to take into account any inner filter effect.

3.4.4.5 Proteolysis with chymotrypsin

Bovine pancreatic α-chymotrypsin (Sigma-Aldrich) at a final concentration of 0.01 units/ml was used to digest approximately 40 µg of protein in 1ml PBS (pH 7.4) and 10% glycerol. The proteins were incubated at either 15°C or 25°C and 20 µl aliquots of the digests were removed at different time intervals and stopped by addition of SDS-loading dye (Sambrook *et al*., 2008) and boiling for five minutes. Samples were stored on ice and visualized by SDS 13% polyacrylamide gel electrophoresis.

3.4.5 Enzyme activity assays

3.4.5.1 Generation of radiolabelled transcripts

The plasmids G73 and pmBsDCCA, both generous gifts from Dr. Alan Weiner (University of Washington), were originally constructed by Cho *et al.* (Cho *et al.*, 2003) and Oh and Pace (Oh and Pace, 1994), respectively and contain the *Bacillus subtilis* $tRNA^{Asp}$ gene modified to produce tRNAs with specific 3' ends through run-off transcription. Digesting plasmid G73 with *Fok*I results in the linearized template used to generate a tRNA-N product. Similarly, restriction digestion of pmBsDCCA with *Fok*I, *Bpi*I or *Bst*OI results in tRNA-NC, tRNA-NCC and tRNA-NCCA run-off products, respectively. Approximately 35 µg of plasmid DNA were digested with the appropriate restriction enzyme (12 U *Fok*I, 40U *Bpi*I or 40U *Bst*OI), phenol extracted and ethanol precipitated. Each 100 µl run-off transcription reaction contained 20 µL of 5X transcription buffer (Thermo Scientific), 5 nmoles ATP, 50 nmoles each of CTP, GTP and UTP (Thermo Scientific) and 50 µCi (~16 pmoles) [α-^{32}P] ATP (10 µCi/µL, 3000 Ci/mmol) (Perkin-Elmer), ~5 µg of linearized DNA template, 60 units of T7 RNA polymerase (Thermo Scientific) and nuclease-free water and was incubated for 2-3 hours at 37^{o}C. The products were phenol extracted, ethanol precipitated and gel purified via polyacrylamide gel electrophoresis. The concentration of each runoff transcription product was determined by scintillation counting (LKB WALLAC-1217 RackBeta liquid scintillation counter) using a standard curve.

3.4.5.2 Standard activity assays

Each 10 µl standard assay mixture (Leibovitch *et al*., 2013) contained 100 mM glycine buffer (pH 9), 10 mM $MgCl_2$, 1 mM ATP, 0.4 mM CTP, 1 µl radiolabeled template, 0.4 µl RNAsecure (25X) (Invitrogen) and 100 ng of protein. When proteins showed dramatically reduced activity it was necessary to add more protein (as indicated). The assay mixtures were heated to 65°C for 10 minutes (to activate the RNAsecure™ reagent) and cooled prior to adding transcript and protein. Assays were initiated with the addition of protein and stopped with 10 µl of Peattie's (2X) loading dye (Peattie, 1979). Samples were heated to 70°C and cooled prior to loading onto a 50 cm X 30 cm 7M urea 12% polyacrylamide gel and electrophoresed for 7-8 hours at 2000 Volts. After electrophoresis the gel was placed against a phosphorImager screen (GE Healthcare) for at least 30 minutes, the screen was developed and the transcripts detected using the Typhoon™ TRIO Variable Mode Imager (GE Healthcare).

3.4.5.3 Standard kinetic assays

For kinetic analysis the standard assay as described in 3.4.5.2 was used except with 10-500 ng protein and increasing amounts of purified Baker's yeast tRNA (0.25 µg - 50 µg). Each assay tube contained enough reagent for six individual assays. At each time point, 10 µl of the assay mixture was removed and treated as described (section 3.4.5.2). The ratio of reactant to product (tRNA-NCC/tRNA-NCCA or tRNA-NC/tRNA-NCC) was measured by densitometry and initial rates were determined for varied amounts of tRNA substrate.

3.4.6 Homology modeling of Arabidopsis tRNA nucleotidyltransferase

The homology model of Arabidopsis tRNA nucleotidyltransferase was derived from four available crystal structures of Class II tRNA nucleotidyltransferases: CCA-adding enzymes from *Thermotoga maritima* (Toh *et al.*, 2009), *Homo sapiens* (Kuhn *et al.*, 2015), *Bacillus stearothermophilus* (Li *et al.*, 2002) and the A adding enzyme from *Aquifex aeolicus* (Tomita *et al.*, 2004). Their PDB numbers

are 3H37, 4X4W, 1MIV and 1VFG, respectively. The primary and secondary structures of the above resolved crystals were aligned using MODELLER v.9.13 (Ben Webb). The homology model was generated using MODELLER (Ben Webb) incorporating the tRNA substrate from *Aquifex aeolicus* (1VFG) and both catalytic magnesium ions and ATP from *Thermotoga maritima* (3H37). The homology model with the lowest energy level was submitted to MolProbity server (http://molprobity.biochem.duke.edu/) where atom-contact and geometry were analyzed (191.51/1000 atom clashes) which was carried over from the available crystal structures. The final homology model was aligned with PDB files 3H37, 1VFG, 4X4W and 1MIV using conserved Motifs A and D generating RMS values of < 0.710 using PyMol (Schrödinger).

3.5 Results

3.5.1 Activity of tRNA nucleotidyltransferase and variants

Preliminary enzyme assays (Fig. 3.2) revealed activity for the native enzyme and no measurable activity for the Mat-Mu1 variant consistent with what we had seen previously (Leibovitch *et al.*, 2013). In addition, these assays indicated that the Mat-Mu1A variant reduces activity only to a small degree whereas the Mat-Mu1B variant reduces activity to the same degree as the larger Mat-Mu1 variant (Fig. 3.2).

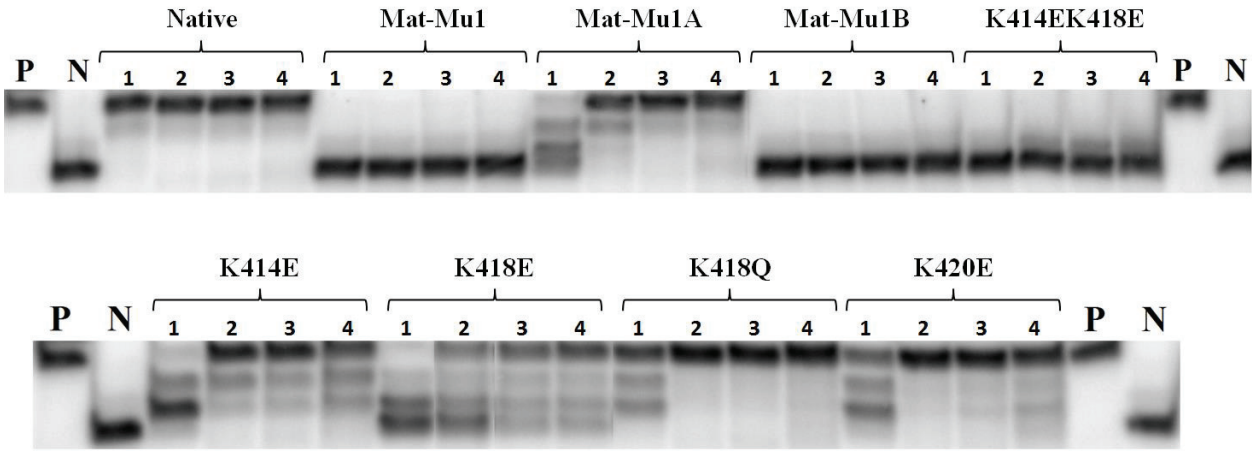


Figure 3.2: Activity assays with *A. thaliana* tRNA nucleotidyltransferases
The proteins indicated were incubated as described with a tRNA substrate for 2, 10, 30 or 60 minutes as indicated by numbers 1-4 respectively. The amount of enzyme used in each of the assays was 1 ng. The reaction products were separated by electrophoresis on 7 M urea/12% polyacrylamide gels and visualized using the Typhoon™ TRIO Variable Mode Imager. P is the active enzyme positive control (tRNA with a complete CCA sequence), N is the boiled enzyme negative control (original tRNA template lacking the CCA sequence).

This is consistent with the *in vivo* data from yeast showing that the Mat-Mu1A variant is viable but the Mat-Mu1B variant is not (Fig. 3.3). To determine which portions of Mat-Mu1B are needed for activity, specific amino acids were changed. A loss of activity is observed for the K414EK418E double variant, and the K418E single variant (Fig. 3.2). In contrast, the K414E and K420E changes show only a minor loss of activity indicating that K418E plays the largest role in reducing enzyme activity.

Interestingly, yeast cells expressing the cytosolic forms (starting from ATG3 and lacking the first 68 amino acids) of any of these variants (K414EK418E, K418E, K414E or K420E) were viable (Fig. 3.3), however, when the same variants were expressed from ATG1 and contained the 68 amino acid amino-terminal targeting signal, cells containing the K414EK418E and K418E variants showed limited growth on both fermentable (glucose) and non-fermentable (glycerol) carbon sources while cells containing the K414E and K420E variants showed growth on both carbon sources (Fig. 3.3). To further address the role of K418 in enzyme activity, a glutamine was introduced at this position and activity was retained (Fig.

3.2). As expected, cells expressing the K418Q variant (from ATG1 or ATG3) which retained a higher level of activity *in vitro* showed growth *in vivo* on both carbon sources (Fig. 3.3).

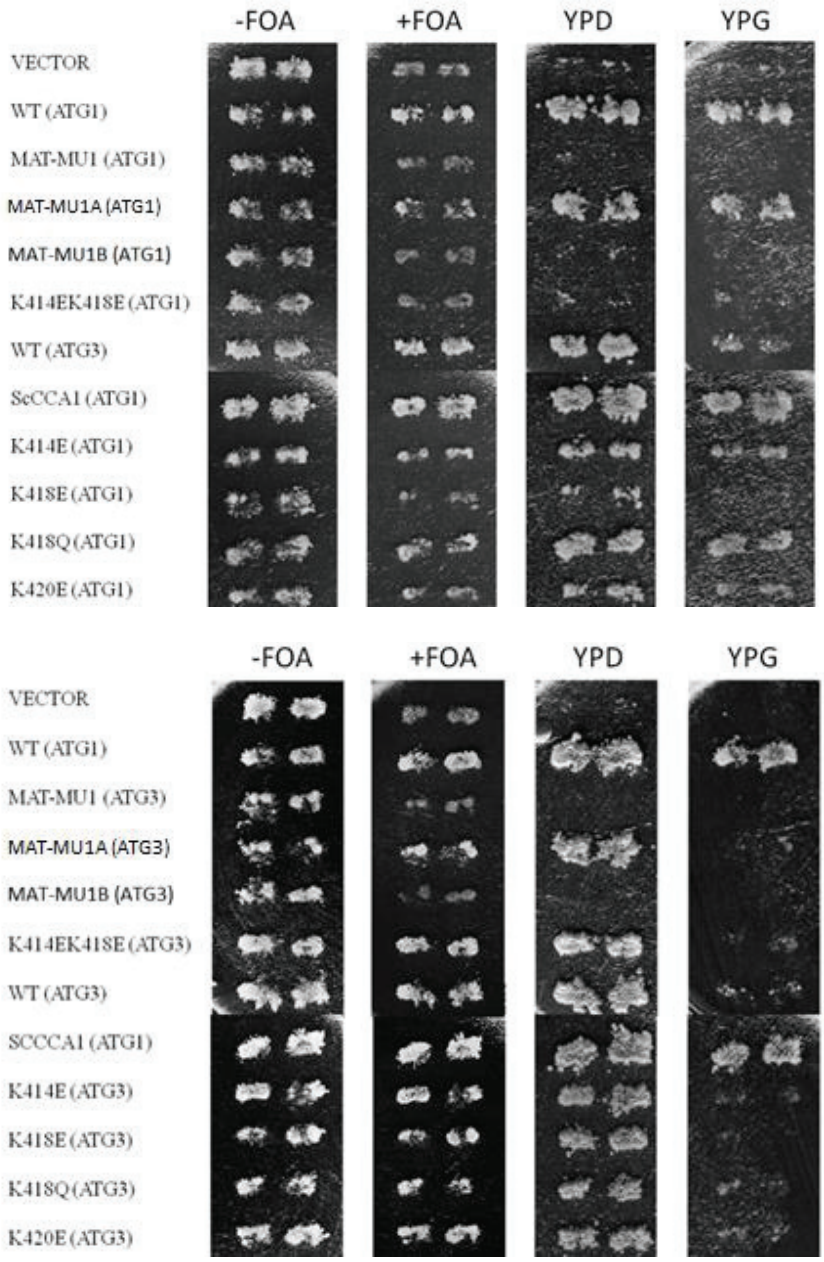


Figure 3.3: Viability Yeast assays of *A. thaliana* tRNA nucleotidyltransferases
An *S. cerevisiae* strain bearing a null mutation in the chromosomal *CCA1* gene and plasmid-borne *CCA1* and *URA3* genes was transformed with the *HIS3* -bearing G131-2 plasmid alone or G131-2 derivatives expressing Arabidopsis tRNA-NT–GFP native or variant fusions under the control of the *TDH3* promoter. To accomplish loss of the *URA3-CCA1* bearing plasmid, two independent His+ Ura+ transformants were patched on SC-ura-his medium, replica-plated to YPD (glucose) medium, incubated for 24 h, and replica-plated to medium with or without FOA. The ability of FOA-resistant yeast to grow on different carbon sources was tested by subsequently replica-plating to YPD (glucose) or YPG (glycerol) medium.

3.5.1.1 Kinetics

To determine which kinetic parameters were affected by these amino acid substitutions, k_{cat} and K_M values for AMP addition to tRNA lacking only the terminal AMP residue and for CMP and AMP addition to tRNA lacking the terminal CMP and AMP residues were determined. Using Michaelis-Menten kinetics as a model, we were able to determine apparent K_M and k_{cat} values using a fixed amount of radiolabelled tRNA template and increasing amounts of purified baker's yeast tRNA. Sample data for the native enzyme and the Mat-MU1, Mat-Mu1A and Mat-Mu1B variants are shown (Fig. 3.4). In agreement with the activity assays, the native and Mat-Mu1A enzymes show comparable K_M values of approximately 500 µM for ATP, 50 µM for CTP and a turnover of 4 s^{-1} (Table. 3.2). There also is a small (approximately two-fold) increase in apparent K_M for tRNA binding such that all of these data are consistent with the preliminary activity assays (Fig. 3.2) and what was observed in yeast (Fig. 3.3). In contrast, although the K_M values for ATP and CTP are similar to the native enzyme for the Mat-Mu1B variant, there is a twenty fold increase in apparent K_M for tRNA binding and a sixteen fold decrease in turnover number (Table. 3.2). These data reflect the *in vivo* viability (Fig. 3.3) and *in vitro* (Fig. 3.2) activity assays.

The more detailed analysis of Mat-Mu1B shows that the K414E and K418Q variants demonstrate an approximately two-fold increase in K_M as compared to the native enzyme while the K414EK418E double variant and the K418E and K420E single variants showed increases in K_M's for tRNA of approximately four-fold, ten-fold and five-fold, respectively. In addition, while the K414EK418E, K414E and K418E variants showed decreases in k_{cat} of approximately 20-fold, four-fold and five-fold, respectively, the K418Q variant shows no difference in k_{cat} and the k_{cat} for the K420E variant increases by 50%. While these data are consistent with the decreases in activity seen in the preliminary *in vitro* activity assays, they raise intriguing questions with respect to the *in vivo* yeast studies (Fig. 3.3).

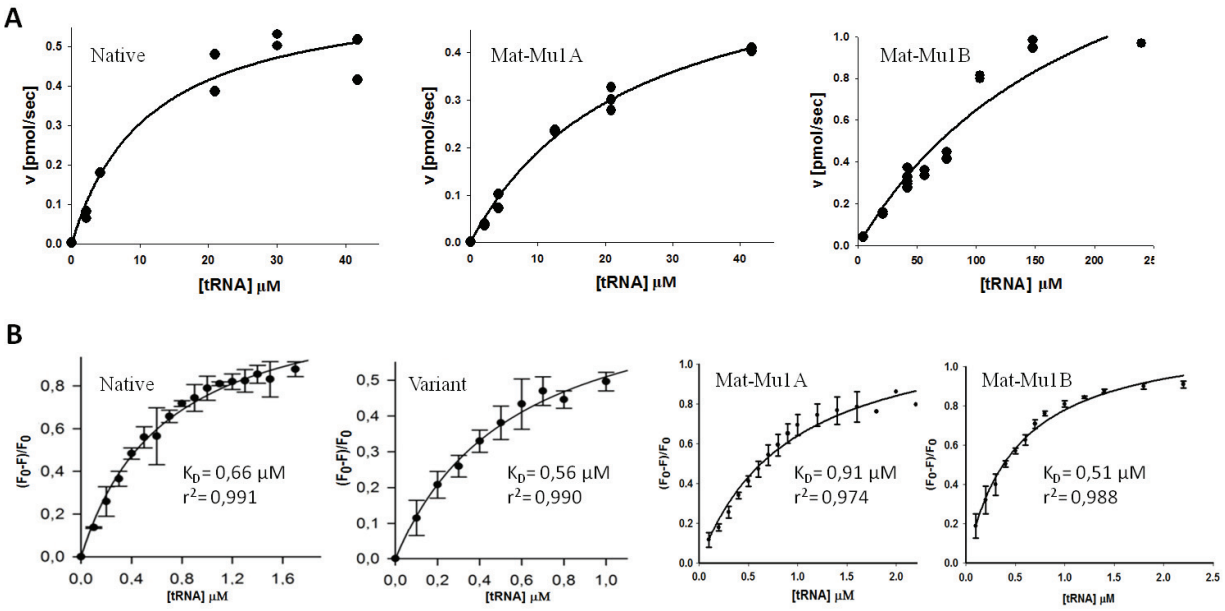


Figure 3.4: Binding and kinetics of *A. thaliana* tRNA nucleotidyltransferases
(A) The rate of addition of ATP to fixed amounts of tRNA-NCC template by a fixed amount of enzyme with increasing amounts of tRNA was determined and plotted. The data were analyzed by a least-squares method using the equation for Michaelis–Menten kinetics, v =Vmax [S]/(Km+[S]), with the results shown as the line. (B) The quenching of Mat-Mu1A or Mat-Mu1B fluorescence by tRNA (solid line) was analysed as previously described (Leibovitch et al., 2013). The data were analysed by least-square fit analysis to $(F_0 - F)/F_0 = MAX \times [tRNA]/(K_D + [tRNA])$, and the determined K_D as well as the r^2 value of the fit is indicated for all proteins.

Table 3.2: Binding and kinetics of tRNA nucleotidyltransferase enzymes

Protein	K_D (µM) for tRNA	K_M (µM) tRNA	k_{cat} (s^{-1}) tRNA	kcat/K_M tRNA	K_M (µM) ATP	K_M (µM) CTP
Native	0.66	11 ± 3	4 ± 0.4	0.36	460 ± 68	48 ± 6
Mat-Mu1A	0.91	24 ± 3	4 ± 0.3	0.17	510 ± 110	43 ± 7
Mat-Mu1B	0.51	210 ± 47	0.25 ± 0.03	0.0012	750 ± 170	34 ± 5
K414EK418E	0.67	42 ± 7	0.16 ± 0.01	0.0038	690 ± 200	51 ± 17
K414E	ND	27 ± 7	1.1 ± 0.2	0.041	1000 ± 350	19 ± 7
K418E	0.18	120 ± 20	0.75 ± 0.05	0.0063	510 ± 180	27 ± 5
K418Q	ND	26 ± 7	4.3 ± 0.5	0.17	510 ± 180	35 ± 13
K420E	0.3	49 ± 14	6.6 ± 1	0.14	470 ± 140	54 ± 17

ND indicates not determined

3.5.1.2 Binding

Fluorescence quenching experiments using baker's yeast tRNA were performed for proteins with Mat-Mu1A or Mat-Mu1B to obtain apparent binding constants (K_{Dapp}) for tRNA of 0.91 µM and 0.51 µM, respectively (Fig. 3.4, Table. 3.2). These values are in good agreement with the values previously obtained for the native (0.66 µM) and Mat-Mu1 (0.56 µM) variant (Leibovitch *et al.*, 2013).

3.5.2 Structure and stability

Our previous studies showed that both the native and Mat-Mu1 enzymes showed comparable circular dichroism and fluorescence spectra revealing no significant changes in secondary or tertiary structure (Leibovitch *et al.*, 2013). All variants tested here (Mat-Mu1A, Mat-Mu1B, K414EK418E, K414E, K418E, K418Q, K420E) showed comparable CD (Fig. 3.5) and fluorescence (Fig. 3.6) spectra with the native enzyme as expected.

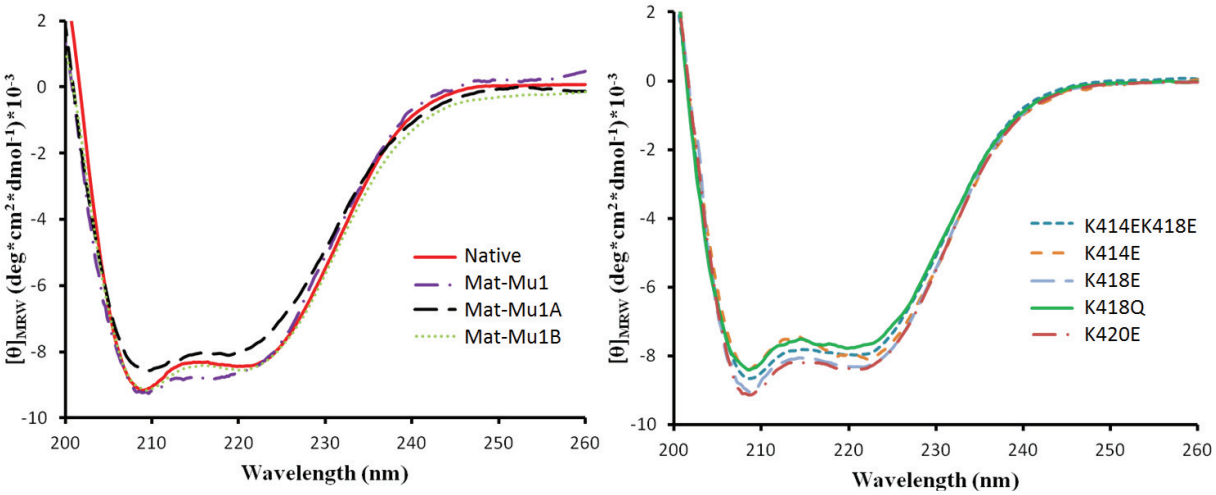
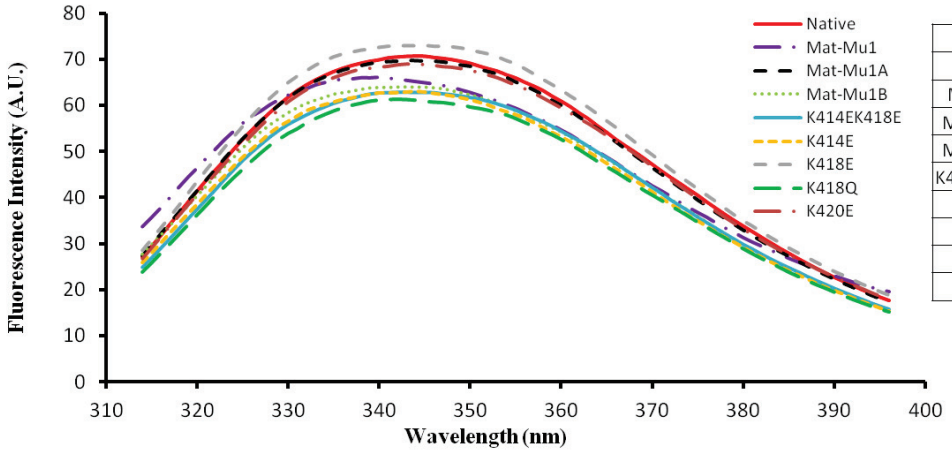


Figure 3.5: Secondary structure comparisons of *A. thaliana* tRNA nucleotidyltransferases
The CD spectra of proteins purified from *E. coli* were recorded at 20°C. All samples were RNaseA treated with extensive dialysis in PBS (pH 7.4) to remove RNA and glycerol.



Protein	Fluor. Max
Native	344
Mat-Mu1	340
Mat-Mu1A	343
Mat-Mu1B	343
K414EK418E	343
K414E	344
K418E	345
K418Q	342
K420E	344

Figure 3.6: Intrinsic fluorescence of *A. thaliana* tRNA nucleotidyltransferases
Approximately 1.2 uM of each protein was excited at 295 nm and the fluorescence emission was monitored from 310 to 400 nm. All samples were RNaseA treated with extensive dialysis in PBS (pH 7.4) to remove RNA and glycerol.

Circular dichroism spectroscopy also was used to assess thermal stability by monitoring the CD signal at 222 nm. Previously, it was demonstrated that the native enzyme undergoes a three state equilibrium with two transition states at 42°C and 80°C respectively and that the Mat-Mu1 variant displays an ambiguous two state equilibrium with a single transition state suggesting a difference in thermal stability between the two proteins (Leibovitch *et al.*, 2013). Here, a closer look at the Mat-Mu1 variant showed that it also undergoes a three state equilibrium (Fig. 3.7) with its first transition state at 10°C lower than the native enzyme (Fig. 3.7) and its second transition state about 12°C lower than the native enzyme (Fig. 3.7). To see if there is any connection between thermal stability and loss of activity, the thermal stability was assessed for each of the new variants. Surprisingly, all of the new variants except for Mat-Mu1A show a clear three state equilibrium with two transition states (Fig. 3.7). The Tm at the first transition state in each case was near that of the native enzyme with the Mat-Mu1B variant showing the largest decrease of only 4.5°C (Fig. 3.7). Almost all of the new variants showed a second transition state at approximately 80°C like the native enzyme except for the Mat-Mu1B and K414EK418E

variants which showed a second transition state at similar temperatures as the Mat-Mu1 variant (Fig. 3.7).

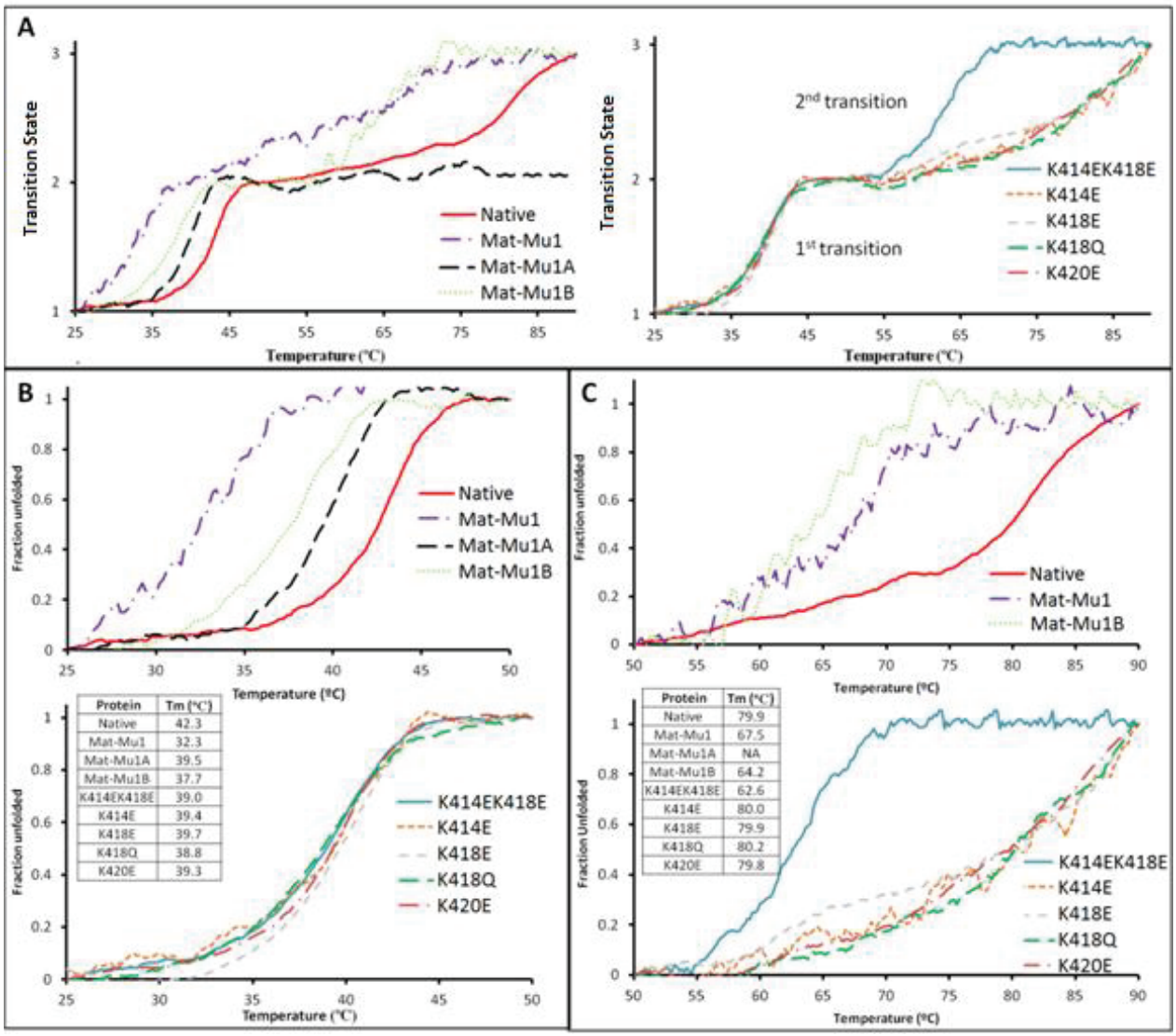


Figure 3.7: Thermal unfolding of *A. thaliana* tRNA nucleotidyltransferases
(A) Ellipticity at 222 nm was monitored between 25°C and 90°C for each of the protein samples illustrating a two or three state unfolding pattern for the enzymes. (B) Thermal denaturation of the first transition state was monitored as previously described. The fraction of protein unfolded was defined by the equation F = 1/ΔθT x θ - θi/ΔθT where F is the fraction of protein unfolded, θ is ellipticity measured in mdeg, and ΔθT is the total change in ellipticity between the initial and final temperatures of interest; ΔθT= θf -θi. θi is the ellipticity value at the start temperature of interest (25°C), and θf is the final ellipticity value at the end temperature of interest (50°C). (C) Thermal denaturation of the second transition was monitored as previously described. Fraction of protein unfolded was defined by the equation in (B) except that the value at the starting temperature is 50°C and the final temperature is 90°C.

To explore further whether any of the amino acid substitutions altered the structure of the protein limited proteolysis with chymotrypsin was used. While the native protein and Mat-Mu1B variant show similar cleavage products of 45 kDa, 40 kDa, 37 kDa, 33 kDa and 27 kDa after the first five minutes, Mat-Mu1A lacks the 27 kDa cleavage products (Fig. 3.8).

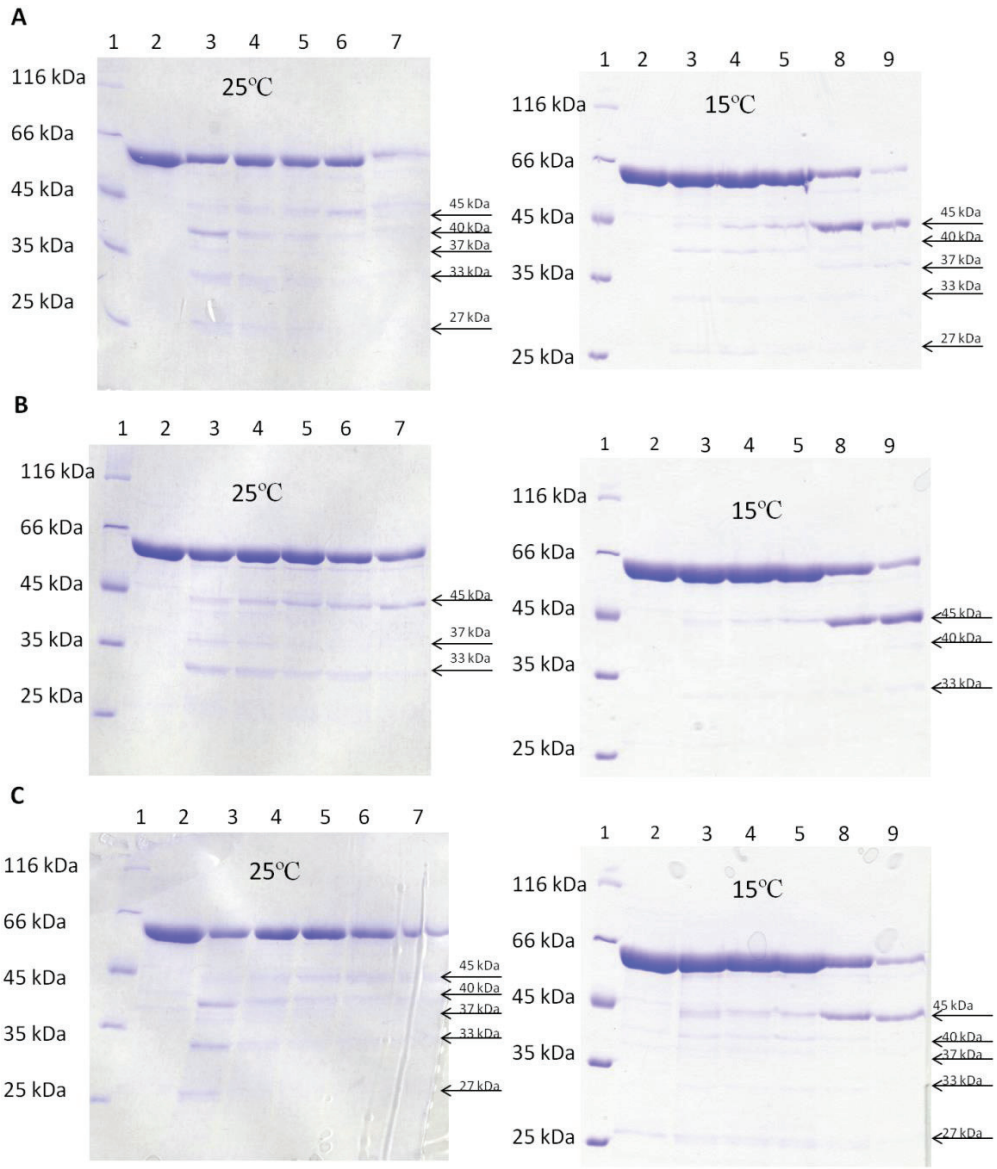


Figure 3.8: Limited proteolysis of *A. thaliana* tRNA nucleotidyltransferases with α-chymotrypsin
The native (A), Mat-Mu1A (B) and Mat-Mu1B (C) proteins were incubated with α-chymotrypsin in PBS (pH 7.4) and 10% glycerol at 15°C or 25°C and aliquots removed at the time points indicated. Lane 1 is the molecular weight ladder (Pierce unstained protein molecular weight marker 26610), Lane 2 is the protein sample prior to proteolysis and lanes 3-9 are protein samples after 5 minute, 15 minute, 30 minute, 1 hour, 6 hour and 21 hour incubation, respectively.

3.6 Discussion

Any protein's function is dependent on the interplay of its structure, stability, activity and localization. Here, a specific region of the Arabidopsis tRNA nucleotidyltransferase (amino acids 399-420 in the body domain) is studied to understand how it affects structure, stability, and activity to explore how changes in this region alter localization. This protein functions in the nucleus, cytosol, mitochondria and plastids of plant cells and all isoforms are encoded by a single nuclear gene (Schmidt von Braun *et al.*, 2007). Simply the presence of the mature domain of the protein has been shown to alter the targeting of a reporter protein (Schmidt von Braun *et al.*, 2007) and amino acid substitutions in the mature domain of the protein far from the amino-terminal targeting signal likewise have been shown to shift distribution from mitochondria to plastids and the nucleocytosol (Leibovitch *et al.*, 2013).

Protein targeting is defined by a combination of factors ranging from the import competence and folding of the protein, *e.g.*, proteins are transported co-translationally across the endoplasmic reticular membrane such that they only fold upon entry (reviewed by Leslie, 2005) to the targeting signals contained on the protein of interest (Lange *et al.*, 2007, Kunze and Berger, 2015) to the transport machinery that must recognize these signals (Kimoto *et al.*, 2015, Saitoh *et al.*, 2007, Radons, 2016). Proteins must adopt an import competent structure (Bauer *et al.*, 2015) and fold such that targeting signals are accessible and arranged in a specific manner to allow chaperones (Deshaies *et al.*, 1988) and/or accessory proteins (Hachiya *et al.*, 1993) and the import machinery to interact with them. Proteins targeted to multiple locations must have an even more complex arrangement of interactions (*e.g.*, signal competition, splice variants, protease cleavage and co-translocation piggyback transport) to enable appropriate targeting to the appropriate destination (Porter *et al.*, 2015).

Two regions of the *Arabidopsis thaliana* tRNA nucleotidyltransferase showed primary sequence characteristics (stretches of basic amino acids) typical of nuclear localization signals (Cokol *et al.*, 2000)

and were surface exposed based on modelling of the available related crystal structures suggesting that they might function as nuclear localization signals. Surprisingly, modifications in these regions of the protein: amino acids 399-420 in the body of the protein (Mat-Mu1) and amino acids 535-540 at the C-terminus of the protein (Mat-Mu2) both affected targeting, but not as expected. In both cases, nuclear localization was actually improved when the positively charged amino acids in these positions were converted to negative charges (Leibovitch *et al.*, 2013). Specifically, in proteins modified at Mat-Mu1 and containing amino acids six to 68 of the amino-terminal targeting signal, protoplasts showing nuclear localization increased from 9% to 31% suggesting that this change increased nuclear localization. When the same protein starting from the first in frame start codon was used, the number of cells showing nuclear localization increased marginally from 3% to 5% and only when the second start codon was removed did nuclear localization decrease (dropping from 11% to 7% of protoplasts showing nuclear localization). In all of these cases, there was a decrease in mitochondrial targeting and a concomitant increase in plastid targeting (Leibovitch *et al.*, 2013). These data suggested that the distribution of this protein was not simply defined by the presence of the amino-terminal targeting sequence and that the region of the protein containing Mat-Mu1 may play some role in protein localization. While the Mat-Mu1 variant showed no detectable structural changes, it did show a reduced transition temperature (10^{o}C lower than native) and a loss of activity below the limits of detection with the standard enzyme assay (Leibovitch *et al.*, 2013). This suggested that there may be a role for protein folding/stability or activity (but seemingly not structure) in defining the localization of this protein. To further explore the roles of specific amino acids in the Mat-Mu1 region of tRNA nucleotidyltransferase in structure, stability, activity and localization, site-directed mutagenesis was used to introduce a number of discrete amino acid changes and the physicochemical properties of these variants explored.

3.6.1 The roles of Mat-Mu1A and Mat-Mu1B

In the Mat-Mu1 variant, each lysine in the stretch of amino acids from 399-420 was changed to a glutamic acid and the single arginine was converted to methionine resulting in two stretches of three negative charges separated by nine amino acids (Fig. 3.1). Here, this region was divided into two parts where one of these stretches of positively charged amino acids was changed to negative charges generating either Mat-Mu1A (amino acids 399-404, now containing three negative charges) or Mat-Mu1B (amino acids 414-420, now containing three negative charges). Based on the model of Arabidopsis tRNA nucleotidyltransferase (Fig. 3.1), Mat-Mu1A and Mat-Mu1B are both surface exposed and are positioned on opposite faces of the protein. Mat-Mu1B is predicted to be located near the acceptor stem and TΨC stem and loop of the tRNA substrate while Mat-Mu1A is located on the other side of the protein away from the tRNA substrate (Fig. 3.1). This suggests that Mat-Mu1B may interact with the tRNA substrate such that changes here may alter the enzyme's activity while Mat-Mu1A seems less likely to affect enzyme activity.

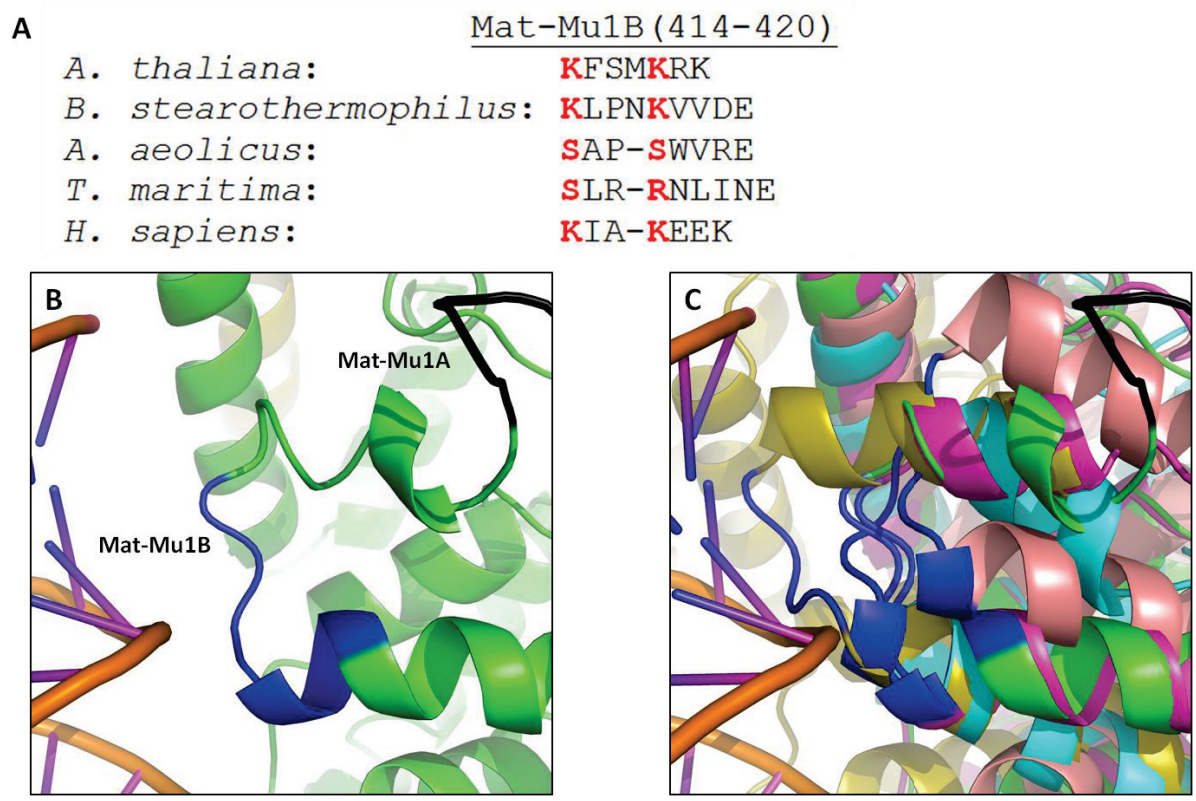


Figure 3.9: Conservation and location of Mat-Mu1
(A) Sequence alignment of the Arabidopsis Mat-Mu1B region based on modeling with the available crystal structures of Class II tRNA nucleotidyltransferases from *Bacillus stearothermophilus, Aquifex aeolicus, Thermotoga maritima* and *Homo sapiens*. In red are the amino acids most likely to interact with the phosphate backbone of the tRNA. (B) A close up of Mat-Mu1A (black) and Mat-Mu1B (blue) in the Arabidopsis homology model viewed in cartoon. The carbon backbone of the protein is green and the phosphate backbone of the tRNA is orange. (C) Overlay of the Arabidopsis model (green) with the crystal structures from *Bacillus stearothermophilus* (1MIW, teal)*, Aquifex aeolicus* (1VFG, yellow)*, Thermotoga maritima* (3H39, salmon) and *Homo sapiens* (4X4W, pink).

The alignment of the Arabidopsis homology model with the Class II structures (Fig. 3.9) shows conservation of a loop and an alpha helix in the region of Mat-Mu1B among tRNA nucleotidyltransferases. In contrast, the Arabidopsis Mat-Mu1A region is found entirely in a flexible loop that is absent from the other tRNA nucleotidyltransferases suggesting that it has no conserved function.

3.6.1.1 Mat-Mu1A

As expected from the model of the protein, changes in Mat-Mu1A, thought to be on the outside of the enzyme, had only minor effects on any of the kinetic parameters associated with enzyme activity (Table 3.2) suggesting that these mutations did not affect enzymatic activity. If these amino acids were not involved in enzyme activity, the question remains as to their role in protein function and how they may have altered protein localization. While the native enzyme and Mat-Mu1B protein showed similar products when subjected to limited proteolysis the Mat-Mu1A variant showed differences as compared to the native and Mat-Mu1B variant (Fig. 3.8). Specifically, the 27 kDa product was evident in both the native and Mat-Mu1B samples but was absent from the Mat-Mu1A sample (Fig. 3.8). This difference in profile suggests a conformational change in Mat-Mu1A that limited the accessibility of the protease to hydrophobic residue(s). These data are consistent with the earlier observation that the thermal denaturation profile for Mat-Mu1A differs from that of the native or Mat-Mu1B variant (Fig. 3.7) and agree well with what was originally observed with the Mat-Mu1 variant (Leibovitch *et al*., 2013). It is interesting that the Mat-Mu1A variant does not show a second transition in thermal denaturation. Based on FTIR experiments (data not shown), the first transition observed defines the complete unfolding of the protein into intermediate soluble aggregates therefore the second transition probably involves the unfolding of those intermediate aggregates. Even though the melting temperatures are comparable between the proteins (Fig 3.7), Mat-Mu1A does not appear to undergo a second transition suggesting that the Mat-Mu1A variant remains in a stable soluble aggregate state after the first transition. This again agrees well with what was observed previously for the Mat-Mu1 variant (Leibovitch *et al.*, 2013). Taken together, all of these data implicate amino acid substitutions in Mat-Mu1A as those responsible for the change in the protein that is responsible for the altered targeting. As was argued previously, this could suggest a more tightly folded protein which is favored for plastid rather than mitochondrial import (Leibovitch *et al.*, 2013). Although these data support a structural

change in the protein as altering targeting, the possibility that changes in Mat-Mu1A change the interaction with *trans*-acting factors required for import either in the cytosol or at the mitochondrial or plastid membrane has not been eliminated. If protein import experiments show that targeting is altered in proteins bearing Mat-Mu1A alone then further experiments could be designed to explore this further. For example, specific point mutations could be introduced to see if the apparent change in structure remains linked to changes in targeting. If changes in sequence are not linked to changes in structure but localization is altered, then experiments will need to be undertaken to try to identify *trans*-acting factors that may interact with this region of the protein. The fact that the flexible loop containing Mat-Mu1A is not conserved in tRNA nucleotidyltransferases in general suggests that it is not important for function, however, its presence in plants may reflect a role in dual import between mitochondria and plastids such that it will only be seen in organisms where both organelles are present.

Changing Mat-Mu1A did not block mitochondrial targeting in yeast as cells expressing this protein were viable both on glucose and glycerol and grew similarly to those expressing the native enzyme (Fig. 3.3) indicating that this sequence is not required for mitochondrial targeting at least in yeast, but perhaps only alters the protein's folding to effect to some degree its mitochondrial targeting. In yeast, sufficient protein still enters the mitochondrion for the cells to remain viable. We have shown previously (Goring *et al*. 2013) that in yeast an enzyme expressing only about 4% of wild-type activity is sufficient for growth, so if only a portion of the Mat-Mu1A variant is entering the mitochondria it might be sufficient to show levels of growth comparable to those of the wild-type enzyme. In fact, reduced mitochondrial targeting might be favorable in yeast expressing this Arabidopsis protein as it contained the Arabidopsis amino-terminal targeting signal. Targeting all of the protein to the mitochondrion would not have allowed sufficient protein to stay in the nucleocytosol to keep the cells alive. Even altering the nucleocytosolic distribution of the yeast tRNA nucleotidyltransferase by targeting it to the nucleus resulted in reduced growth (Wolfe *et al.*, 1996). The hypothesis that these changes in Mat-Mu1A result

in reduced mitochondrial targeting agrees well with what we saw in the Arabidopsis Mat-Mu1 variant and one could envision that if the protein was not being targeted efficiently to mitochondria, it could begin to accumulate in the cytosol and be directed to other locations, *e.g*., plastid and nucleus as our data showed (Leibovitch *et al.*, 2013).

3.6.1.2 Mat-Mu1B

Based on the model that was generated (Fig. 3.1) Mat-Mu1B appears to be located near the tRNA binding region of the protein. An overlay of available crystal structures with the Arabidopsis homology model suggests that Mat-Mu1B is conserved and interacts with the tRNA substrate via serine, arginine or lysine residues in all proteins (Fig. 3.9). The kinetic data support this idea. The Mat-Mu1B variant shows large changes in both K_M and k_{cat} (Table. 3.2). In fact, most substitutions made in Mat-Mu1B resulted in decreased catalytic efficiency (K_M/k_{cat}), particularly the single K418E substitution and the K414EK418E double mutant (Table. 3.2).

The single lysine to glutamate substitution at position 418 shows native-like structure and thermostability suggesting that the drop in catalytic efficiency is a direct result of reduced substrate binding and not to a major change in the organization of active site residues. Further supporting this is the observation that the K_M's reported for both CTP and ATP do not change dramatically for any of the variants (Table. 3.2). As, ATP and CTP bind in motif D found between the head and the neck of the protein (Li *et al.*, 2002, Cho *et al.*, 2007) nearly 200 amino acids away from the region of interest, this suggests that the change at 418 is not transmitted through the protein as a structural change. The approximately 5-fold drop in k_{cat} seen in the K418E variant likely results as the tRNA is oriented less precisely at the active site making catalysis less efficient.

These data suggest that K418 plays an important role in tRNA binding and catalysis but not in protein structure or stability. It was first proposed that the body and tail domains were involved in tRNA

binding (Li *et al.*, 2002, Tomita *et al.*, 2004) as crystal structures showed exclusive interactions between the phosphate backbone of the tRNA and the protein. Experimental evidence also showed that the C-terminal portion of the enzyme must play an important role in catalysis as a deletion at the C-terminus abolished activity (Zhu and Deutscher, 1987). In addition, C-terminal domain swapping experiments between bacterial poly(A) polymerase and the tRNA nucleotidyltransferase showed that this region of the tRNA nucleotidyltransferase played an anchoring role to limit polymerization to three nucleotides (Betat *et al.*, 2004). Our homology model of the Arabidopsis tRNA nucleotidyltransferase as compared to the available crystal structures suggests that Mat-Mu1B is near the tRNA and indicates that both K414 and K418 directly interact with the sugar phosphate backbone of the tRNA (Fig. 3.10). Even though we cannot exclude the possibility of an interaction between the tRNA and residue K414, the 10 fold increase in K_M suggests that of the residues in that motif, mainly K418 is interacting with the tRNA substrate. The model constructed shows a distance of 3.74 Å (Fig. 3.10) between K418 and the sugar phosphate backbone of the tRNA indicative of a medium to weak electrostatic interaction (Jeffrey, 1997). We suggest that the change at position 418 from lysine to glutamate not only eliminates this electrostatic interaction but actually introduces electrostatic repulsion to destabilize the tRNA:protein complex. This is similar to what was observed in the active site of the *B. stearothermophilus* tRNA nucleotidyltransferase for the binding of the nucleotide triphosphates in motif D. Changing the conserved arginine 157 to glutamate resulted in a dramatic decrease in catalytic efficiency (Cho *et al.*, 2007). To support the hypothesis that electrostatic repulsion between glutamate 418 and the phosphate backbone of the tRNA is responsible for the dramatic drop in tRNA binding and reduced catalytic activity, the lysine at position 418 was converted to glutamine. While the glutamine and the glutamate are approximately the same size, the glutamine lacks the negative charge. As predicted, the K418Q substitution restored activity to near native level and altered K_M only two fold (Table. 3.2). Restoring the turnover number to native levels supports the idea that the negative charge could repel the tRNA such

that it was not aligned optimally at the active site. The change in K_M back to the near native level further supports the role of the lysine at position 418 in tRNA binding.

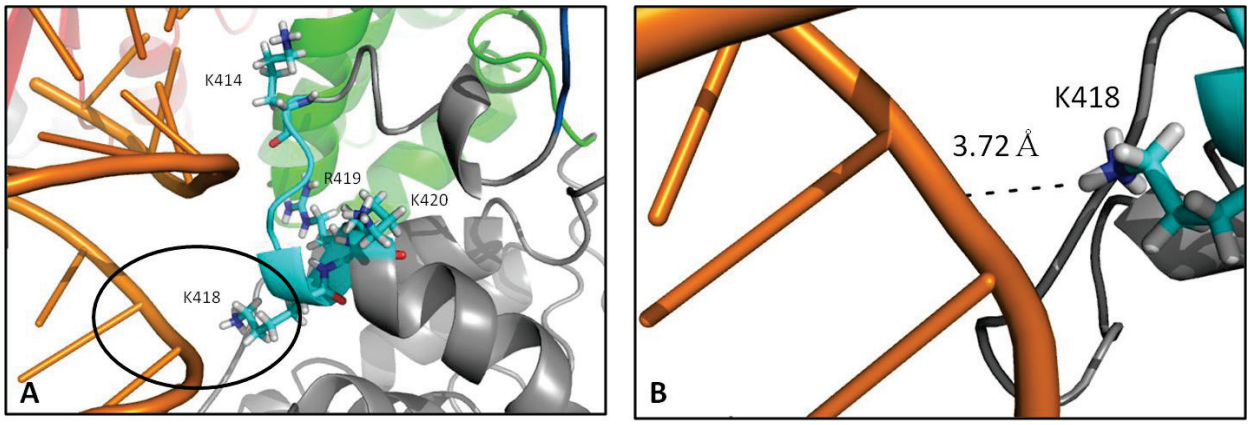


Figure 3.10: Model of *A. thaliana* tRNA nucleotidyltransferase depicting the amino acids in Mat-Mu1B
(A) Amino acids 414-420 found in Mat-Mu1B (blue) are represented as sticks and are near the tRNA substrate (orange). (B) A close up of a potential interaction between K418 and the sugar phosphate backbone of the tRNA substrate. The distance measured between K418 and the sugar phosphate backbone is shown in black

It would be interesting to define the role of the lysine at position 420. When it is converted to glutamate the K_M for tRNA increases by approximately four-fold but the turnover number increases by approximately 50% (Table. 3.2). As lysine 420 is positioned on the opposite face of the helix to lysine 418, one could imagine that a modification at position 420 could reorient K418 thereby altering binding of the tRNA substrate as reflected by the four-fold increase in K_M. However, if a glutamate at position 418 pushes the tRNA substrate away from the active site such that the k_{cat} is reduced, perhaps the K420E change moves the helix containing K418 such that the tRNA when bound is oriented more optimally in the active site thus increasing the reaction rate. This change does not alter any of the physicochemical properties of the protein so it is unclear as to how it increases catalytic activity. It would be of interest to see what affect the K418EK420E variant would have on activity.

Although Mat-Mu1B residues seem to be involved in substrate binding, we cannot exclude that it may have a role in targeting that is related to substrate binding. For example in yeast, overproduction

of tRNA nucleotidyltransferase can suppress a defect in Los1p, the exportin used to transport tRNA from the nucleus to the cytosol, indicating that tRNA nucleotidyltransferase with tRNA bound can be efficiently exported from the nucleus (Feng and Hopper, 2002). When methotrexate is bound to dihydrofolate reductase, it folds into a conformation that makes it difficult to be imported into mitochondria (Vestweber and Schatz, 1988). Perhaps tRNA binding to tRNA nucleotidyltransferase makes the protein more (or less) import competent. Although this seems unlikely, this possibility could be addressed by *in vitro* import experiments in the presence and absence of added tRNA. It might also be possible to compare the protease sensitivity of the native and Mat-Mu1B variant tRNA nucleotidyltransferases in the presence and absence of tRNA.

If the protein import experiments to be carried out in plants show that targeting is altered in proteins bearing Mat-Mu1B alone or in proteins with only the specific single amino acid substitutions, *e.g*., K418E, K420E, then this will suggest that this region is involved in targeting and in substrate binding. To determine if substrate binding can be separated from targeting, additional amino acid substitutions could be introduced in this region as a means to distinguish between the two processes.

3.7 Conclusion

These data suggest that the previously identified Mat-Mu1 sequence can be divided into two parts, Mat-Mu1A and Mat-Mu1B, which play roles in protein targeting and catalysis, respectively. Altering Mat-Mu1A causes a change in protein conformation such that it is imported less well into mitochondria (or more efficiently into plastids) while changes in Mat-Mu1B alter tRNA binding and catalysis. In fact, a single amino acid substitution K418E results in a 60-fold drop in the specificity constant suggesting an important role for this amino acid in enzyme activity. This provides some of the most detailed analysis of a region of tRNA nucleotidyltransferase where tRNA binds. Protein import experiments are required to confirm the hypothesis with respect to targeting.

Chapter 4: The ability of an arginine to tryptophan substitution in *Saccharomyces cerevisiae* tRNA nucleotidyltransferase to alleviate a temperature-sensitive phenotype suggests a role for motif C in active site organization

4.1 Preface

The work presented in Chapter 4 was publish in: Goring M. E., Leibovitch M., Gea-Mallorqui E., Karls S., Richard F., Hanic-Joyce P.J., and Joyce P.B.M. (2013). **The ability of an arginine to tryptophan substitution in *Saccharomyces cerevisiae* tRNA nucleotidyltransferase to alleviate a temperature-sensitive phenotype suggests a role for motif C in active site organization**. Biochim Biophys Acta. 1834(10):2097-2106. Paul Joyce designed the experiments and drafted the manuscript. Mark Goring and Matthew Leibovitch performed the biophysical and biochemical characterization, Pamela Hanic-Joyce performed the yeast studies, Shawn Karls and Francis Richard cloned and sequenced *CCA1* genes from suppressor strains and Ester Gea-Mallorqui showed R64W suppresses the E189F temperature-sensitive phenotype. All authors contributed during manuscript preparation and approved the content.

4.2 Abstract of Manuscript

We report that the temperature-sensitive (ts) phenotype in *Saccharomyces cerevisiae* associated with a variant tRNA nucleotidyltransferase containing an amino acid substitution at position 189 results from a reduced ability to incorporate AMP and CMP into tRNAs. We show that this defect can be compensated for by a second-site suppressor converting residue arginine 64 to tryptophan. The R64W substitution does not alter the structure or thermal stability of the enzyme dramatically but restores catalytic activity *in vitro* and suppresses the ts phenotype *in vivo*. R64 is found in motif A known to be involved in catalysis and nucleotide triphosphate binding while E189 lies within motif C previously thought only to connect the head and neck domains of the protein. Although mutagenesis experiments indicate that residues R64 and E189 do not interact directly, our data suggest a critical role for residue

E189 in enzyme structure and function. Both R64 and E189 may contribute to the organization of the catalytic domain of the enzyme. These results, along with overexpression and deletion analyses, show that the ts phenotype of *cca1-E189F* does not arise from thermal instability of the variant tRNA nucleotidyltransferase but instead from the inability of a partially active enzyme to support growth only at higher temperatures.

4.3 Introduction

All mature transfer RNA (tRNA) molecules contain an invariant cytidine-cytidine-adenosine (CCA) sequence at their 3'-ends that is required for amino acid attachment and protein synthesis. In many organisms this CCA sequence is not encoded by the tRNA genes so addition of CMP and AMP residues is catalyzed by the enzyme ATP(CTP):tRNA nucleotidyltransferase (EC 2.7.7.25). The CCA-adding enzyme has been the subject of considerable study because the stepwise addition of specific nucleotides occurs without a nucleic acid template (Betat *et al.*, 2010, Vörtler and Mörl, 2010). CCA-adding enzymes belong to the nucleotidyltransferase (NT) superfamily (Holm and Sander, 1995, Aravind and Koonin, 1999) and share a similar structure with four domains (head, neck, body and tail indicated in Fig. 4.1A) of similar dimensions (Li *et al.*, 2002, Augustin *et al.*, 2003, Okabe *et al.*, 2003, Xiong *et al.*, 2003) and a conserved active site signature (Holm and Sander, 1995). Based on sequence similarity, the nucleotidyltransferase superfamily can be divided further into two classes (Yue *et al.*, 1996). Class I enzymes (including the archaeal CCA-adding enzymes, eukaryotic poly (A) polymerases, DNA polymerase β and kanamycin nucleotidyltransferase) do not show significant overall similarity to each other outside of the shared active site signature. In contrast, class II enzymes (representing eubacterial and eukaryotic CCA-adding enzymes) share a conserved N-terminal region containing several conserved motifs (labeled A to E, Fig. 4.1B, C) with, in many cases, defined roles in nucleotide specificity and catalysis (Li *et al.*, 2002).

Figure 1A

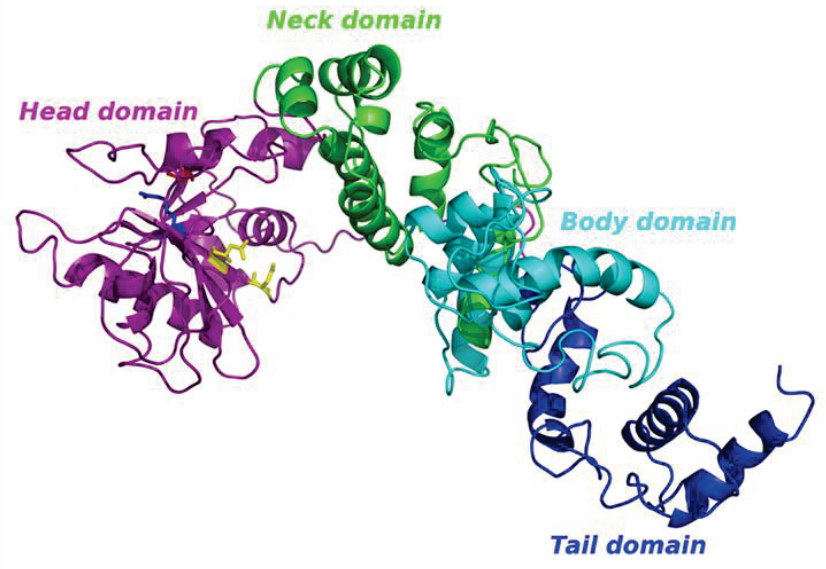


Figure 1B

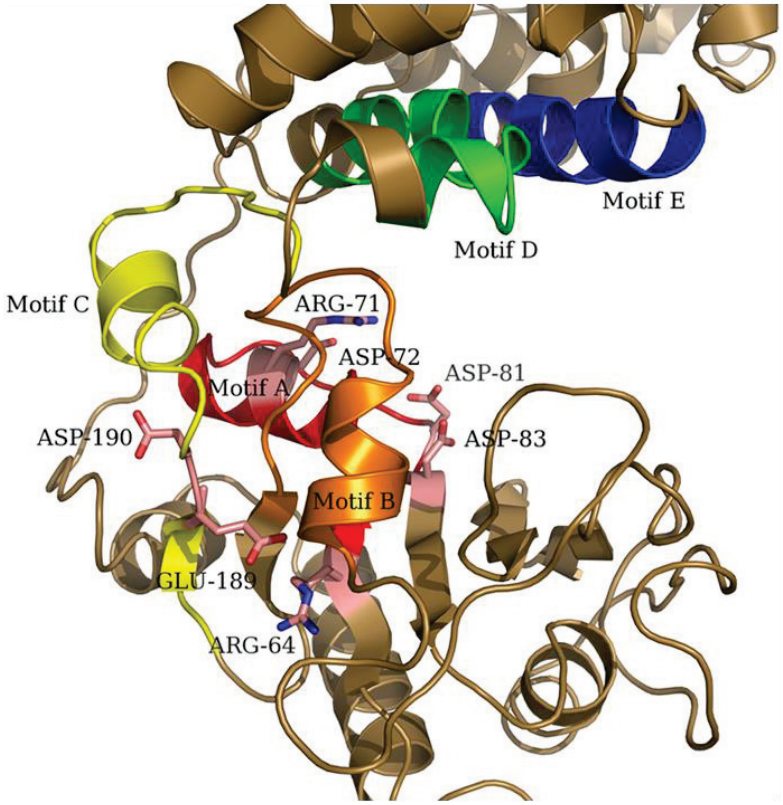


Figure 1 C

Figure 4.1. **Structure and sequence conservation of tRNA nucleotidyltransferase.** Model of the overall structure of *Saccharomyces cerevisiae* tRNA nucleotidyltransferase showing the conserved head (purple), neck (green), body (cyano), and tail (sand) domains. E189 (red), R64 (blue), and the catalytic residues; Asp 81 (yellow) and Asp 83 (yellow), are also shown. Phyre2 intensive modelling was used to generate the model based on the known crystal structure of the *Bacillus stearothermophilus* tRNA nucleotidyltransferase (Li *et al.*, 2002). (B) The head and neck regions of *Saccharomyces cerevisiae* tRNA nucleotidyltransferase modeled on the *Bacillus stearothermophilus* structure showing the positions of Arg 64 and Glu 189 relative to conserved motifs A (red), B (orange), C (yellow), D (green), and E (blue). The rest of the enzyme is colored sand. Key residues are also shown including the catalytic aspartic acids 81 and 83. Arg 71 and Asp 72 are part of the conserved GGxVRD sequence involved in binding the triphosphate motif of NTP substrates (Li *et al.*, 2002). Phyre2 intensive modelling was used to generate the model. (C) ClustalW (Torres *et al.*, 2010) alignment of the head and neck regions of tRNA nucleotidyltransferases. The yeast enzyme and four other class II tRNA nucleotidyltransferases (*Aquifex aeolicus (Tomita et al., 2004)*, *Thermotoga maritime* (Toh *et al., 2009)*, *Bacillus stearothermophilus* (Li *et al., 2002)*, human (Augustin *et al.*, 2003)) for which crystal structures have been solved are shown. Motifs A-E first identified in the *Bacillus stearothermophilus* enzyme (Li *et al.*, 2002) are boxed and labeled. Residues R64 and E189 in the yeast enzyme are in bold. (*) amino acid identity, (:) strongly conserved amino acid, (.) weakly conserved amino acid.

Examination of solved crystal structures (Augustin *et al.*, 2003) and co-crystal structures of class II tRNA nucleotidyltransferases complexed with one (Li *et al.*, 2002, Toh *et al.*, 2009) or more (Tomita *et al.*, 2004, Tomita *et al.*, 2006) substrates revealed roles for residues in these conserved motifs. Motif A (found in all nucleotidyltransferases) includes a highly conserved GGxVRD sequence (where x is any amino acid) and the two metal-binding carboxylates, DxD, involved in catalysis and the binding of incoming nucleotide triphosphates (Li *et al.*, 2002, Betat *et al.*, 2010, Steitz, 1998). This conserved motif is found as part of a consensus five-stranded antiparallel β-sheet flanked by two α-helices common to all tRNA nucleotidyltransferases (Li *et al.*, 2002, Augustin *et al.*, 2003, Okabe *et al.*, 2003). Motif B plays a critical role in ribose recognition with the conserved RRD sequence recognizing the 2' hydroxyl group of the ribose in each incoming ribonucleotide triphosphate (Li *et al.*, 2002, Cho *et al.*, 2007). Motif A and motif B are separated by a sequence which may contain a highly flexible loop that is important in the addition of the terminal A residue but not in the addition of the two C residues (Just *et al.*, 2008, Neuenfeldt *et al.*, 2008). Motif C, found at the C-terminal end of the head domain, shows little sequence conservation and is thought to act only as a connection between the head and neck domains of the protein (Li *et al.*, 2002). Motif D in the neck domain, contains the conserved EDxxR sequence which acts as an amino-acid based template for CCA addition by forming the nucleotide binding pocket for each incoming CTP or ATP, while discriminating against GTP and UTP through the generation of specific Watson-Crick-like hydrogen bonds with the incoming nucleotides (Li *et al.*, 2002, Cho *et al.*, 2007, Xiong and Steitz, 2006, Lizano *et al.*, 2008). Motif E may help to stabilize a helix-turn structure in motif D (Li *et al.*, 2002) and/or interact with the growing 3' end of the tRNA substrate (Cho *et al.*, 2007). In contrast to the conserved N-terminal region, sequences of the C-terminal portion of Class II enzymes bear little sequence conservation and are thought to be involved in RNA substrate selection through interactions with the T stem and loop (Tomita *et al.*, 2004, Betat *et al.*, 2004). As first demonstrated in *Aquifex aeolicus* (Tomita and Weiner, 2001) certain bacteria accomplish CCA addition by using two tRNA

nucleotidyltransferases with partial activity, a CC-adding enzyme and an A-adding enzyme. Recently, Tretbar *et al.*, (2011) demonstrated that a deletion in the C-terminal region of an A-adding enzyme conferred CCA-adding activity and proposed that the C-terminal region acts to physically constrain and position the tRNA substrate.

In *Saccharomyces cerevisiae* there is a single gene, *CCA1*, coding for tRNA nucleotidyltransferase. The essential role for this gene was demonstrated by Aebi *et al.*, (1990) who isolated a temperature-sensitive (ts) allele, *ts352*. When cells bearing this mutation are shifted to the non-permissive temperature of 37^{o}C there is an abrupt cessation of protein synthesis attributed to a reduction in the pool of functional tRNAs (Peltz *et al.*, 1992). We (Shan *et al.*, 2008) previously determined that the *ts352* mutation results from a glutamate to lysine substitution at position 189 (E189K), which is located in motif C (Fig. 4.1). Based on the available crystal structures of other class II tRNA nucleotidyltransferases, E189 is predicted to be near the end of a β-strand such that it could interact with the adjacent α-helix to support a β-turn near the junction of the head and neck domains. We also studied the effect of the E189K substitution and showed reduced activity at both the permissive and non-permissive temperatures *in vitro* and a 6^{o}C reduction in melting temperature as measured by circular dichroism spectroscopy (Shan *et al.*, 2008). Studies of additional E189 variants (E189F, E189Q, E189H) revealed a range of structural and catalytic effects *in vitro* and a variety of growth phenotypes. The E189F variant was similar to E189K and showed reduced activity (approximately 25 fold less than native enzyme) and thermal stability and both mutants exhibited a temperature-sensitive phenotype. The E189Q variant showed a 2.5-fold drop in activity, no change in thermal stability *in vitro* and mutants exhibited a 75% reduction in growth rate at 37^{o}C. The most interesting variant, E189H, exhibited no detectible change in thermal stability but showed a 13-fold reduction in activity and conferred a temperature-sensitive phenotype. This result did not support the classical notion that the temperature-sensitive phenotype was caused by thermolability of the gene product (Tan *et al.*, 2009, Poultney *et al.*,

2011). Taken together these results suggested that the size of the residue at position 189 was important to enzyme activity and that larger residues were deleterious. Therefore, we (Shan *et al.*, 2008) proposed that the E189F substitution altered the organization of the head and neck domains generally by disrupting (through lost hydrogen-bonding potential or through steric hindrance) a turn connecting a β-strand and an α-helix in the region of the yeast enzyme corresponding to motif C in the *Bacillus stearothermophilus* enzyme (Li *et al.*, 2002). Motif C is a region with little sequence conservation and is thought to provide only a linking role in the organization of the head and neck regions of tRNA nucleotidyltransferase. However, given the important catalytic residues in the head and neck regions of the protein it would not be surprising if altering the local arrangement of this region could affect enzyme activity. How the E189F substitution affected interactions with any or all of the tRNA nucleotidyltransferase substrates, *i.e.*, tRNA, CTP, or ATP remains largely unexplored. Moreover, the wide range of effects conferred by E189 substitutions did not distinguish between whether the temperature-sensitive phenotypes were caused by reduced thermal stability of the variant proteins or by partial defects in enzyme activity which only affected growth at higher temperatures.

To further address the role of glutamate 189 in yeast tRNA nucleotidyltransferase structure and function and to investigate whether defects in thermal stability or activity or both were responsible for the temperature-sensitive phenotype, we used UV mutagenesis to generate an intragenic suppressor (*cca1-E189F,R64W*) of the rarely reverting *cca1-E189F* allele. By comparing the sequence of the yeast enzyme to that of the *Bacillus stearothermophilus* protein for which a crystal structure is available (Li *et al.*, 2002), we suggest that R64 of the yeast enzyme is found in a position corresponding to the end of a β-strand at the start of motif A (Fig. 4.1B, C). While the R64W substitution alone does not affect enzyme activity or stability, it restores the E189F protein to an active conformation without restoring thermal stability, suggesting an important structural association between residue 64 in motif A and residue E189 in motif C. The protein encoded by the *cca1-E189F,R64W* allele retains the reduced melting temperature

exhibited by the *cca1-E189F* variant but shows increased activity at the permissive temperature suggesting that it is the reduced activity of the *cca1-E189F* variant and not its altered stability that confers temperature-sensitivity. We also show that the temperature-sensitive phenotype can be suppressed *in vivo* by overexpressing the *cca1-E189F* variant from a heterologous promoter and that diploids bearing the *cca1-E189K* mutation in *trans* to a *cca1* null allele exhibit more extreme heat sensitivity than *cca1-E189K*/*cca1-E189K* diploids. We conclude that although the *cca1-E189F* mutation produces a classical temperature-sensitive phenotype, this mutation is actually hypomorphic and renders the gene product partially inactive at all temperatures. This reduction in enzyme activity, however, is deleterious only at the high temperature.

4.4 Materials and Methods

4.4.1 Strains

Yeast strains 352-1A, E189F, NT33-5, YPH500, BY4743 are as described (Aebi *et al.*, 1990, Shan *et al.*, 2008, Shanmugam *et al.*, 1996, Brachmann *et al.*, 1998), respectively. Strain SCDT8 (*MATα cca1::TRP1 leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*) is a derivative of W303-1B (R. Rothstein) and carries a pRS316 derivative containing *CCA1* and *URA3* genes. Yeast strain NT1-34 (*MATa ura3 leu2-3,112 lys2 his3 ade2 cca1-E189K*) was a product of a cross between W303-1A (R. Rothstein) and 352-1A. Diploid yeast strains (relevant genotypes *cca1-E189K*/*cca1-E189K* and *cca1-E189K*/*cca1::TRP1* are the results of matings between haploid strains (NT1-34 with NT33-5 and NT1-34 with SCDT8, respectively) with selection of diploids on SC medium minus Lys and Trp. Diploidy was confirmed by formation of asci on sporulation medium. *E.coli* strains XL2-Blue and BL21 (DE3) were purchased from Stratagene and used for plasmid construction and protein expression, respectively.

4.4.2 Plasmid Construction

Standard molecular biology techniques were employed throughout (Ausubel, 1989). The open reading frame (ORF) of the *CCA1* gene was amplified by polymerase chain reaction using primers MUTCCA5-4 and MUTCCA3 (Table 4.1) and digested with *Bam*HI and *Sal*I for insertion into a modified pGEX-2T (GE Healthcare) vector (Shan *et al.*, 2008) for expression in *E. coli*. Plasmid A163-2 was constructed by transferring the *Sac*I-*Kpn*I expression cassette of p415-ADH (Mumberg *et al.*, 1995) into *Sac*I-*Kpn*I digested pRS316 (Sikorski and Hieter, 1989). Plasmid SCR9-1 containing the 5'-noncoding region and ORF of *CCA1* fused to the CYC terminator was constructed as follows: oligonucleotides SCCAPR5 and SCCAAM3 (Table 4.1) and *Pfu* polymerase were used to amplify the target region from the genomic DNA of strain BY4743. The PCR product was digested with *Sac*I and *Sal*I and inserted into similarly digested pA163-2. For expression of Cca1p from the *TDH3* or *ADH1* promoters, oligonucleotides SCCCA1GFPATG3 and SCCAAM3 (Table 4.1) were used to amplify the *CCA1* open reading frame without the mitochondrial targeting signal from template SCR9-1. The PCR product was digest with *Bam*HI and *Sal*I and ligated to similarly digested p426GPD (Mumberg *et al.*, 1995) or A163-2. Plasmid SCR9-1 was used as template in QuikChange™ (Stratagene) site-directed mutagenesis reactions utilizing *Pfu* polymerase. Oligonucleotide pairs are listed in Table 4.1. Oligonucleotides were purchased from Biocorp Inc. (Montreal, QC). All plasmid constructs were confirmed by sequence analysis (McGill University and Génome Québec Innovation Centre).

4.4.3 UV mutagenesis

Strain E189F (Shan *et al.*, 2008) was chosen for UV mutagenesis as it is less likely to revert than strain E189K (with the sequence AAA at codon 189). Moreover, any single point mutation at the phenylalanine codon 189 (TTT) is likely (seven out of nine possible substitutions) to result in the introduction of another large hydrophobic amino acid, which, based on our previously-proposed size-

constraint model (Shan *et al.*, 2008), would result in loss of enzyme activity and thereby further decrease the possibility of isolating revertants. UV mutagenesis was a modification of(Lawrence, (1991). In brief, cells were exposed to 50 J/m^2 UV light (254 nm) in 0.9% phosphate buffered saline with a Stratalinker® UV Crosslinker Model 1800 (Stratagene), pelleted by centrifugation, resuspended in 100 µl phosphate buffered saline, and spread onto YPD plates. All manipulations were carried out under dim red light. The plates then were incubated in the dark at 37°C. Temperature-resistant colonies were patched onto YPD, incubated at room temperature, and replica plated to 37°C to confirm loss of the temperature-sensitive phenotype.

4.4.4 Growth assays

Transformation of yeast strain E189F (Shan *et al.*, 2008) was by the method of Schiestl and Gietz, (1989) with selection on synthetic complete medium lacking uracil (Sherman, 2002) at 22°C. Two independent transformants were replica-plated to the same medium and incubated at 22°C or 37°C to assess temperature sensitivity.

Table 4.1. Oligonucleotides used in this study

Oligonucleotide	Sequence
SQCE189F5	CATAAAGGTGAAGTGTTTGATTTCACTAAGAGAGGTCTGCAAG
SQCE189F3	CTTGCAGACCTCTCTTAGTGAAATCAAACACTTCACCTTTATG
SQCE189K5	CATTCATAAAGGTGAAGTTAAAGATTTCACTAAGAGAGGTCTGCAAG
SQCE189K3	CTTGCAGACCTCTCTTAGTGAAATCTTTAACTTCACCTTTATGAATG
R64W5	GTACCACAATAAGCCTGAACCATTGACTCTTTGGATCACGGGCGG
R64W3	CCGCCCGTGATCCAAAGAGTCAATGGTTCAGGCTTATTGTGGTAC
R64E5	CCTGAGCCATTGACTCTCGAGATCACGGGCGGATGGG
R64E3	CCCATCCGCCCGTGATCTCGAGAGTCAATGGCTCAGG
SCQCE189R5	CATTCATAAAGGTGAAGTACGAGATTTCACCAAGAGAGG
SCQCE189R3	CCTCTCTTGGTGAAATCTCGTACTTCACCTTTATGAATG
SCCAPR5	ATACCCGGGAGCTCCAGCAGATCTATAAAATAAGTGG
SCCAAM3	CCATCGATGTCGACGCTACTACAGGTATTTTGGTAGTATG
SCCCA1GFPATG3	GCTCAGGGATCCATGACGAATTCTAATTTTG
MUTCCA5-4	ACTAGTGGATCCATGACGAA
MUTCCA3	ATCGATGTCGACTACTACAG

4.4.5 Heterologous protein expression and purification

The procedure was modified from that of Shan *et al.* (2008) with minor modifications. Protein expression was induced in 1.3 l of *E. coli* BL21 (DE3) cells by adding IPTG to a final concentration of 0.5 mM or 1 mM with growth at 18°C for 16-20 hours. After lysis and centrifugation, the cleared lysate was loaded onto a Glutathione Fast Flow 4B (GE Healthcare) column and the GST-tagged protein eluted using elution/cleavage buffer (50 mM Tris-HCl, 140 mM NaCl, 15 mM glutathione, 2.5 mM CaCl2, pH 8.0). Cleavage by thrombin (Amersham Biosciences, 1 unit/g protein) was carried out overnight at 4°C with dialysis against Tris-HCl/NaCl buffer (50 mM Tris-HCl, 140 mM NaCl, pH 8.0). After cleavage had been verified, additional dialyses were used to remove the glutathione and GST tag and to change the buffer to PBS (pH 7.3). The samples then were loaded again onto the Glutathione Fast Flow 4B column to remove any remaining GST tag and glycerol was added to the pooled proteins to a final concentration of 10% (v/v) before storage at -80°C.

4.4.6 Enzyme Assays

Plasmids G73 and pmBSDCCA were constructed by Cho *et al.* (2002) and Oh and Pace (1994) respectively and generously provided by Dr. Alan Weiner (University of Washington). By digesting with the appropriate restriction enzymes, these plasmids generated templates ending in the discriminator base (N), the antepenultimate base (C), the penultimate base (C) or the 3' terminal base (A) of a tRNA with a complete CCA sequence. Each run-off transcription reaction (100 µl) contained 20 µl of 5x transcription buffer (Fermentas), 5 µl of 10 mM UTP, 5 µl of 10 mM CTP, 5 µl of 10 mM ATP, 5 µl of 1 mM GTP, 50 µCi [α^{32}P] GTP (10 µCi/µl, 3000 Ci/mmol, Perkin Elmer), at least 1 µg of linearized DNA template, 60 units of T7 RNA polymerase (Fermentas) and RNase-free water. After incubation at 37°C for 3 hours, each reaction was phenol extracted, ethanol precipitated and resuspended in nuclease-free water for purification after electrophoresis and exercised from a 7M urea, 20% acrylamide gel. Enzyme

assays were modified from Cudny *et al.*, (1978) with each 10 µl reaction mixture containing 100 mM glycine buffer (pH 9), 10 mM $MgCl_2$, 1 mM ATP, 0.4 mM CTP, 1 µl of the appropriate [α-^{32}P] GTP-labelled transcript after gel extraction, and the quantity of enzyme indicated. Standard assays were performed for 2 minutes at 22^{o}C and 37 ^{o}C. Reactions were terminated by adding 10 µl of 2X loading buffer containing 7M urea (Peattie 1979) with heating at 70^{o}C for 10 min and the reaction products characterized after electrophoresis on 7M urea/20% polyacrylamide gels with detection by autoradiography or using the Typhoon™ TRIO Variable Mode Imager (GE Healthcare).

4.4.7 Circular Dichroism Spectroscopy

Circular dichroism studies were performed on all proteins, in the 200-280 nm far-UV region, at 22°C using a JASCO 815 spectropolarimeter. A Jasco-815 circular dichroism spectrophotometer along with the Variable Temperature program (Jasco) was used to measure CD signals. At standard sensitivity, a bandwidth of 1 nm was used, along with a response time of 0.25 seconds, data pitch of 0.2 nm and a 0.2 cm cell. Temperature was regulated using a Pelletier water bath accessory while nitrogen was set to flow through the instrument at a constant rate of 3 l/min. Aliquots of 0.5 ml at 0.85 µM (0.061 mg/ml) of protein were used for the measurements. For scans a scan speed of 20 nm/min was used. Five accumulations were collected. The 1 x PBS buffer (pH 7.3) was scanned under identical conditions in order to establish the base line and was subtracted from each spectrum in order to reveal the protein spectrum alone. For thermal experiments a temperature slope of 30°C/hour was used. The cell was stoppered to limit evaporation. The data obtained were smoothed using the smoothing function in the Spectra Analysis (Jasco) program.

4.5 Results

4.5.1 Identification of a mutation in the *CCA1* gene that suppresses the temperature-sensitive phenotype

To isolate suppressors of the temperature-sensitive *cca1-E189F* mutation, cells were UV irradiated (54-60% cell survival) to generate cells which could grow at the non-permissive temperature (37^{o}C). Of the cells that grew at the non-permissive temperature (0.0003%), the *CCA1* open reading frame was amplified by polymerase chain reaction and the products were sequenced. Multiple clones (approximately 25%) showed, alone or in combination with another mutation, the same CGG to TGG transition resulting in an arginine to tryptophan substitution at position 64 in the gene product. To confirm that this single mutation was sufficient to suppress the temperature-sensitive phenotype, this nucleotide change was introduced by site-directed mutagenesis into the *cca1-E189F* allele on plasmid SCR9-1. Introduction of the resulting plasmid into yeast strain E189F (Fig. 4.2) conferred growth at the non-permissive temperature.

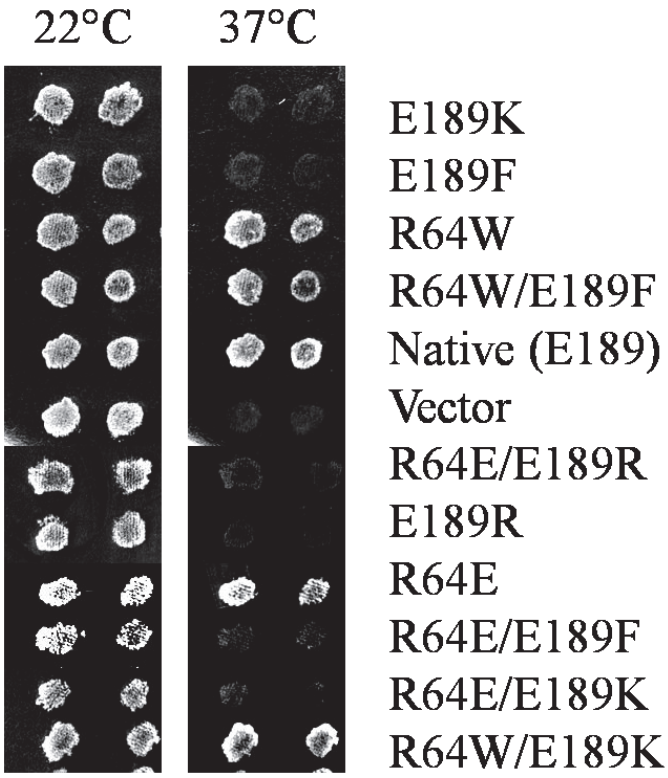


Figure 4.2. Growth of yeast expressing native and variant tRNA nucleotidyltransferases. Yeast strain E189F was transformed with plasmids coding for the tRNA nucleotidyltransferase proteins indicated and two independent transformants were replica-plated to SC medium lacking uracil and incubated at 22°C or 37°C to assess temperature sensitivity.

4.5.2 Effect of the R64W substitution on tRNA nucleotidyltransferase activity

Native and variant proteins were produced in *E. coli* as GST fusion proteins, the GST-tags removed and the proteins purified to near homogeneity (Fig. S4.1). As the *cca1-E189F,R64W* mutant was able to grow at the non-permissive temperature, we expected to see an increased level of enzyme activity as compared to the *cca1-E189F* mutant. Indeed, when provided with a template lacking the CCA sequence, the *cca1-E189F,R64W* mutant showed enzyme activity near native levels at both the permissive and non-permissive temperatures with reduced efficiency in incorporation of AMP at the third position (Fig. 4.3A). In contrast, the *cca1-E189F* variant showed reduced activity with little or no full extension at both the permissive and non-permissive temperatures within 30 min incubation period supporting our previous results (Shan *et al.*, 2008). We also found that the *cca1-R64W* mutation alone

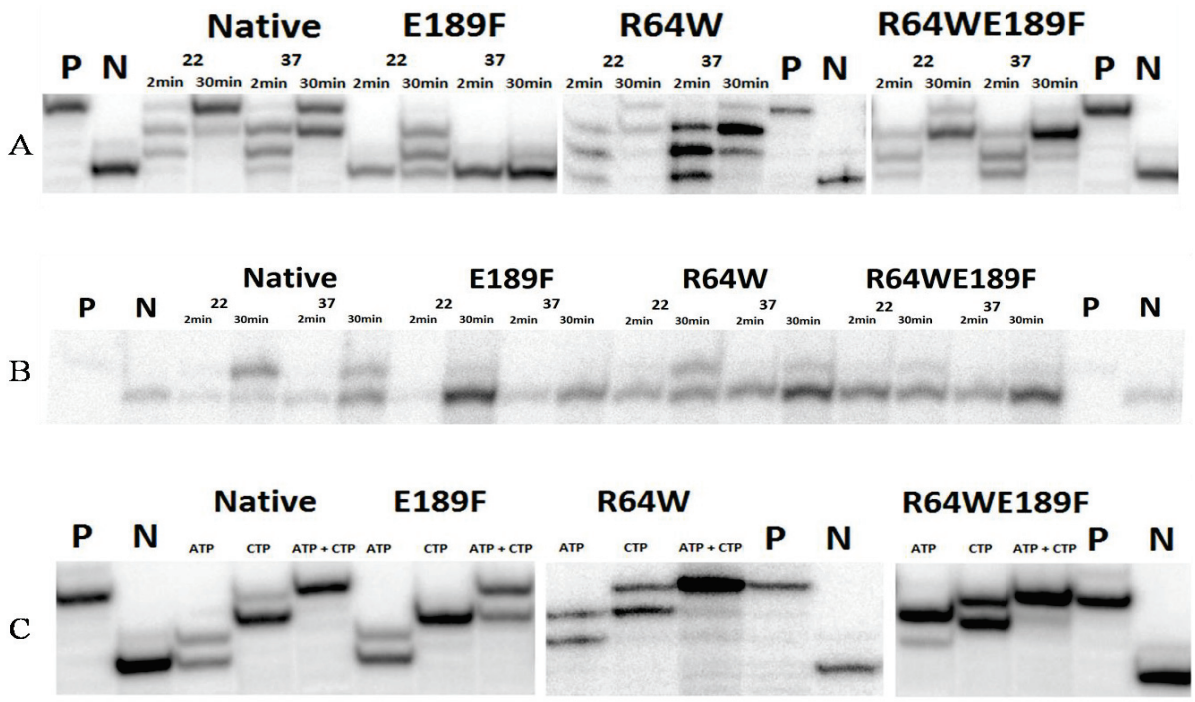


Figure 4.3. Nucleotide addition by tRNA nucleotidyltransferase variants. (*A*) The enzymes indicated (5 ng) were incubated with a tRNA template ending at the discriminator base for 2 or 30 min at the temperatures shown and the products separated by polyacrylamide gel electrophoresis. P, positive control (active enzyme generating a product containing a complete CCA sequence), N, negative control (boiled enzyme leaving a product lacking the CCA sequence). (B) The enzymes (2 ng) indicated were incubated for 2 or 30 min at the temperatures shown and the products separated by polyacrylamide gel electrophoresis. Controls as in panel A. (C) The reactions were carried out to completion at 22°C with the tRNA template ending at the discriminator base and the nucleotides indicated and the products separated by polyacrylamide gel electrophoresis. Controls as in panel A.

did not eliminate enzyme activity indicating that an arginine to tryptophan substitution at position 64 does not dramatically affect enzyme activity. These data also confirm the observation that, with increased exposure to 37°C, the activity of the native enzyme is reduced *in vitro* (Shan *et al.*, 2008). To address differences in AMP incorporation observed at the third position we repeated the reactions using a template lacking only the terminal adenosine (Fig. 4.3B). Supporting what we already had observed, we noted that at the permissive temperature the E189F variant added AMP inefficiently, while the R64W and R64WE189F variants showed an intermediate level of ATP incorporation as compared to the native enzyme (Fig. 4.3B). When the reactions were performed at 37°C the native enzyme added AMP efficiently, while the E189F variant again was unable to add to the third position and the R64W and

R64WE189F variants showed an intermediate level of incorporation (Fig. 4.3B). These results show that while the E189F variant has little or no activity, the R64WE189F suppressor has restored enzyme activity to a level similar to that of the R64W variant and intermediate between that of the native enzyme and the E189F variant. Furthermore, the R64W change also restores activity to near native levels at the non-permissive temperature for the E189K variant (data not shown).

To detect differences in the ability to discriminate between ATP and CTP, we carried out end-point analyses at 22^{o}C to measure the incorporation of either AMP or CMP (Fig. 4.3C). As was expected, the native enzyme showed low levels of misincorporation of AMP (at the first position) or CMP (at the third position). The E189F variant showed a low level of AMP incorporation at the first position and little or no addition of either AMP or CMP at the third position reflecting this enzyme's reduced activity. Intriguing results were obtained for the R64W and R64WE189F variants indicating that AMP is misincorporated at both the first and the second positions and CMP is misincorporated at the third position more frequently than for the native enzyme (Fig. 4.3C).

4.5.3 Effects of the R64W substitution on tRNA nucleotidyltransferase structure and thermal stability

To determine what effect the R64W suppressor mutation had on enzyme structure and thermal stability we used circular dichroism and fluorescence spectroscopy. The circular dichroism spectra were similar for all proteins at 22^{o}C (Fig. 4.4) and showed the characteristic peak minima at 208 nm and 222 nm expected for proteins with α-helical character as seen previously for E189F and other E189 variants (Shan *et al.*, 2008). These data were supported by fluorescence experiments where the wavelength of fluorescence intensity maximum at 22^{o}C differed by less than 3 nm between any of the proteins (data not shown). The thermal stability of the native and variant enzymes was monitored using circular dichroism spectroscopy by following the peak minima at 222 nm (Fig. 4.5).

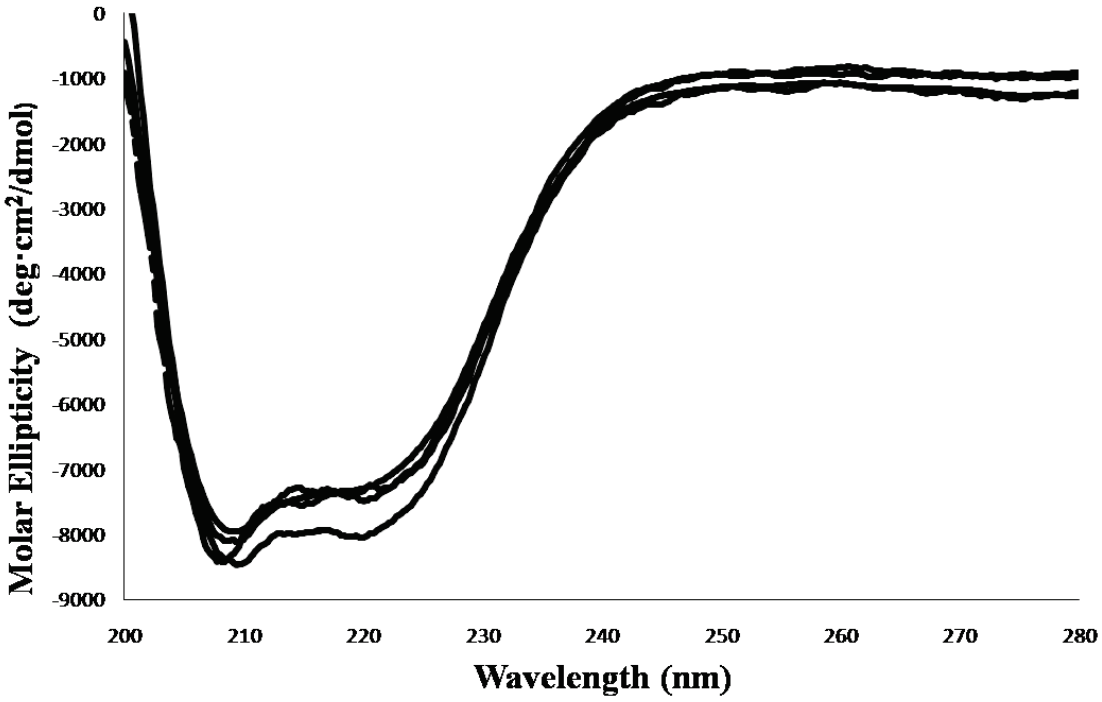


Figure 4.4. **Circular dichroism spectra for native and variant tRNA nucleotidyltransferases**. All spectra were recorded at 22 °C with proteins at ~1 µM. native -·-·-**.,** R64W ·····, R64WE189F **- - -**, E189F ——.

Using either melting temperature (Tm) or percent remaining α-helicity at 37°C, the values for the R64WE189F variant were in good agreement with the E189F variant while the R64W variant showed values more similar to those of the native enzyme (Fig. 4.5, Table 4.2). Although the changes in Tm are small (maximum of 4.5°C), the differences are reproducible (standard deviation ± 1°C or less in each case) with the native enzyme and the R64W variant being more similar to each other (Tm of ~40-42°C), while E189F and R64WE189F are more similar to each other and show a Tm of ~38°C and reduced remaining α-helicity. The results for the native enzyme and the E189F variant are in good agreement (± 1°C) with those obtained previously (Shan *et al.*, 2008).

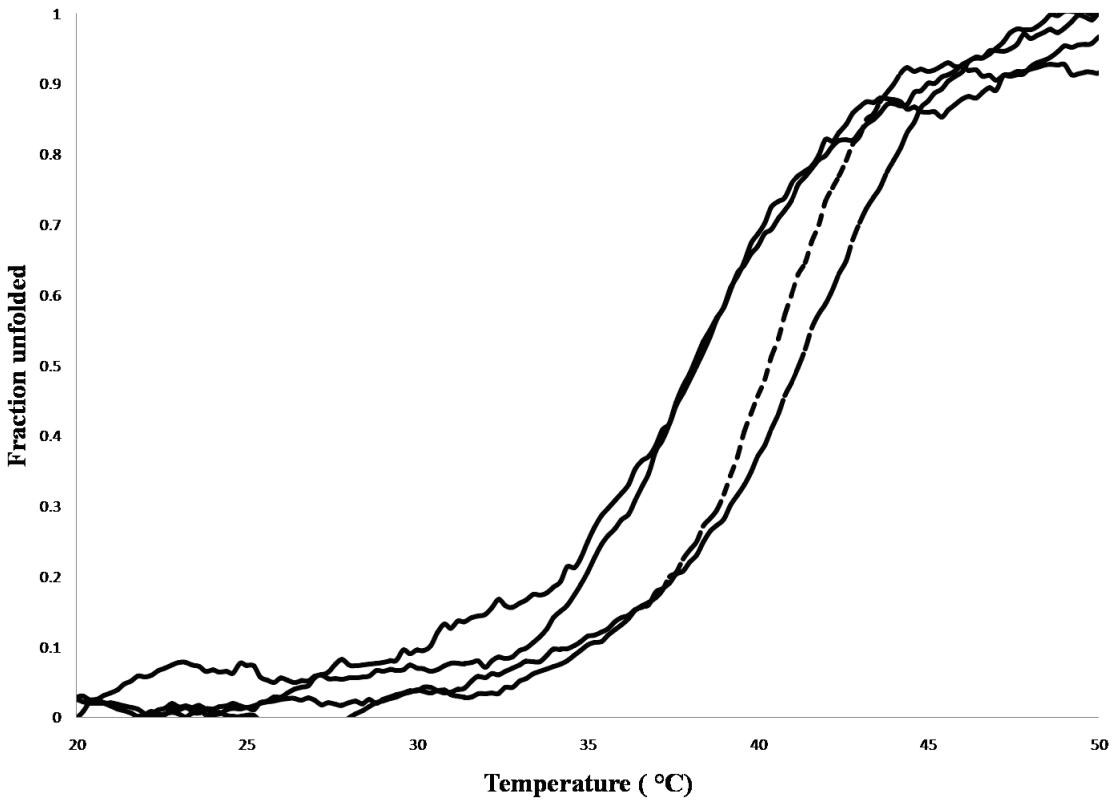


Figure 4.5. Representative thermal denaturation curves of native and variant tRNA nucleotidyltransferases. Signal intensity at 222 nm was monitored between 20°C and 50°C for proteins at ~1 µM. native -·-·-., R64W ·····, R64WE189F - - -, E189F ——. Fraction of protein unfolded was defined by the equation F = 1/ΔθT x θ - θi/ΔθT where F is the fraction of protein unfolded, θ is ellipticity measured in mdeg, and ΔθT is the total change in ellipticity between the initial and final temperatures of interest; ΔθT= θf -θi. θi is the ellipticity value at the start temperature of interest (20°C), and θf is the final ellipticity value at the end temperature of interest (50°C).

TABLE 4.2. Biophysical characteristics of native and variant tRNA nucleotidyltransferases

Protein	Mean Tm (^{o}C)	Remaining α-helicity at 37^{o}C (%)
native	41.8 ± 0.8	83.0
E189F	37.4 ± 0.9	59.7
R64W	39.5± 0.7	82.9
R64WE189F	38.6 ± 1.1	61.4

Each experiment was performed at least twice and the data represent the mean and standard deviation from the mean. Tm is defined as the temperature at which the fraction of unfolded protein is 50% and remaining α-helicity is defined as the θ_{222} at 37°C as compared to 22°C multiplied by 100%.

4.5.4 Roles of R64 and E189

To further explore the roles of positions 64 and 189, we produced several additional amino acid substitutions at these two positions and tested the abilities of these variant proteins, expressed from plasmid-borne genes, to suppress the ts phenotype of *cca1-E189F* (Fig. 4.2). As we have shown previously, transformants expressing the E189F and E189K variant proteins were unable to grow at the non-permissive temperature. In contrast, transformants expressing the R64W variant, the R64WE189F variant, or the native enzyme all grew at the non-permissive temperature. These data are consistent with the comparable levels of CCA-adding activity obtained for these variants *in vitro* (Fig. 4.3). We then reversed the residues normally found at positions 64 and 189 in the native enzyme to produce a plasmid expressing an R64EE189R variant enzyme and these transformants did not grow at the non-permissive temperature. While the R64E substitution alone allowed for growth at the non-permissive temperature, expression of the R64EE189K or R64EE189F double substitution variants yielded a temperature-sensitive phenotype. In contrast, the expression of the R64W substitution in combination with either the E189K or E189F change conferred growth at 37^{o}C. Finally, expression of the single E189R amino acid substitution variant also did not suppress the temperature-sensitive phenotype.

4.4.5 The *cca1-E189F* and *cca1-E189K* mutations are hypomorphic *in vivo*

As the *cca1-E189F* variant protein showed reduced activity at the permissive and non-permissive temperatures, and the R64W substitution restored activity *in vitro* at both temperatures without significantly affecting the overall structure or stability of the protein, we reasoned that the *cca1-E189F* mutation was hypomorphic rather than conditional. To address this hypothesis we overexpressed the mutant protein *via* heterologous promoters and showed that the temperature-sensitive phenotype of *cca1-E189F* could be suppressed (Fig. 4.6A). Therefore, increasing the amount of the partially active enzyme was sufficient to restore growth at the non-permissive temperature. We also confirmed the hypomorphic nature of the first-isolated ts allele, *cca1-E189K* (ts352), by comparing the growth of *cca1-*

E189K/*cca1-E189K* diploids to *cca1-E189K*/*cca1::TRP1* diploids. While neither of these strains grew at 37°C, the *cca1-E189K/cca1::TRP1* was uniquely sensitive to the lower temperature of 30°C (Fig. 4.6B).

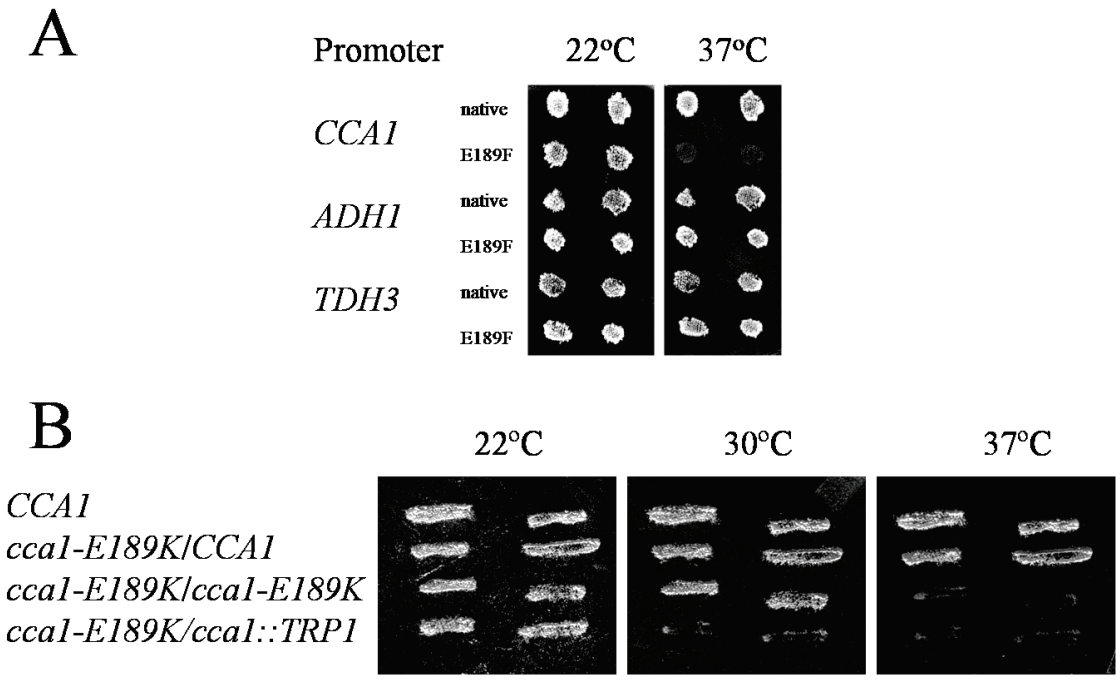


Figure 4.6. Growth of yeast expressing varying amounts of tRNA nucleotidyltransferase. (A) The native or E189F variant of tRNA nucleotidyltransferase was expressed from the promoters indicated in yeast strain E189F and incubated at 22°C or 37°C on SC medium lacking uracil to assess temperature sensitivity. (B) Diploid yeast strains (relevant genotypes shown) were replica plated to YPD medium and incubated at the temperatures indicated.

4.6 Discussions

4.6.1 The R64W substitution restores tRNA nucleotidyltransferase activity

Our previous work (Shan *et al.*, 2008) showed that substitutions at position 189 conferred a wide range of phenotypes and suggested a key role for E189 in enzyme structure and function. To further address the role of glutamate 189 in tRNA nucleotidyltransferase activity, we subjected strain E189F to UV mutagenesis and 25% of the resulting temperature-resistant isolates showed a CGG to TGG transition at codon 64 (resulting in an arginine to tryptophan substitution) in addition to the original E189F substitution. Introducing a plasmid-borne copy of this gene into the temperature-sensitive strain E189F

resulted in suppression of the temperature-sensitive phenotype (Fig. 4.2) indicating that this mutation alone defined the suppressor. To assess how the R64W substitution alone or in combination with E189F affected CCA addition *in vitro,* we tested variant proteins for the ability to add CMP and/or AMP to tRNA substrate templates. When a tRNA template ending at the discriminator base (Fig. 4.3A) was used, the native enzyme exhibited complete CCA addition at both the permissive and non-permissive temperatures. In contrast, the E189F variant showed little or no complete CCA addition at either the permissive or non-permissive temperature. This observation was in good agreement with our earlier results (Shan *et al.*, 2008), although those assays had measured nucleotide incorporation as acid-precipitable counts and therefore, did not assess specific addition at each of the 3 positions in the CCA sequence. It has been shown (Peltz *et al.,* 1992, Aebi *et al.*, 1990) that tRNAs generated at the non-permissive temperature by the *cca1-E189K* strain lack a complete 3' terminus, suggesting that it is the reduced ability to incorporate nucleotides and not misincorporation that defines the phenotype. The presence of multiple species in the E189F samples indicates either incomplete CCA addition or misincorporation of nucleotides at the antepenultimate or penultimate positions, which could block further extension. The latter possibility seems unlikely because increasing the incubation time at the permissive temperature (compare 2 min with 30 min) increases the amount of the extended products with a concomitant reduction in the intermediate species (Fig. 4.3B). The data suggest that the loss of efficiency of E189F is not linked to reduced specificity in nucleotide selection but instead is due to a lower nucleotide addition rate. When the reactions were repeated in the presence of ATP alone, some AMP was incorporated into the template at the antepenultimate position to block further extension particularly in the case of the native and E189F variant (Fig 4.3C). Intriguingly, both the R64W variant and the R64W/E189F variant were able to incorporate a second AMP (Fig. 4.3C) suggesting that the R64W substitution does have an effect on nucleotide selection. However, these data combined with our earlier observation of decreased incorporation of AMP as measured by acid-precipitable counts (Shan *et*

al., 2008) support the idea that the phenotype observed in the E189F variant reflects reduced enzyme activity and not misincorporation into the growing tRNA. Interestingly, the enzymes showing complete CCA addition also are able to incorporate CMP instead of the 3'-terminal AMP when provided with CTP alone (Fig. 4.3C). Although CMP addition may be the default addition reaction for many class II tRNA nucleotidyltransferases (Hoffmeier *et al.,* 2010), when the enzyme is given both ATP and CTP, AMP incorporation is favored at the third position (Fig. 4.3C, compare ATP + CTP lanes with CTP lanes). In any event, addition of CMP to the final position by the E189F variant is greatly reduced. These data indicate that at both the permissive and non-permissive temperatures the rate of nucleotide addition of the CCA sequence is reduced in the E189F variant. This suggests that the reduction in enzyme activity is not due to the inability of the enzyme to recognize one or the other nucleotide, but instead is due to a reduced ability to add nucleotides to the tRNA substrate. This implicates changes either in substrate binding (affinity and/or positioning) or in the organization of the catalytic residues at the active site, in the loss of activity.

Interestingly, both the R64W and R64WE189F variants show some loss of efficiency in terms of nucleotide selectivity at all positions (Fig. 4.3C). Based on the available crystal structure of the *Bacillus stearothermophilus* tRNA nucleotidyltransferase (Li *et al.*, 2002) we expect residue R64 to be in motif A which contains both the highly conserved GGxVRD sequence and the two carboxylate residues important for catalysis in nucleotidyltransferases (Li *et al.*, 2002, Hoffmeier *et al.*, 2010). Position 64 is expected to be in the first β-strand in the nucleotidyltransferase catalytic domain (Hoffmeier *et al.*, 2010) while the two catalytic carboxylates are in the adjacent β-strand (Li *et al.*, 2002). Converting arginine at this position to tryptophan restores activity to the E189F variant, and has only a limited effect on the activity of the native enzyme.

4.6.2 The R64W substitution does not dramatically alter tRNA nucleotidyltransferase structure

To assess the structural effects of the R64W and E189F substitutions, we used circular dichroism spectroscopy to detect changes in the secondary structure of the proteins. The CD spectra were similar for all of the proteins at 22°C (Fig. 4.4) and showed characteristic peak minima at 208 nm and 222 nm that were indicative of high α-helical character. These results supported the fluorescence data (data not shown) and indicated that the proteins shared similar secondary structures confirming what we had seen previously for native enzyme and E189F, E189K, E189H and E189Q variants (Shan *et al.,* 2008).

4.6.3 The R64W substitution does not dramatically alter thermal stability

Given that the R64W amino acid substitution restored enzyme activity and eliminated the temperature-sensitive phenotype in cells carrying the *cca1-E189F* mutation, we wanted to determine whether the R64W substitution restored thermal stability to the E189F variant. To assess thermal stability we followed the CD signal at 222 nm which is indicative of the α-helical content of the protein. All of the proteins showed similar two-state thermal denaturation curves (Fig. 4.5) but exhibited different Tms (Table 4.2). In analyzing these curves we found that the R64W substitution did not alter dramatically the thermal stability or percent remaining α-helicity of either the native protein or the E189F variant (Table 4.2), *i.e*., both the E189F and R64WE189F variants had reduced values when compared to the R64W and native proteins. Taken together these results suggest that the R64W substitution does not alter the structure or the stability of either the native or E189F variant to any major degree. Furthermore, it is clear that suppression of temperature-sensitivity does not correlate with increased thermal stability. Previously, we showed that the E189H variant showed a Tm, percent remaining α-helicity and thermal denaturation curve in good agreement with those of the native enzyme, but yet displayed a temperature-sensitive phenotype (Shan *et al.*, 2008). Here we show that R64WE189F and E189F also have comparable thermal stability parameters, yet R64WE189F suppresses

the temperature-sensitive phenotype while E189F does not, unless it is over-expressed from the *ADH1* or *TDH3* promoter (Fig. 4.6A). Consequently, the suppression of the temperature-sensitive phenotype is not expected to correlate with any differences in denatured protein level caused by the observed differences in thermal stability. However, the suppression of the temperature-sensitive phenotype correlates well with near-native levels of activity, while temperature-sensitivity correlates with, and appears to be characterized by, low specific activity. Moreover, in further support of these observations we have shown that generating the R64WE189K variant restores activity of the E189K variant to near native levels (data not shown). Finally, even with thermal stability parameters effectively identical to those of the native enzyme, but with an intermediate level of activity (~40% of native), the E189Q variant displays an intermediate temperature-sensitive phenotype (Shan *et al.*, 2008). Taken together these observations suggest that the differences in thermal stability of the variant proteins that we have characterized appear not to play a major role in the development of the temperature-sensitive phenotype or its subsequent. If protein stability is not the criterion that defines the temperature-sensitive phenotype then how does the R64W mutation restore enzyme activity and suppress the temperature-sensitive phenotype?

4.6.4 The R64W substitution suppresses the temperature-sensitive phenotype

To explore the role of R64W in the suppression of the temperature-sensitive phenotype, we generated other variant alleles and tested their ability to alleviate the temperature-sensitivity of the *cca1-E189F* mutation (Fig. 4.2). Based on the available crystal structures of class II tRNA nucleotidyltransferases, it is possible to generate a model of the yeast enzyme which posits an electrostatic interaction between R64 and E189 important for enzyme structure or function. This model was supported by our initial observations that enzyme activity was reduced when position 189 was changed to either a positively charged or hydrophobic residue (potentially disrupting this interaction). Moreover, the ability of the R64W substitution to suppress the temperature-sensitive phenotype

conferred by the *cca1-E189F* mutation suggested that this proposed electrostatic interaction was replaced by a hydrophobic interaction. However, this model does not stand up to scrutiny as transformants carrying plasmid borne *cca1-R64W* or *cca1-R64W,E189K* alleles are temperature resistant (Fig. 4.2). The change from R64 to W64 would have disrupted the predicted electrostatic interaction in the native enzyme, and would not have allowed for a hydrophobic interaction in the E189K temperature-sensitive variant. As we had shown that a hydrophobic residue and a positively charged residue could be tolerated at position 64 we also investigated how a negative charge at this position would affect activity. Transformants bearing the plasmid borne *cca1-R64E* allele grew at the non-permissive temperature (Fig. 4.2), thereby allowing us to explore whether we could replace the proposed E189 to R64 salt bridge with an R189 to E64 salt bridge. We found that cells transformed with plasmids bearing *cca1-E189R*, *cca1-R64EE189R*, *cca1-R64EE189F*, or *cca1-R64EE189K* alleles were unable to grow at 37^{o}C (Fig. 4.2), indicating that the R64E substitution cannot suppress temperature-sensitivity.

The data presented here are more consistent with the structural model for E189 that we proposed initially (Shan *et al.*, 2008). Our model showed E189 positioned at a turn between a β-strand and an α-helix in the head and neck region of the protein and we suggested that replacement of the glutamate residue at position 189 by larger residues would alter the organization of this turn and thereby affect enzyme activity. Our observation that a plasmid-borne *cca1-E189R* allele could not suppress the temperature-sensitive phenotype of the *cca1-E189F* mutation, supports this model since arginine has a Van der Waals' volume of 148 Å^3 (Martin and Keller, 1996) which is larger than the previously studied E189K and E189F substitutions that conferred a ts phenotype. An interesting question remains as to how an alteration at position 64 can restore activity and alleviate the temperature-sensitive phenotype. Our data indicate that there is no essential direct electrostatic or hydrophobic interaction between amino acid 64 and amino acid 189. Therefore, we propose that the

change in structure conferred by E189 substitutions is compensated for by a second minor alteration in structure conferred by R64W. The R64W substitution alone does not alter the structure of the protein sufficiently to abrogate enzyme activity (Fig. 4.3) and clearly, charge is not essential at position 64 as in addition to the native arginine, both tryptophan and glutamate substitutions are tolerated at this position to generate an active enzyme. Moreover, some variability in the relative Van der Waals' volumes (Martin and Keller, 1996) is also acceptable at this position (148 $Å^3$ for arginine, 163 $Å^3$ for tryptophan, but only 109 $Å^3$ for glutamate). The importance of a larger amino acid residue at this position may be reflected by the fact that although tryptophan, arginine and glutamate generate a functional enzyme, the position corresponding to residue 64 in the *Saccharomyces cerevisiae* protein is usually occupied by either an arginine or a tyrosine (Van der Waals' volume 141$Å^3$) in many other tRNA nucleotidyltransferases, including those for which crystal structures have been solved (Fig. 4.1). Moreover, only the amino acid with the largest Van der Waals' volume (tryptophan) is able to suppress the temperature-sensitive phenotype (Fig. 4.2). Based on computer modeling and crystal structures of the related eubacterial enzyme, we predict that position 64 in the yeast enzyme is within the catalytically important motif A (Li *et al.*, 2002) and in the first β-strand of the nucleotidyltransferase catalytic domain (Hoffmeier *et al.*, 2010). A glutamate substitution at this position would be expected to alter the organization of the β-strand in a minor manner and this is supported both by the ability of the *cca1-R64E* allele and the inability of *cca1-R64E,E189F*, *cca1-R64E,E189K*, or *cca1-R64E,E189R* alleles to complement the temperature-sensitivity of *cca1-E189F*. In contrast, we suggest that the larger and more hydrophobic tryptophan substitution at position 64 is able to alter the organization of this β-strand in such a way that the deleterious effects of the E189F substitution are suppressed. Regardless of the changes in this β-sheet, they do not measurably affect the thermal denaturation of the native or variant enzymes when loss of α-helical structure is the criterion by which structure is measured. We propose that the substitutions at position 189 result in minor conformational changes in the active site of the

enzyme which reduce enzyme activity. These conformational changes may modify the orientation of the catalytically important β-sheet, or the organization in space of the flanking motifs; A, B, D, and E. The result of this may be to reduce either the amount of the tRNA and/or the nucleotide triphosphate substrates that are bound or their orientation relative to the catalytic residues. We propose that changes at position 189 result in a minor conformational change in the protein that alters the ability of the tRNA substrate to bind to the enzyme. This alteration could affect either the amount or the orientation of the tRNA that is bound, but in any case reduces enzyme activity. Subsequently, introducing the R64W substitution in motif A, known to be directly involved in catalysis and nucleotide triphosphate binding, may subtly adapt the organization of the active site to restore enzyme activity at the expense of some fidelity. This may be achieved by reorganizing the β-sheet to recover substrate binding, or by reorganizing catalytic residues. In fact, we do see some loss in nucleotide selectivity with both the R64W and R64WE189F variants. These data demonstrate the critical role motif C residue 189 has in the organization of the head and neck regions of the enzyme's active site.

4.6.5 The temperature-sensitive *cca1* alleles are hypomorphic

From our data it is clear that the temperature-sensitive phenotype of E189F results not from decreased thermal stability of the mutant protein but primarily from a partial loss of tRNA nucleotidyltransferase activity that affects cell growth only at higher temperatures. Furthermore, this temperature-sensitive phenotype can be suppressed by restoring the activity of the tRNA nucleotidyltransferase but not necessarily its thermal stability. The loss of temperature-sensitivity conferred by overexpression of the *cca1-E189F* allele (Fig. 4.6A) and the more extreme temperature-sensitivity (no growth at 30^{o}C) exhibited by *cca1-E189K/cca1::TRP1* diploids (Fig. 4.6B) support our conclusion that the ts *cca1* mutations are hypomorphic. Furthermore, the ability of the R64W substitution to suppress the deleterious effects of the E189F and E189K substitutions by restoring the activity of the tRNA nucleotidyltransferase but not its thermal stability supports the conclusion that the

substitutions at E189 cause partial loss of activity at both permissive and non-permissive temperatures. While temperature-sensitive (ts) mutants of yeast were first defined by their inability to form colonies at 36°C but to grow "normally, or nearly so," at 23°C (Richards, 1974), in the intervening years, this phenotype has been attributed to a variant gene product that is functional at one temperature (permissive) but which is rendered non-functional at a higher (non-permissive) temperature (Tan *et al.*, 2009, Poultney *et al.*, 2011). It is generally assumed that at the permissive temperature the activity of the mutant protein is similar to wild-type (Hartwell, 1967). The simple idea that a ts phenotype arises because a gene product is functional at the permissive temperature but not functional at the non-permissive temperature is clearly an oversimplification.

We suggest here that the ts phenotype arises not because tRNA nucleotidyltransferase is unstable at the non-permissive temperature, but because its inherent reduced activity (which is sufficient for growth at the permissive temperature) does not meet the needs of the cell at the non-permissive temperature. This phenotype may result simply because the defective enzyme cannot generate functional tRNAs at a sufficient rate for protein synthesis at the non-permissive temperature or it may involve other differences in cells grown at higher temperatures. For example, *CCA1* transcript levels are decreased at the higher temperature (Chakshusmathi *et al.*, 2004). Moreover, in response to stress, tRNA transcription is also down-regulated and endonucleolytic cleavage of cytoplasmic tRNAs increases (Castells-Roca *et al.*, 2011). Our observations are in good agreement with data indicating that rather than producing a thermolabile protein, the widely used *cdc13-1* mutation produces a gene product that is impaired for function even at the permissive temperature of 23°C (Thompson and Parker, 2009). These authors propose that the growth phenotype results from an additive effect of this partial-loss-of-function mutation (whose effect is present at all temperatures) combined with some additional aspects of telomere metabolism that are affected by high temperatures and not simply from a temperature-dependent impairment of a single gene. Moreover, Paschini *et al*. (2012) point out that

other cellular pathways could be affected by higher temperatures and could make a specific hypomorphic mutation generate a temperature-sensitive phenotype. Based on the results presented here, we would argue that this is indeed the case with processes such as mRNA stability and protein synthesis, both of which are affected by tRNA nucleotidyltransferase activity (Peltz *et al.*, 1992). Interestingly, certain null mutations, including genes involved in RNA metabolism, *e.g*., *GLE2* (Paschini *et al.*, 2012), *LRP1 (Miao et al., 2006)* and telomeres, *e.g*., *HPR1 (Mitchell et al., 2003)*; protein folding and degradation, *e.g*., *RPN9 (Jimeno et al., 2002)*, *BUL1 (Takeuchi and Toh-e, 2001)*, *YDJ1* (Yashiroda *et al.*, 1996), *CKS1* (Caplan and Douglas, 1991); and assembly and structure of multiprotein complexes, *e.g*., *FMC1* (Kaiser *et al.*, 1999)*, NUP84* (Lefebvre-Legendre *et al.*, 2001)*, SNF1* (Rout *et al.*, 2000), generate ts phenotypes. This also suggests that a loss of protein function may manifest itself only at the non-permissive temperature due to some change in the intracellular conditions at the higher temperature. Moreover, aneuploids often exhibit a temperature-sensitive phenotype (Tachibana *et al.*, 2007, Hanic-Joyce, 1985). Torres *et al*. (2007, 2008) suggest that, at higher temperatures, imbalanced stoichiometries for members of multiprotein complexes cause proteotoxic stress which, depending on the genes involved, can lead to a temperature-sensitive phenotype. Data are accumulating to indicate that a temperature-sensitive phenotype is not always indicative of a thermolabile gene product. Further research will elucidate what additional factors are involved in generating the ts phenotype in cells with reduced tRNA nucleotidyltransferase activity.

4.7 Supplementary Figures

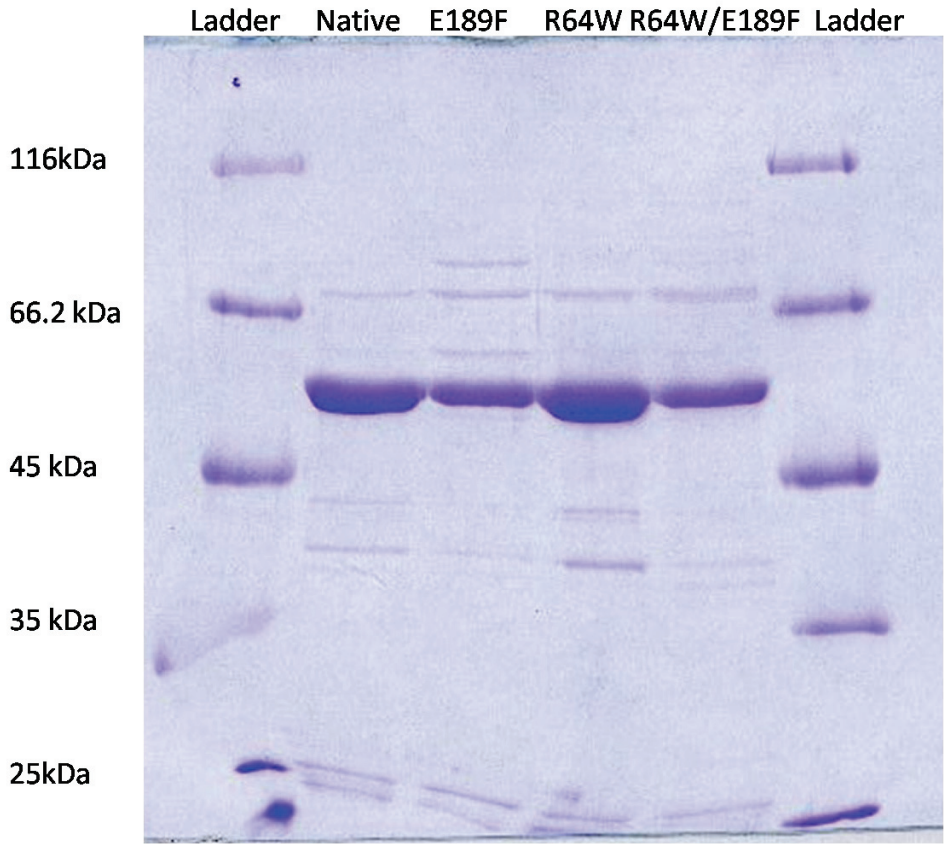


Figure S2.1. Quality of proteins used in this study. Proteins were separated by SDS-PAGE and stained with Coomassie blue. Ladder, PageRuler^{TM} unstained Broad Range Protein Ladder #SM1881, native (3.2 µg); E198F (2.8 µg), R64W (3.2 µg), R64W/E189F (2.8 µg). Sizes of the molecular weight standards are indicated.

Chapter 5: Establishing the link between human tRNA nucleotidyltransferase and congenital sideroblastic anemia with B-cell immunodeficiency, periodic fevers and developmental delay (SIFD)

5.1 Preface

The work presented in Chapter 5 corresponds to the following manuscript in preparation: Leibovitch M. *et al*. (2016). **Finding the link between human tRNA nucleotidyltransferase and congenital sideroblastic anemia with B-cell immunodeficiency, periodic fevers and developmental delay (SIFD)**. Paul Joyce and Matthew Leibovitch designed the experiments and drafted the manuscript. Matthew Leibovitch carried out all studies described here. Experiments using human cell lines to complete this study will be performed by our collaborators at the Children's Hospital of Eastern Ontario.

5.2 Abstract of manuscript

Recently (Chakraborty *et al.*, 2014), seven alleles of the TRNT1 gene (p.T154I, p.M158V, p.L166S, p.R190I, p.L166S, p.I223T, p.I326T, and p.K416E) were identified in patients with congenital sideroblastic anemia with B-cell immunodeficiency, periodic fevers and developmental delay (SIFD). The TRNT1 gene encodes tRNA nucleotidyltransferase which is essential for protein synthesis as it plays a vital role in tRNA maturation. Five of the seven alleles defined amino acid substitutions in the well-characterized motifs A-E of class II tRNA nucleotidyltransferases, while two were located in the less well conserved body and tail domains (Li *et al.*, 2002). We define here the biophysical and biochemical characteristics of the variant proteins to relate changes in tRNA nucleotidyltransferase to the SIFD phenotypes that result. Our data suggest that the SIFD phenotype is linked to poor stability of the T154I, L166S and I223T variant proteins, and to a combination of both reduced stability and catalytic function in both the M158V and R190I variants. Our results also show mutations mapped to the body and tail domains of the protein are interesting. In the case of the I326T variant, the substitution leads to a loss of stability, reduced catalytic activity, lower affinity for the tRNA substrate and a loss of quaternary structure. In

contrast, the K416E substitution results in poor binding and affinity for the tRNA substrate. Together these data suggest that the complex phenotype seen in SIFD patients may result from different defects in tRNA nucleotidyltransferase which affect some or all facets of its structure, stability, activity, and location.

5.3 Introduction

Sideroblastic anemias are a diverse group of diseases defined by the presence of ring sideroblasts (erythroid precursors with perinuclear mitochondrial iron deposition) in bone marrow (Wiseman *et al.*, 2013). While in some instances anemia is the only symptom of the disease, in other cases clinical features may be much broader and include neuromuscular and metabolic phenotypes (Fleming, 2011). Several defective genes responsible for sideroblastic anemia have been identified and they map into three main groups: those involved in heme production, those involved in Fe-S cluster biogenesis and transport, and those involved in mitochondrial protein synthesis (Iolascon, 2014). These genes may be encoded in either the nuclear or the mitochondrial genome (Fleming, 2011). Recently (Wiseman *et al.*, 2013), a clinical subgroup was identified with congenital sideroblastic anemia with B-cell immunodeficiency, periodic fevers and developmental delay (SIFD). As the name suggests, SIFD is a severe multiorgan syndrome with life-threatening complications such that many patients die in the first decade of life (DeLuca *et al.*, 2016).

When whole exome sequencing and descent mapping techniques were used on SIFD patients, an autosomal recessive mode of inheritance was detected and the causative gene responsible for this syndrome was identified as *TRNT1* (Chakraborty *et al.*, 2014). This nuclear gene codes for ATP(CTP):tRNA nucleotidyltransferase (TRNT1), the enzyme responsible for the addition of the conserved cytidine-cytidine-adenosine (CCA) sequence to the 3' end of both nuclear and mitochondrially-encoded transfer RNAs (tRNAs). The identified mutations result in partial loss of function of TRNT1 and lead to metabolic

defects in both the mitochondria and cytosol, which can account for the phenotypic pleiotropy observed in human patients (Chakraborty *et al.*, 2014).

Exome sequence analysis revealed three frameshift alleles, three splicing variants, and seven unique missense TRNT1 alleles characterized here as p.T154I, p.M158V, p.L166S, p.R190I, p.I223T, p.I326T, and p.K416E (Chakraborty *et al.*, 2014). Five of these seven alleles define amino acid substitutions found in the well-characterized motifs A-E (Li *et al.*, 2002) located in the conserved head and neck portions of class II tRNA nucleotidyltransferases (Yue *et al.*, 1996) (Fig. 5.1). None of these mutations map to conserved Motif A which plays a key role in catalysis by coordinating two catalytic magnesium ions and binding the triphosphate moiety of each nucleotide triphosphate (Joyce and Steitz, 1994, Steitz, 1998). However, both the T154I and M158V substitutions map to Motif B which functions by discriminating between deoxyribonucleotides and ribonucleotides at the active site (Li *et al.*, 2002). The L166S substitution is found in the less well conserved Motif C which acts as a connection point between the head and neck region of the protein (Li *et al.*, 2002) but may also act as a flexible spring element modulating the relative orientation of the head and body domains to help accommodate the growing 3' end of the tRNA substrate (Ernst *et al.*, 2015). The R190I substitution is found in Motif D which is responsible for forming a binding pocket to accommodate incoming CTP or ATP (Li *et al*., 2002) and discriminates against UTP and GTP binding (Yue *et al.*, 1996, Shi *et al.*, 1998, Cho *et al.*, 2007). In contrast, the I223T substitution is found in the less well conserved spacer between motifs D and E and the I326T and K416E substitutions were found in the poorly conserved body and tail portions of the protein, respectively (Fig. 5.1). While one can imagine how mutations in the conserved catalytic portions of the protein might result in structural changes that lead to changes in stability or activity resulting in the SIFD phenotype, it is more difficult to understand how mutations in the less well conserved portions of the protein can effect structure or function leading to disease.

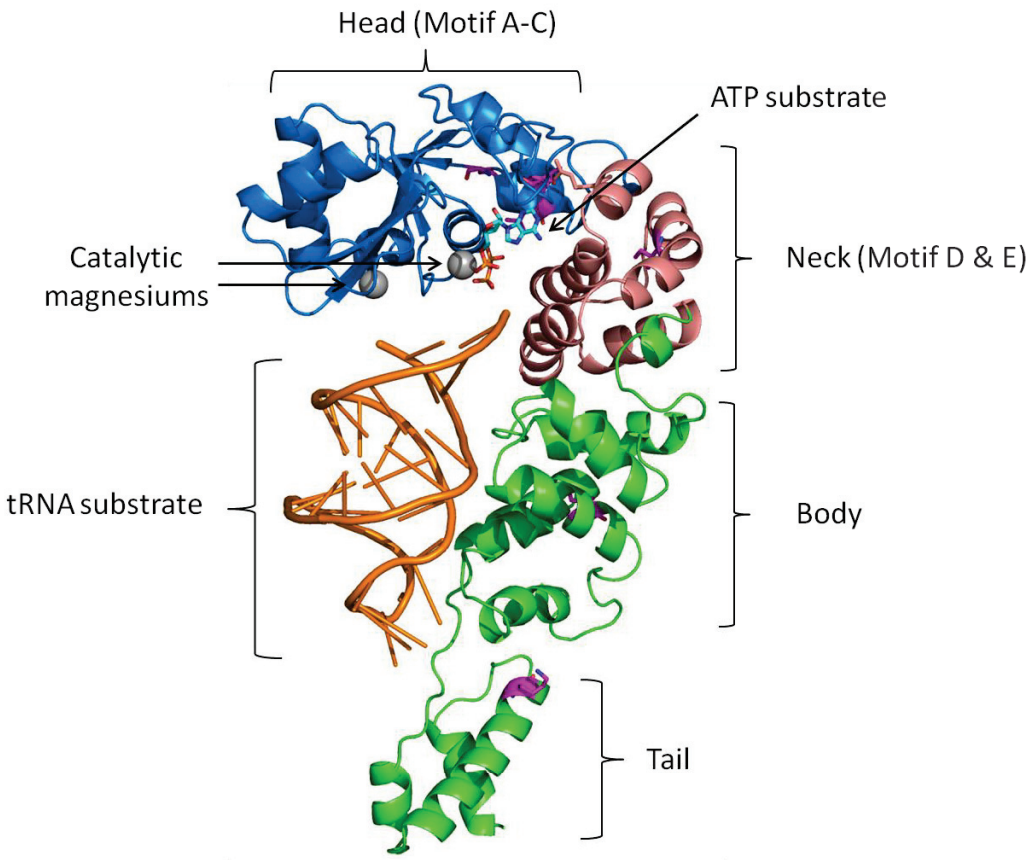


Figure 5.1: Human tRNA nucleotidyltransferase (Kuhn *et al.*, 2015)

The human tRNA nucleotidyltransferase (PDB 4X4W) protein was remodelled with two magnesium ions (grey balls) and ATP (PDB 3H39, Toh *et al.*, 2009) and a tRNA (orange) (PDB 1VFG,Tomita *et al.*, 2004) at the active site. The head (blue), neck (salmon), body and tail (green) domains are visualized showing the polypeptide backbone (cartoon). The ATP substrate is shown in stick (blue) and mutations mapped to the human tRNA nucleotidyltransferase protein are represented in pink.

The ATP(CTP):tRNA nucleotidyltransferase is an enzyme with multiple functions in addition to its role in adding the CCA sequence required for aminoacylation. The tRNA nucleotidyltransferase enzyme also may play a repair function in the cytosol (Rosset and Monier, 1965), be involved in tRNA quality control (Kuhn *et al*, 2015), in the stress response for targeting tRNAs for decay (Wilusz *et al.*, 2011) and in nucleocytosolic shuttling of tRNAs (Feng and Hopper, 2002). Given all of these possible roles in the

cell and the multiple locations where the enzyme must function, it is not surprising that changes in tRNA nucleotidyltransferase structure, stability, activity or localization could lead to the phenotypes observed.

In this study, we show that there are no measurable changes in the secondary and tertiary structures of any of the variants, however single amino acid substitutions at T154I, M158V, L166S, R190I, I223T or I326T all lead to a decrease in stability. Surprisingly, all changes made in the conserved head and neck regions of the protein (motifs A-E) responsible for catalytic function, had no major effect on activity for the tRNA, ATP and CTP substrates. In contrast, both I326T and K416E showed changes catalytically with a 10-fold drop in activity and poor binding of the tRNA substrate. When activity assays were repeated with a modified mini-tRNA substrate, the native and T154I, M158V, L166S, I223T variants showed comparable activity while under these conditions the R190I, I326T and K416E variants showed less activity. Here, we also show that *in vitro* all of the proteins except for the I326T variant form a homodimer stabilized by a disulfide bond. Collectively, these data suggest that changes made in the head and neck portions lead to altered stability and changes in the body and tail affect tRNA binding, activity and dimerization. Ultimately, an understanding of how these changes affect tRNA nucleotidyltransferase function may aid in developing treatments for SIFD.

5.4 Materials and methods

5.4.1 Plasmids and cloning

The cDNAs coding for the native and variant human tRNA nucleotidyltransferases in pTrueORF (Origene) (Chakraborty *et al.*, 2014) were used as templates for the polymerase chain reaction. The open reading frame for the native or mutant cDNA was amplified by PCR as follows: 200 µmoles of dNTPS, 5X Phusion HF buffer (New England Biolabs), 0.25 units of Phusion DNA Polymerase, 25 pmoles of primers and 25 pmoles of template as described previously (Leibovitch *et al.,* 2013) with an annealing temperature of 51°C. The primers used are 5'-ATCCCATATGAAGTTGCAGTCTCCCG-3' (start codon

underlined) and 5'-ATGTCTGGATCCTTAGGTCTTCTTTATGTAACTCAG-3' (stop codon underlined). The resulting PCR products were digested with *Bam*HI and *Nde*I and ligated into pET-15b digested with the same enzymes.

The G73 and pmBsDCCA plasmids, both generous gifts from Dr. Alan Weiner (University of Washington), were originally constructed by Cho *et al.* (Cho *et al.*, 2003) and Oh and Pace (1994), respectively and both contain the *Bacillus subtilis* $tRNA^{Asp}$ gene modified to produce specific 3' ends after restriction digestion and *in vitro* transcription. Plasmids pMG73 and pMpmBsDCCA expressing mini-tRNA templates containing only the TψC stem and loop and the aminoacyl stem of the $tRNA^{Asp}$ were generated from G73 and pmBsDCCA by removing 31 nucleotides including the DHU, anticodon and variable stems and loops from the $tRNA^{Asp}$ gene by using QuikChange™ mutagenesis to amplifying only a portion of the gene of interest. The following primers were used:

5'-GCGGGTTCGAGTCCCGTCCGG-3' and 5'-CCGGACCTATAGTGAGTCGTATTAGGGGATCC -3' and the resulting products were ligated and transformed into *E.coli* XL2 cells.

The identity of each product was confirmed by sequence analysis (McGill University and Génome Québec Innovation Centre).

5.4.2 Expression and purification of tRNA nucleotidyltransferase

E.coli BL21 (DE3) carrying the plasmid of interest was grown in 10 ml of YT + ampicillin (100 µg/ml) medium at 37°C and 225 rpm for 16-18 hours. Cultures were transferred in 5 ml aliquots into flasks containing 1.3 l of YT + ampicillin (100 µg/ml) medium and returned to the shaker at 37°C and 225 rpm for 2-3 hours until an OD_{600} of 0.4-0.8 was achieved. The cells then were induced with a final concentration of 0.5 mM IPTG and returned to the shaker at 19°C and 193 rpm for 16-20 hours. Cells were collected by centrifugation at 4°C in a Beckman JA-10 rotor for 15 minutes at 6 000 rpm. Cell pellets were stored at -80°C for a minimum of 1 hour before used.

Frozen cell pellets were thawed on ice while being resuspended in cold lysis buffer (PBS + 1 mM EDTA) at approximately 2 ml/g of pellet. The resuspended cell suspension was passed through a French Pressure cell press (Thermospectonic) three to four times at 1000 psi. The resulting cell lysate was centrifuged at 18 000 rpm in a Beckman JA-20 rotor once for 30 min at 4^{o}C. The cleared lysate was kept and centrifuged once more as above before being passed two to four times through a 10 cm Bio-Rad column containing 2-3 ml of NTA-Ni resin (Novagen). The column was washed first with 50-100 ml of PBS (pH 7.5) + 10% glycerol and then with 50-100 ml 10 mM imidazole in PBS (pH 7.5) + 10% glycerol to eliminate non-specific binding. Bound protein was eluted stepwise with 6-8 ml each of 50 mM, 100 mM, 150 mM and 250 mM imidazole in PBS (pH 7.5) + 10% glycerol. Fractions were analyzed by SDS-PAGE and appropriate fractions were pooled, transferred to a 50 kDa dialysis bag (SpectroPor) and dialysed against 1-2 l of PBS (pH 7.5) + 10% glycerol for 1-2 hours at 4^{o}C. Dialysis was repeated as above for 3-5 hours and then overnight. The protein was removed from the dialysis bag and stored at -80^{o}C in aliquots.

Protein concentrations were determined by Bradford assay (Bradford, 1976) and based on the molar extinction coefficient (38515 $M^{-1}cm^{-1}$) at 280 nm for tRNA nucleotidyltransferase (Gasteiger, 2006). Absorption readings were determined in a 1.0 cm quartz cell using the Varian Cary 100 Bio UV-Visible Spectrophotometer. All samples were treated with a final concentration of 20 ng/ml RNAseA (Bioshop) and dialyzed extensively in PBS (pH 7.5) to remove any nucleic acid associated with the protein.

5.4.3 Biophysical characterization of tRNA nucleotidyltransferase

5.4.3.1 Circular dichroism spectroscopy

Secondary structure determination was done by using circular dichroism spectroscopy. Protein samples, prepared as described above (section 5.4.2), were diluted in PBS to approximately 0.4 mg/ml

and placed in a 0.2 cm cell cuvette. Prior to taking measurements, samples were centrifuged at 14 000 x g for 5 minutes at 4^{o}C. A Jasco-815 circular dichroism spectrophotometer along with the "Spectrum Measurement" program (Jasco) was used and was set to scan over the region of 200 nm to 260 nm. At standard sensitivity, a bandwidth of 1 nm was used, along with a response time of 0.25 seconds, data pitch of 0.2 nm, and a scanning speed of 20 nm/min. Data were accumulated five times at 20°C using a Pelletier water bath accessory while nitrogen was set to flow through the instrument at a constant rate of 3 l/min. The data obtained were smoothed using the smoothing function in the Spectra Analysis (Jasco) program.

5.4.3.2 Thermal denaturation monitored by circular dichroism

Temperature denaturation was monitored by circular dichroism for each protein. A sample of approximately 0.5 mg/ml was prepared as described above (section 5.4.2). The program used was "Variable Temperature" where the signal was monitored at 222 nm at standard sensitivity using a bandwidth of 1 nm, response time of 0.25 seconds, and a data pitch of 0.2°C. A temperature slope of 30°C/hour was used while the start and end temperatures were 20°C and 60°C. A 0.2 cm cell was used and was sealed to limit evaporation. Data obtained were smoothed using the smoothing function in the Spectra Analysis (Jasco) program.

5.4.3.3 Tryptophan fluorescence spectroscopy

Tertiary structure determination was done by using fluorescence spectroscopy. The instrument used was a Varian Cary Eclipse Fluorescence Spectrophotometer along with a Varian single cell Pelletier Cary accessory (water bath). The application "Scan" was used to obtain emission spectra at room temperature averaged over 10 scans between 310 nm to 400 nm with an excitation wavelength of 280 nm or 295 nm. Scan speed was set to medium (600 nm/min) with a 1.0 nm data sampling interval, 5 nm excitation and emission slit widths, and voltage was set to medium (600 Volts). Samples (1 ml) were

prepared as described above (section 5.4.2) and diluted to a final concentration of 2.2 µM determined by OD_{280} (38 515 $M^{-1}cm^{-1}$ molar extinction coefficient) for the spectrum measurements and placed in a 1 cm cuvette. The data obtained were smoothed using the moving average smoothing function (Cary Eclipse software, Agilent).

5.4.3.4 Gel filtration chromatography

The quaternary structure and the binding capacity of the tRNA substrate of tRNA nucleotidyltransferase was assessed by gel filtration chromatography using a Superdex 200 10/300 GL column (GE Healthcare Life Sciences). The flow rate was set to 0.65 ml/min and samples were monitored by absorption at 280 nm. The column was equilibrated with at least two volumes of 50 mM Tris-HCl (pH 7.6) and 100 mM NaCl prior to sample injection. All protein samples 500 µl (0.3-1 mg/ml) were centrifuged for five minutes at 4^{o}C prior to injection. RNA-free samples were prepared as described above (section 5.4.2). To see the effect of RNA on the proteins, the RNaseA treatment was not used and 6 µM purified baker's yeast tRNA (Roche) was added with preincubation for 1 h at 4^{o}C prior to injection. When needed, samples were preincubated with reducing agent (2 mM TCEP, Sigma-Aldrich) for 1 hour at 4^{o}C prior to injection. Bio-Rad Molecular weight standard (151-1901) was used to generate calibration curve and assign the molecular weight of each species (see S5.1).

Equations:

[Dimeric protein] = (Absorbance at 1250 seconds)(extinction coefficient of protein at 280 nm over extinction coefficient of tRNA and protein at 280 nm)/((Extinction coefficient of protein)x(pathlength))

[Total protein] = (Absorbance at 1475 seconds + Absorbance at 1350 seconds of RNaseA treated sample) /((Extinction coefficient of protein)x(pathlength))

[Total tRNA] = (Absorbance at 1250 seconds + Absorbance at 1350 seconds + (absorbance at 1475 seconds - absorbance at 1475 seconds of RNaseA treated sample) + (absorbance at 1700 seconds - absorbance at 1700 seconds for RNaseA treated sample))x(extinction coefficient of tRNA at 280 nm over extinction coefficient of tRNA and protein at 280 nm)/(Extinction coefficient of tRNA)x(pathlength)

[tRNA complexed/protein complexed] = (Absorbance at 1250 seconds + Absorbance at 1350 seconds)x(extinction coefficient of tRNA at 280 nm over extinction coefficient of tRNA and protein at 280 nm)/(Extinction coefficient of tRNA)x(pathlength)

k_d=(Free protein)x(Free Ligand)/ (Bound complex)
k_d = ((Total protein) – (protein-tRNA complex))x((Total tRNA – protein-tRNA complexed))/ ((protein-tRNA complex))

5.4.4 Enzyme activity assays

5.4.4.1 Run-off Transcription and purification of transcripts

Each 100 µl run-off transcription reaction contained 20 µL of 5X transcription buffer (Thermo Scientific), 5 nmoles ATP, 50 nmoles each of CTP, GTP and UTP (Thermo Scientific) and 50 µCi (~16 pmoles) [α-^{32}P] ATP (10 µCi/µl, 3000 Ci/mmol) (Perkin-Elmer), ~5 µg of linearized DNA template, 60 units of T7 RNA polymerase (Thermo Scientific) and nuclease-free water and was incubated for 2-3 hours at 37^{o}C. The products were phenol extracted, ethanol precipitated and gel purified via polyacrylamide gel electrophoresis. The concentration of each runoff transcription product was determined by scintillation counting (LKB WALLAC-1217 RackBeta liquid scintillation counter) using a standard curved.

5.4.4.2 Standard Activity Assays

Each standard assay mixture contained 100 mM glycine buffer (pH 9), 6 mM $MgCl_2$, 1.2 mM ATP, 0.1 mM CTP, 10-20 ng of the [α-^{32}P] ATP labelled tRNA transcript and 0.1-100 ng of protein (modified from Leibovitch *et al*., 2013). Assays were initiated with the addition of protein and stopped with 10 µl of Peattie's (2X) loading dye (Peattie, 1979) after two minutes. Samples were heated to 70°C and cooled prior to loading onto a 50 cm x 32 cm 7M urea 12% acrylamide gel and electrophoresed for 7-8 hours at 2000 Volts. The gel was place against a phosphorImager screen (GE Healthcare) for at least 30 minutes, the screen was developed and transcripts were detected using the Typhoon™ TRIO Variable Mode Imager (GE Healthcare).

5.4.4.3 Standard Kinetics Assays

For kinetic analyses, the standard assay as described in 5.4.4.2 was used except with a maximum of 1 ng of protein and increasing amounts of $tRNA^{Asp}$. Each assay tube contained enough reagents for two to three individual assays. At each time point, 10 µl of the assay mixture was removed and treated as described (section 5.4.4.2). The ratio of reactant to product (tRNA-NCC/tRNA-NCCA or tRNA-NC/tRNA-NCC) was measure by densitometry and initial rates were determined for varied amounts of tRNA substrate. Similarly, the assays were repeated with varying amounts of ATP or CTP with a fixed amount of $tRNA^{Asp}$ lacking the terminal CMP residue and/or AMP residue.

5.4.5 Homology Modeling of Human tRNA nucleotidyltransferase

The homology model of human tRNA nucleotidyltransferase with bound magnesium ions, ATP and tRNA was derived from the *Homo sapiens* (PDB 4X4W) (Kuhn *et al.*, 2015), *Thermotoga maritima* (PDB 3H39) (Toh *et al.*, 2009) and *Aquifex aeolicus* (PDB 1VFG) (Tomita *et al.*, 2004) crystal structures of Class II tRNA nucleotidyltransferases. The primary and secondary structures of these proteins were aligned using MODELLER v.9.13 (Ben Webb). The homology model was generated using MODELLER (Ben Webb) incorporating the *Thermotoga martima* $tRNA^{Phe}$ substrate modeled at the *Aquifex aeolicus* active site and both catalytic magnesium ions and ATP from *Thermotoga maritima* crystal structure into the human enzyme. The homology model with the lowest energy level was viewed by PyMol (Schrödinger).

5.5 Results

5.5.1 Biophysical characterization of the tRNA nucleotidyltransferase variants

Seven different disease-causing mutations have been mapped to TRNT1. These mutations result in the following amino acid substitutions: T154I or M158V in motif B, L166S in motif C, R190I in motif D, I223T between motifs D and E, and I326T or K416E in the more carboxy-terminal portion of the protein (Fig 5.1). To link these mutations to the disease phenotype, it is important to define how these amino

acid changes affect protein structure, stability, activity and location. Here, any changes in structure, stability and activity will be explored. To determine whether any of the seven single amino acid substitutions alter the secondary structure or melting temperature of tRNA nucleotidyltransferase, we used circular dichroism (CD) spectroscopy. The spectrum of the native enzyme has the characteristic peak minima at 208 nm and 222 nm expected for a protein with primarily α-helical character (Fig. 5.2). This was as expected based on the available crystal structures of the native enzyme (Augustin *et al.*, 2003, Kuhn *et al.*, 2015). The circular dichroism spectra of all variants showed excellent agreement with the spectrum of the native enzyme indicating that no single amino acid substitution dramatically altered the secondary structure of this protein (Fig. 5.2).

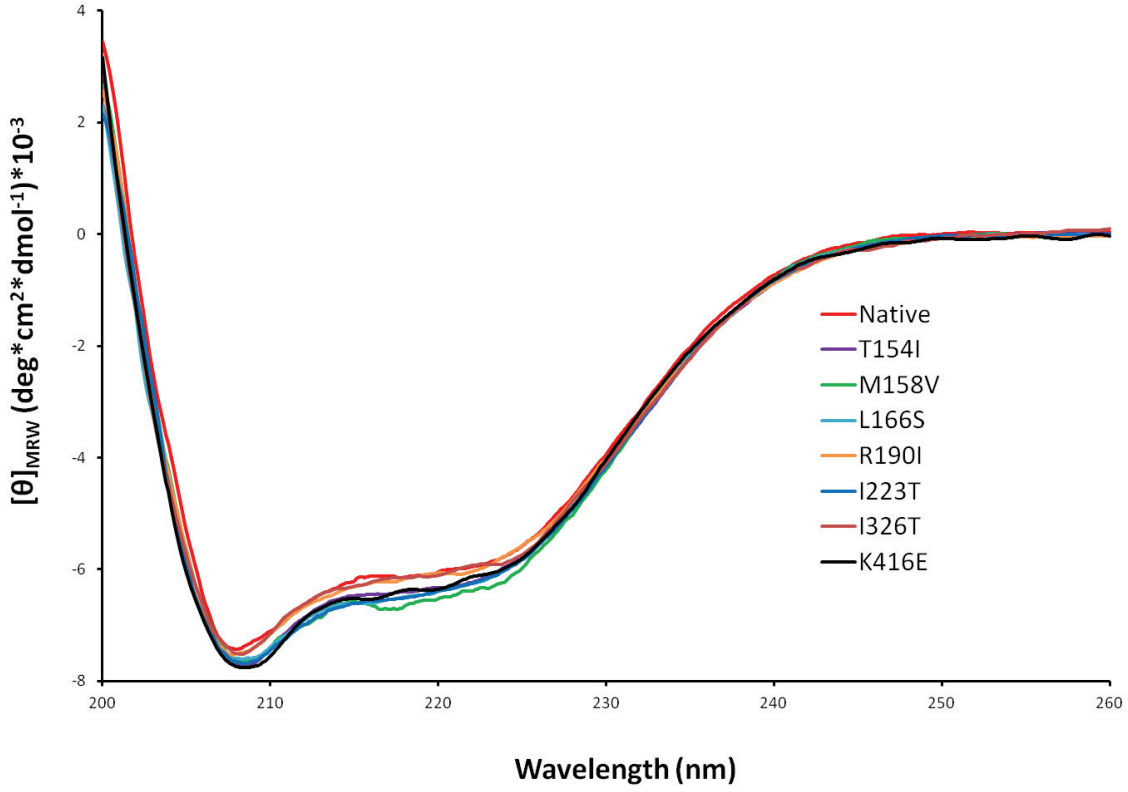
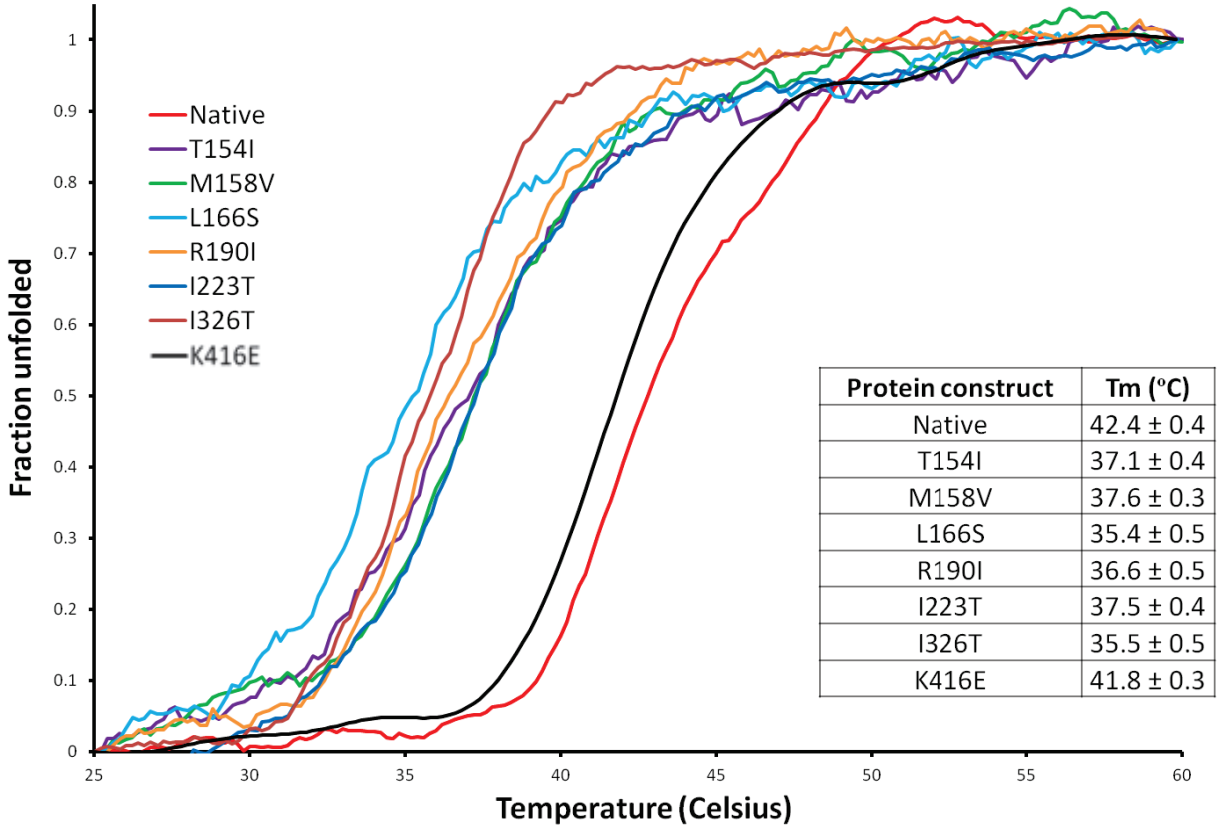


Figure 5.2: Secondary structure comparisons of human tRNA nucleotidyltransferases
The CD spectra of the native and variant proteins as purified from *E. coli* were recorded at 20^{o}C. The CD spectra were taken after the removal of RNA by RNaseA treatment and extensive dialysis in PBS (pH 7.5).

In contrast, when the melting temperature of the native and variant enzymes was monitored by following the change in CD peak minimum at 222 nm with increasing temperature, all of the variants showed a 5-7°C decrease in melting temperature as compared to the native enzyme (Fig. 5.3) except for the K416E variant which showed a melting temperature similar to that of the native enzyme.

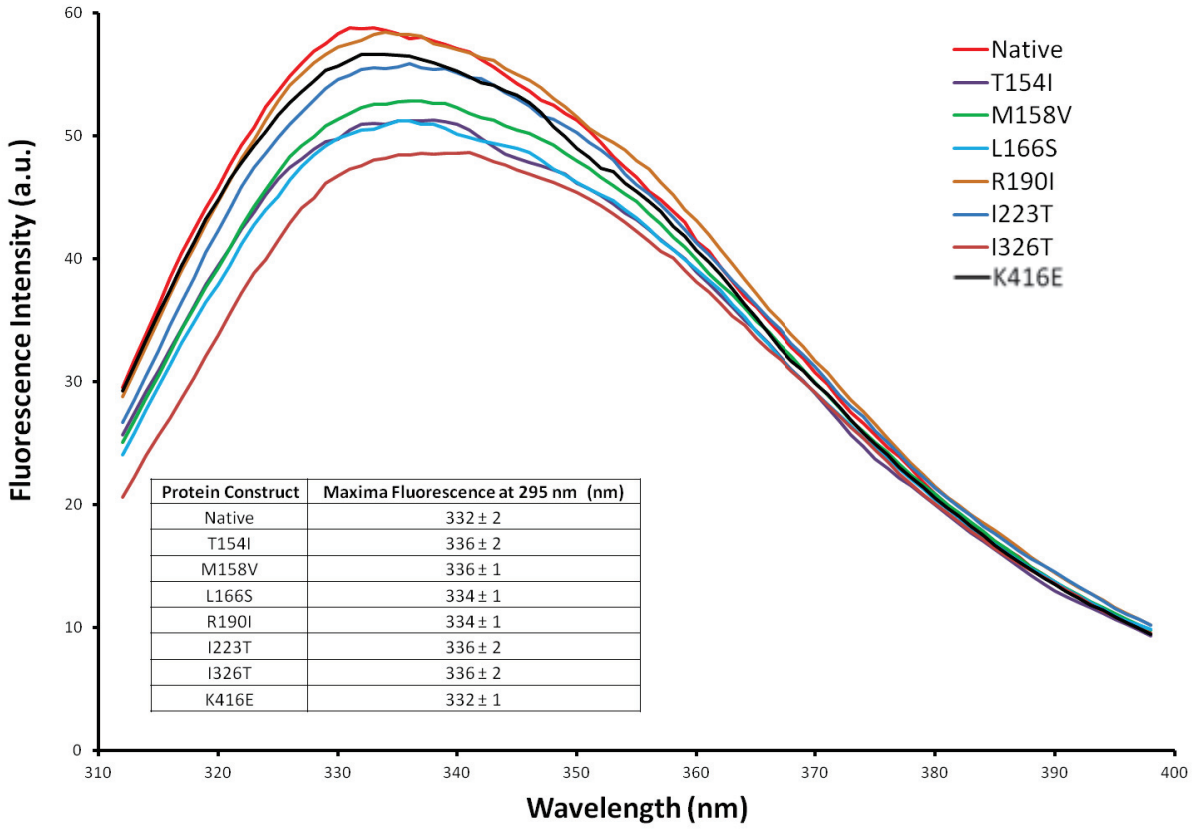
As the mature native protein contains only four tryptophan residues (one in the region between conserved motifs A and B in the head domain of the protein, one in the highly conserved motif E in the neck domain and two in the less well conserved tail domain of the protein), fluorescence spectroscopy (upon excitation of tryptophan residues at 295 nm) was used to determine if any of the amino acid alterations resulted in changes in the tertiary structure of the protein.



Protein construct	Tm (°C)
Native	42.4 ± 0.4
T154I	37.1 ± 0.4
M158V	37.6 ± 0.3
L166S	35.4 ± 0.5
R190I	36.6 ± 0.5
I223T	37.5 ± 0.4
I326T	35.5 ± 0.5
K416E	41.8 ± 0.3

Figure 5.3: Thermal unfolding of human tRNA nucleotidyltransferases
Signal intensity at 222 nm was monitored between 25°C and 60°C for the native and variant proteins after the removal of RNA by RNaseA treatment and extensive dialysis in PBS (pH 7.5). Fraction of protein unfolded was defined by the equation $F = 1/\Delta\theta T \times \theta - \theta i/\Delta\theta T$ where F is the fraction of protein unfolded, θ is ellipticity measured in mdeg, and ΔθT is the total change in ellipticity between the initial and final temperatures of interest; $\Delta\theta T = \theta f - \theta i$. θi is the ellipticity value at the start temperature of interest (25°C), and θf is the final ellipticity value at the end temperature of interest (60°C).

The native protein gave a spectrum with a fluorescence emission maximum at 332 nm which is within experimental error (1-2 nm) of the fluorescence emission maxima seen for all of the variants (Fig. 5.4). The spectra for the T154I, M158V, L166S and I326T variants showed a 15 - 20% drop in fluorescence intensity at maximum (Fig. 5.4) while the I223T, R190I and K416E variants showed almost no difference in fluorescence intensity (~5% lower as compared to native).



Protein Construct	Maxima Fluorescence at 295 nm (nm)
Native	332 ± 2
T154I	336 ± 2
M158V	336 ± 1
L166S	334 ± 1
R190I	334 ± 1
I223T	336 ± 2
I326T	336 ± 2
K416E	332 ± 1

Figure 5.4: Fluorescence emission spectra of human tRNA nucleotidyltransferases
Approximately 2.2 µM of each protein was excited at 295 nm and the fluorescence emission was measured from 310 to 400 nm. The lines represent the native and variant proteins after the removal of RNA by RNaseA treatment and extensive dialysis in PBS (pH 7.5).

Gel filtration chromatography was used to assess the quaternary structure of the native and variant tRNA nucleotidyltransferases. When the heterologously-expressed enzymes were isolated from *E. coli*, the native enzyme showed four absorbance peaks: one at 22 kDa, another at 50-52 kDa, a third at 78 kDa and a fourth at ~136-144 kDa (*e.g.*, Fig. 5.5 A, 5.6 A). The T154I, M158V, L166S and R190I

variants showed four absorbance peaks similar to the native enzyme: one at 19-24 kDa, another at 50-52 kDa, a third at 71-84 kDa and a fourth at 131-138 kDa (Fig. 5.5 B-E, respectively). In contrast, the I223T variant showed absorbance peaks at 24 kDa, 52 kDa and 162 kDa (Fig. 5.5 F), the I326T variant showed absorbance peaks at ~22 kDa, 56 kDa and 84 kDa, (Fig. 5.6 B) and the K416E variant showed absorbance peaks at 19 kDa, 50 kDa and 100 kDa (Fig. 5.6 C).

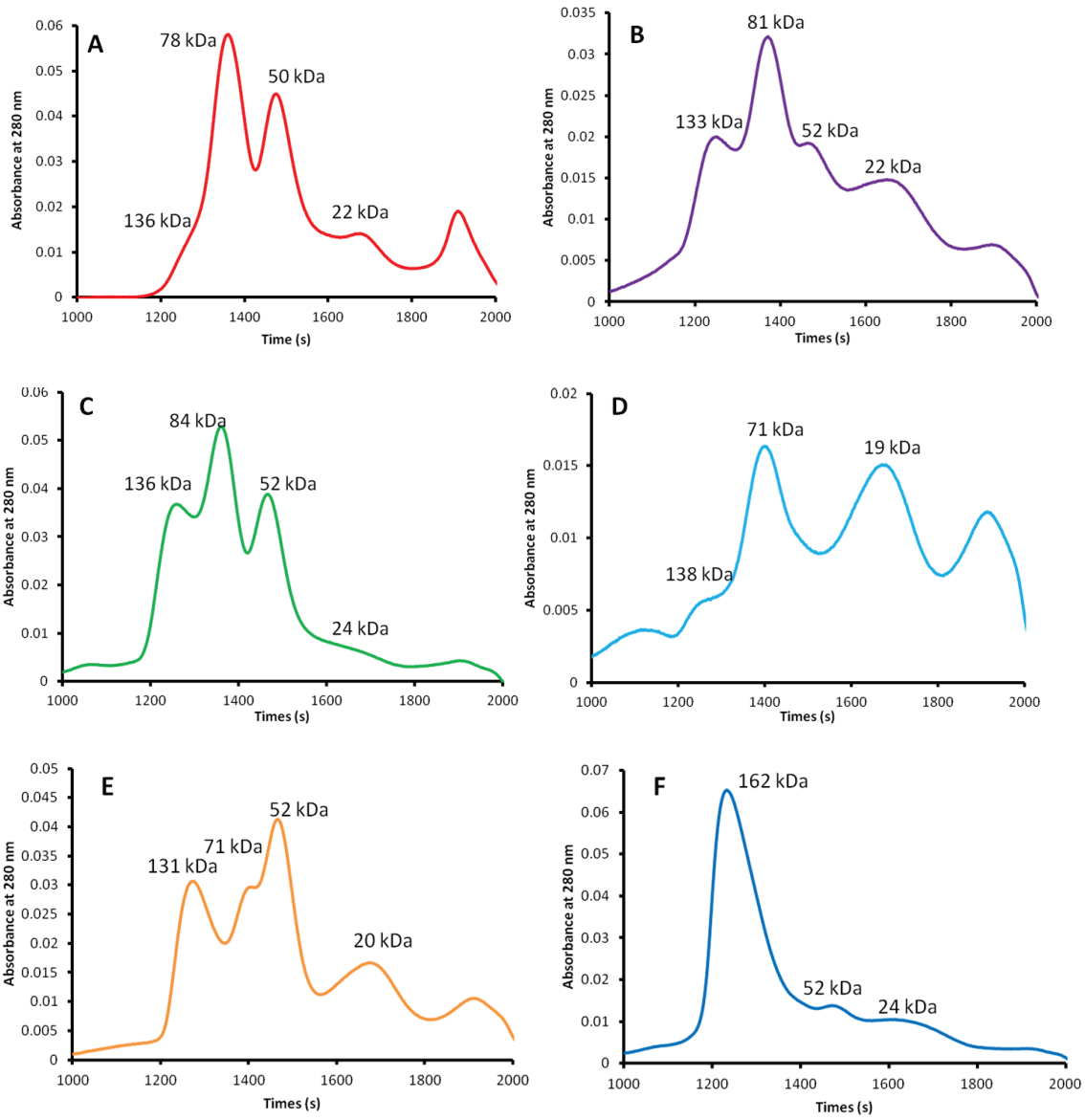


Figure 5.5: Gel filtration chromatography of human tRNA nucleotidyltransferases
A) native, B) T154I, C) M158V, D) L166S, E) R190I, and F) I223T proteins purified from *E. coli* and eluted from a Superdex 200 10/300 GL gel filtration column at 22^{o}C. Molecular masses were calculated as described in section 5.4.3.4.

As we previously showed that heterologously-expressed tRNA nucleotidyltransferase isolated from *E. coli* has bacterial tRNA associated with it (Leibovitch *et al.*, 2013), we treated the native enzyme and the I326T variant with RNaseA to remove any tRNA to better define what forms of the protein were present in each absorbance peak. If the protein migrates close to its expected molecular mass of 50 kDa and the 78-84 kDa peak represents the protein plus bound tRNA then we should expect that the relative amounts of these two peaks will change upon RNaseA treatment. This was indeed the case for both of the enzymes (Fig. 5.6 A-B). The shift in molecular mass after addition of RNaseA suggests that the 50 kDa species represents a monomer and the 78-84 kDa species represents the monomer plus tRNA. Interestingly, the ~140 kDa species in the native enzyme was lost upon RNaseA treatment and a new species at 101 kDa appeared (Fig. 5.6 A). This suggests that the ~140 kDa species represents a dimer with one to two tRNAs attached while the 101 kDa species represents the dimer with no tRNA associated with it. No species of greater than about 84 kDa is apparent in the elution profile of the I326T variant (Fig. 5.6 B). The disappearance of the 22 kDa peak in both spectra after RNaseA treatment suggests that this peak represents free tRNA (Fig. 5.6 A-B).

To further assess the role that tRNA may play on the elution profiles of these two enzymes, the proteins were incubated with 6 µM of commercial baker's yeast tRNA. This is approximately twice the amount of tRNA associated with the heterologously-expressed enzymes and when added it enriched both the 78 kDa and ~140 kDa peaks evident in the native enzyme and the 78 kDa peak (to a small extent) in the elution profile for the I326T variant (Fig. 5.6 A-B). Interpretation of these data is complicated by the fact that the molar extinction coefficient (ε) for tRNA at 280 nm ($3x10^5 M^{-1}{\cdot}cm^{-1}$) is about 10 times greater than that of tRNA nucleotidyltransferase ($3.8x10^4$ $M^{-1}{\cdot}cm^{-1}$) meaning that peak size is defined by both a contribution from the protein and from the tRNA. Using the known extinction coefficients and deconvoluting the absorption peaks, these data suggest that the native enzyme and the K416E variant exist as both monomers and dimers with the monomer predominating in each case with a

dimer:monomer ratio of 1:10 for the native enzyme and 1:2 for the K416E variant. These data further suggest that the tRNA complexes more efficiently with the native monomer (with an equal ratio of protein with and without bound tRNA) than with the I326T monomer (with a 1:10 ratio of protein with and without tRNA bound). Using the ratio of protein complex to free protein and knowing the amount of protein and tRNA added, apparent dissociation constants of 2.4 μM ± 0.3, 86 μM ±16 and 316 μM ± 55 were calculated for the native enzyme, and the I326T and K416E variant enzymes, respectively.

As the available crystal structure of the human enzyme (Augustin *et al.*, 2003) showed a disulfide bond between the Cys373 residues of the two monomers, a reducing agent (2 μM TCEP) was added to the native enzyme and the K416E variant to see what affect it would have on the higher order structure (Fig. 5.6 A and C). In both cases the absorbance peaks for the higher molecular weight species (at ~140 kDa and 104 kDa) suggestive of dimers, disappeared and the 78 kDa and 50-52 kDa peaks indicative of monomers increased (Fig. 5.6 A and C). As no apparent dimer was seen in the elution profile of the I326T variant no reducing agent was added to that sample. Again, using the ratio of protein complex to free protein and knowing the amount of protein and tRNA added, an apparent dissociation constant of 1.8 μM ± 0.14 was calculated for the native enzyme in the presence of reducing agent. This is in good agreement with the value of 2.4 μM ± 0.3 seen without added reducing agent.

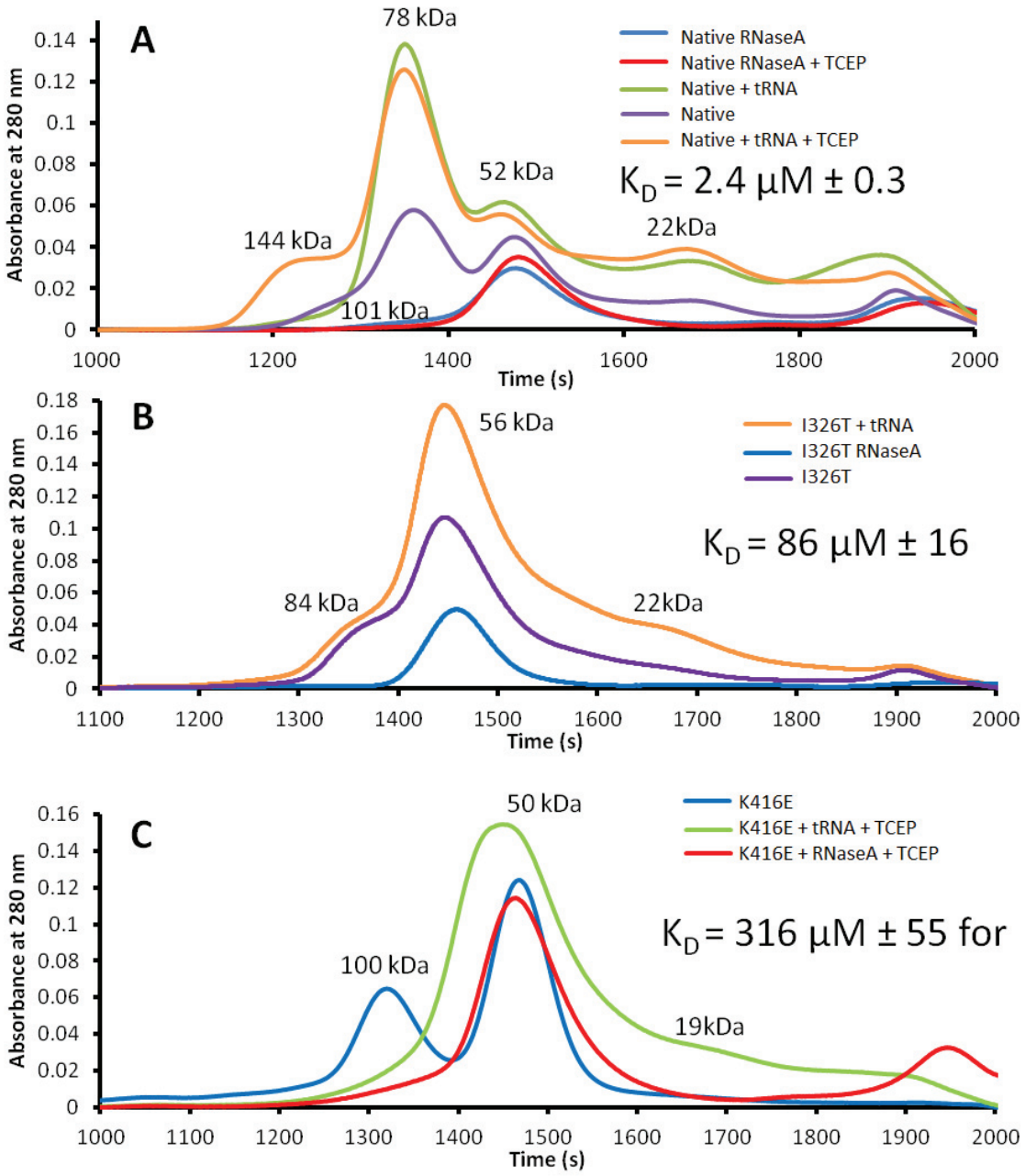


Figure 5.6: Quaternary structure and tRNA binding of human tRNA nucleotidyltransferases
The gel filtration (Superdex 200 10/300 GL) chromatograms of A) native, B) I326T, and C) K416E (C) proteins as purified from *E. coli* at 22°C. Proteins were monitored with co-purifed *E.coli* tRNA (protein alone), after addition of Baker's yeast tRNA (tRNA), after RNaseA treatment (RNaseA) and in the presence of reducing agent (TCEP).

To support the presence of dimers and monomers predicted from the gel filtration observations, a standard 13% SDS-PAGE was carried out in the absence of any reducing agent (Fig. 5.7). With the reducing agent (panel A), all proteins migrate at a similar molecular mass of approximately 48 kDa except for I223T which appears slightly larger. Mass spectrometry confirmed that all of the proteins had the expected molecular masses of 48.5 ± 0.1 kDa (depending on the amino acid substitution). In the absence of reducing agent (panel B), all of the proteins show a major band at about 48 kDa and two additional bands at approximately 100 kDa and 115 kDa suggestive of a protein dimer. Only I326T lacks these higher molecular weight species in good agreement with what was seen in the gel filtration chromatography.

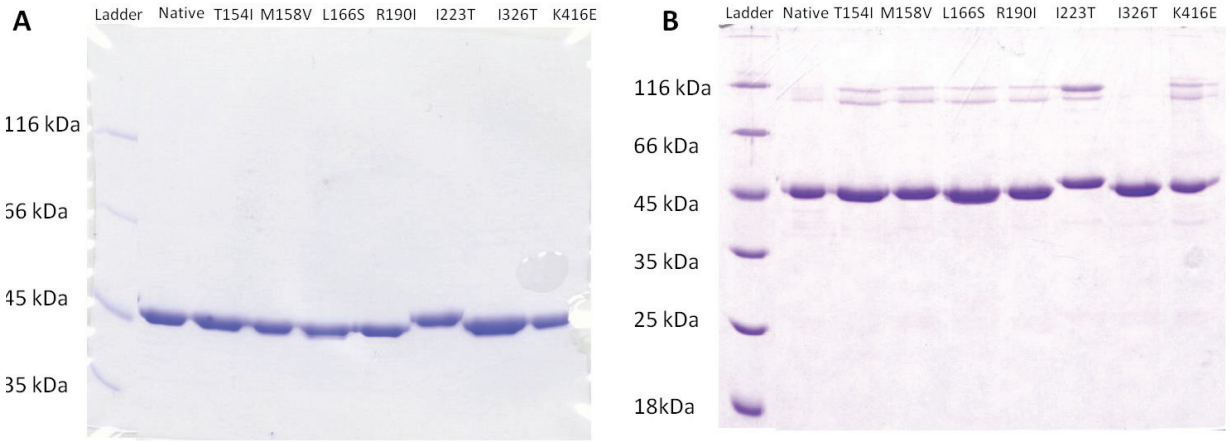


Figure 5.7: SDS-PAGE of human tRNA nucleotidyltransferases
Approximately 4 µg of the native and variant proteins were incubated in SDS loading dye with (A) and without reducing agent, boiled and analyzed by SDS-PAGE. Under reducing conditions (A), protein bands were detected at approximately 48 kDa. Under non-reducing conditions (B), protein bands were detected around 48 kDa, 100 kDa and 115 kDa. Molecular weights were assessed based on the molecular weight markers (Pierce unstained protein molecular weight marker 26610).

5.5.2 Enzymatic characterization of tRNA nucleotidyltransferase variants

5.5.2.1 Activity

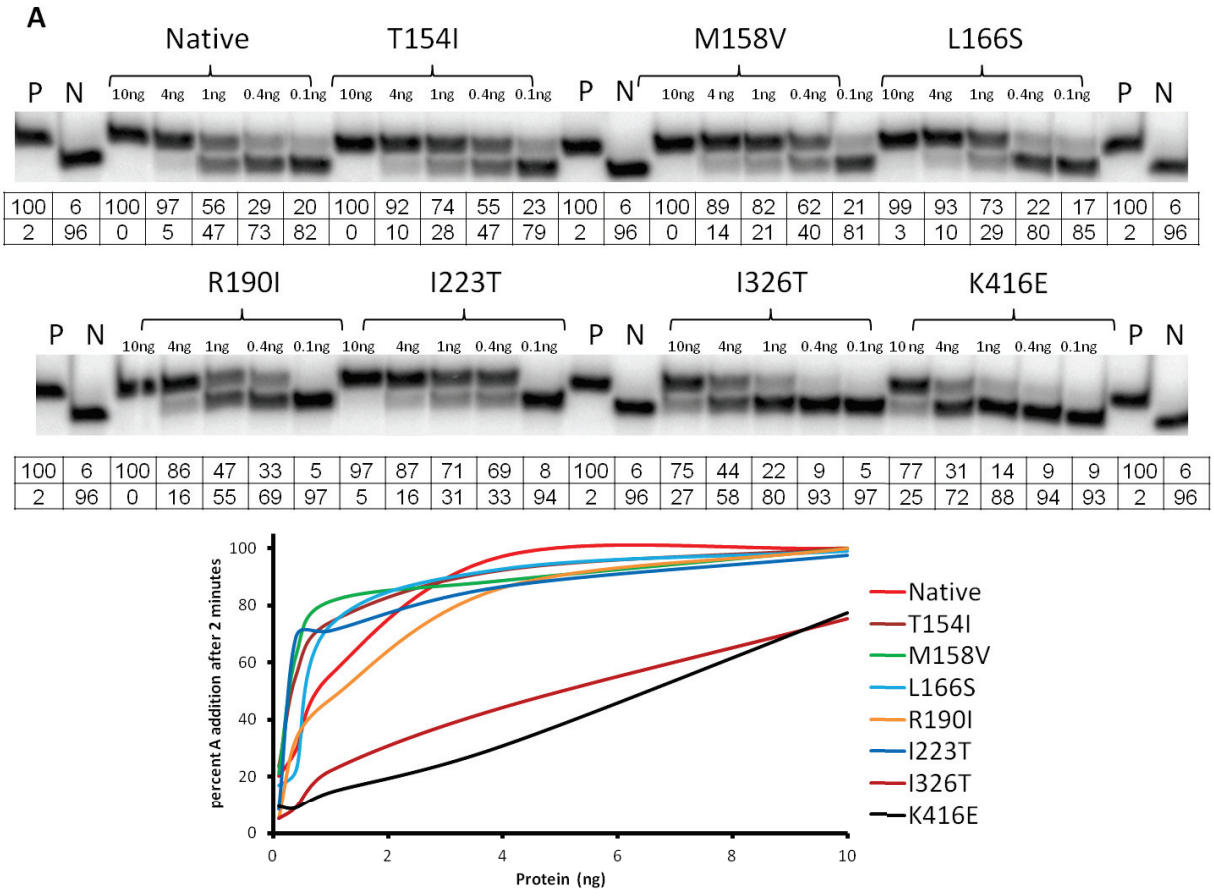
To assess what affect the amino acid substitutions had on enzyme activity, each variant was assayed using increasing amounts of enzyme in the standard (Leibovitch *et al.*, 2013) enzyme assay.

Initially, all enzymes were incubated with a tRNA substrate lacking only the terminal AMP residue to monitor the addition of AMP (Fig. 5.8 A). While all of the substrate is converted to product over the time frame tested at 10 ng of protein for the native and the T154I, M158V, L166S, R190I and I223T variant enzymes, both the I326T and K416E variants show only about 75% conversion of substrate to product under these conditions. If one compares relative activity to protein concentration (Fig. 5.8 A) it suggests that the amount of protein required to convert 50% of the substrate to product in this standard assay is 0.4-0.5 ng of enzyme for the native and T154I, M158V, L166S, R190I and I223T variants but about 10 times higher (~5-7 ng) for the I326T and K416E variants.

When the same assay was repeated using a mini-tRNA containing only the TψC stem and loop and the aminoacyl stem of the $tRNA^{Asp}$, again the native and T154I, M158V, L166S and I223T variant enzymes showed similar levels of activity at 10 ng protein required to convert all of the substrate to product while in this case the R190I, I326T and K416E variants showed less activity. In comparing relative activity to protein concentration, the native and T154I, M158V L166S, and I223T variant enzymes show similar amounts of enzyme (~20 ng) required to convert half of the substrate to product while the R190I, I326T and K416E variants show that more enzyme is needed (~100 ng) (Fig. 5.8 B). For this mini-tRNA, both the M158V and R190I variants show reduced conversion of substrate to product as compared to the native enzyme in contrast to what was seen for the full-length tRNA. While the R190I variant looks similar to the I326T and K416E variants in terms of enzyme required to convert half of the substrate to product (~100 ng), the M158V variant requires approximately the same amount of protein (~30 ng) to reach the halfway mark, but it never achieves 100% conversion of substrate to product over the course of the reaction (Fig. 5.8 B).

When the proteins were incubated with tRNA lacking its terminal CMP and AMP residues to monitor the addition of both CTP and ATP a more complex picture results (Fig. 5.8 C). Again, at 10 ng of protein more than 95% of substrate is converted to product (containing both CMP and AMP) for the

native and T154I, M158V, L166S, R190I and I223T variants while both the I326T and K416E variants show less product (~75%). The analysis is more complicated in this case because of the presence of an intermediate substrate (that contains the CMP residue but lacks the terminal AMP residue). At 0.1 ng of protein for the native enzyme, 18% of the substrate was converted to the final product whereas 17% of the substrate was present as an intermediate species (lacking the AMP residue). Only slightly less activity was observed for variants T154I, M158V and L166S variants (approximately 10% final product with less than 10% intermediate product). In contrast, the R190I, I223T, I326T and K416E variants show no final product at 0.1 ng and only 0-6% intermediate species.



100	6	100	97	56	29	20	100	92	74	55	23	100	6	100	89	82	62	21	99	93	73	22	17	100	6
2	96	0	5	47	73	82	0	10	28	47	79	2	96	0	14	21	40	81	3	10	29	80	85	2	96

100	6	100	86	47	33	5	97	87	71	69	8	100	6	75	44	22	9	5	77	31	14	9	9	100	6
2	96	0	16	55	69	97	5	16	31	33	94	2	96	27	58	80	93	97	25	72	88	94	93	2	96

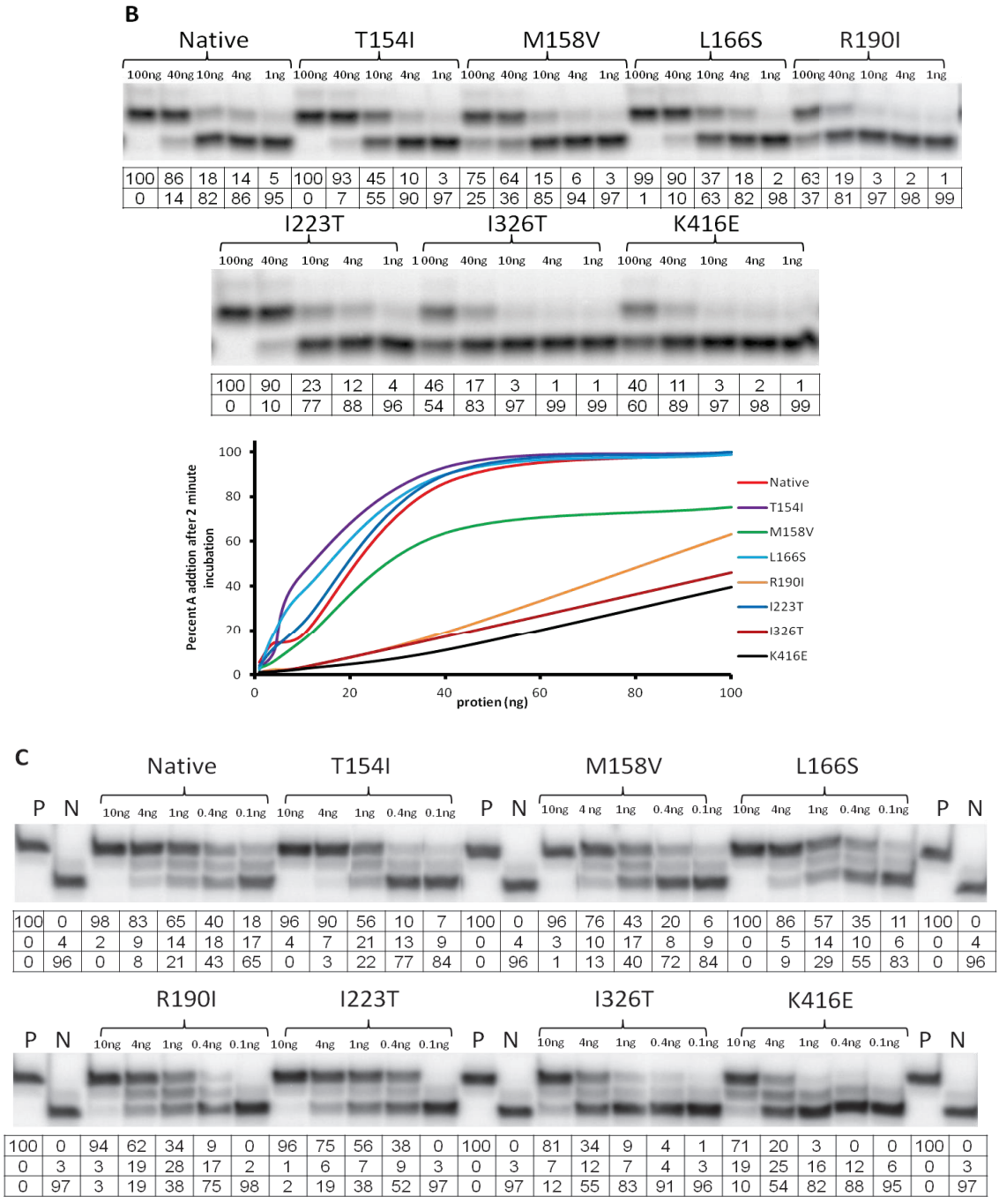


Figure 5.8: Enzymatic activity of human tRNA nucleotidyltransferases
The proteins indicated were incubated at 37^{o}C for 2 minutes with a (A) tRNA-NCC substrate or (B) mini tRNA-NCC substrate or (C) tRNA-NC substrate as described in materials and methods. The amount of enzyme added is indicated above each assay. The reaction products were separated by electrophoresis on 7M urea/12% polyacrylamide gels and visualized using the Typhoon™ TRIO Variable Mode Imager. P is the active enzyme positive control (tRNA with a complete CCA sequence), N is the boiled enzyme negative control (original tRNA template lacking the CA or A sequence). Products were analyzed by densitometry and results are listed under each assay and the rates were graphically plotted for each enzyme.

Since the substitutions T154I and M158V are found in Motif B which plays a role in discriminating between the binding of ribonucleotides and deoxyribonucleotides, the native and T154I and M158V variant enzymes were assayed for the ability to incorporate dAMP at the terminal position or dCMP at the penultimate position (Fig. 5.9). At all protein concentrations tested the T154I variant used dCTP or dATP most effectively as a substrate and the M158V variant used dATP or dCTP least, with the native enzyme showing a rate of incorporation between the variants.

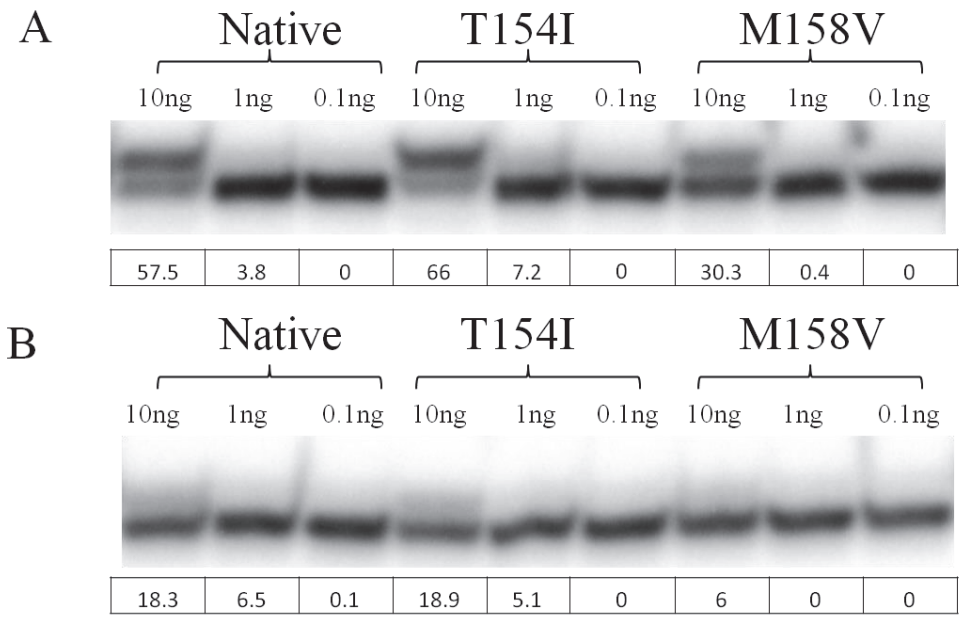
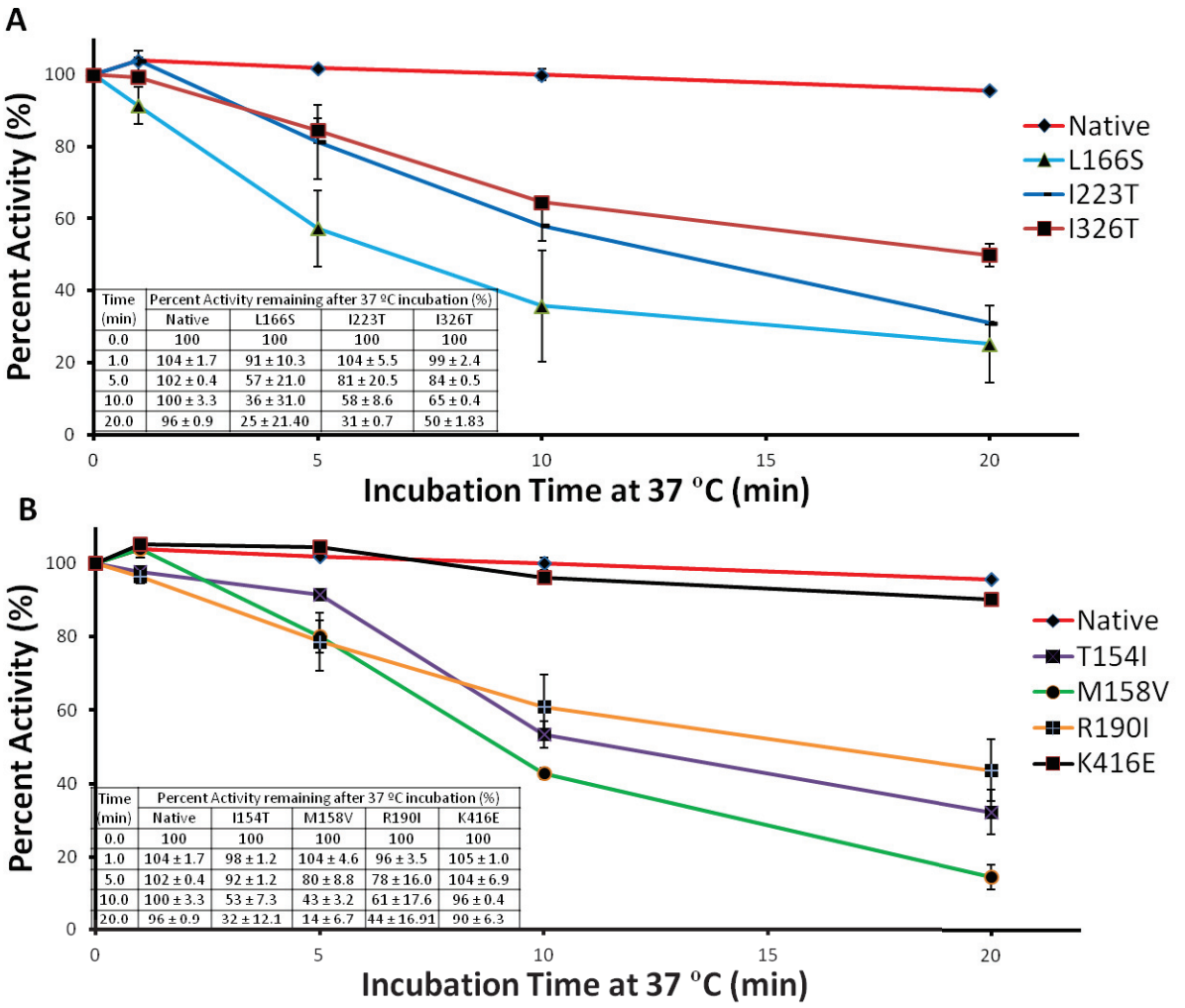


Figure 5.9: Deoxyribonucleotide incorporation of native and variant enzymes
The proteins indicated were incubated at 37^{o}C for 2 minutes with (A) tRNA-NC substrate and dCTP or (B) tRNA-NCC substrate and dATP as described in materials and methods. The amount of enzyme added is indicated above each assay. The reaction products were separated by electrophoresis on 7M urea/12% polyacrylamide gels and visualized using the Typhoon™ TRIO Variable Mode Imager. Products were analyzed by densitometry and the percent product is listed under each assay.

To determine if there was any correlation between the decrease in melting temperature seen in most variants and the activity levels measured, samples were checked for activity in the standard assay after preincubation at 37^{o}C for different times. The activity measured after each preincubation was compared to the activity in the absence of heat treatment (Fig. 5.10). For up to 10 min of preincubation, the native and K416E enzymes showed approximately 100% activity while the remaining enzymes

showed considerably less activity (40-70%). After 20 minutes preincubation, the native enzyme retained ~96% of its initial activity, the K416E variant retained about 90% of its activity, both the R190I and I326T variants retained about 50% of their activity and the T154I, M158V, L166S and I223T variants retained between 14% and 32% of their activity (Fig. 5.10). These data are in good agreement with what was seen in the thermal denaturation experiments (Fig. 5.3) and suggest a link between stability and activity.



Time (min)	Percent Activity remaining after 37 ºC incubation (%)			
	Native	L166S	I223T	I326T
0.0	100	100	100	100
1.0	104 ± 1.7	91 ± 10.3	104 ± 5.5	99 ± 2.4
5.0	102 ± 0.4	57 ± 21.0	81 ± 20.5	84 ± 0.5
10.0	100 ± 3.3	36 ± 31.0	58 ± 8.6	65 ± 0.4
20.0	96 ± 0.9	25 ± 21.40	31 ± 0.7	50 ± 1.83

Time (min)	Percent Activity remaining after 37 ºC incubation (%)				
	Native	I154T	M158V	R190I	K416E
0.0	100	100	100	100	100
1.0	104 ± 1.7	98 ± 1.2	104 ± 4.6	96 ± 3.5	105 ± 1.0
5.0	102 ± 0.4	92 ± 1.2	80 ± 8.8	78 ± 16.0	104 ± 6.9
10.0	100 ± 3.3	53 ± 7.3	43 ± 3.2	61 ± 17.6	96 ± 0.4
20.0	96 ± 0.9	32 ± 12.1	14 ± 6.7	44 ± 16.91	90 ± 6.3

Figure 5.10: Thermal stability activity assays of human tRNA nucleotidyltransferases
Activity, defined as AMP addition to the tRNA-NCC template in a standard assay, is measured after preincubation at 37°C for the times indicated. For the variants (I326T and K416e) with the least activity 1ng of protein was used while in all other cases 0.1 ng of protein was used. Reaction products were analyzed by densitometry and rates were calculated. Percent activity is the rate of a specific protein after preincubation divided by the rate for the same protein in the absence of preincubation multiplied by 100%.

5.5.2.2 Kinetic analyses

Detailed kinetics analyses were carried out on the native enzyme and L166S, I326T and K416E variants. To first determine the K_M and k_{cat} values for tRNA, AMP addition to a tRNA lacking only its terminal AMP residue was performed at varying concentrations of added baker's yeast tRNA at 37^{o}C. Intriguingly, the data for the native, L166S and K416E variant enzymes fit best to a sigmoidal plot (Fig. 5.11) when the reaction rate was plotted as a function of compared to the concentration of added tRNA. This suggested an allosteric effect involved in tRNA binding and cooperativity constants of 2.13 ± 0.46, 2.08 ± 0.34 and 2.50 ± 0.47 were calculated for the native, L166S and K418E variant enzymes, respectively (Table 5.1).

Table 5.1: Kinetics of tRNA nucleotidyltransferase enzymes

Construct	Varying substrate	K or Km (μM)	n (cooperativity coefficient)	turnover (s^{-1})
Native	tRNA-NCC	0.061 ± 0.008	2.13 ± 0.46	0.99 ± 0.08
	tRNA-NCC (TCEP)	0.114 ± 0.012	NA	1.325 ± 0.08
	mini tRNA-NCC	1.14 ± 0.34	1.85 ± 0.3	0.72 ± 0.2
	ATP	230 ± 40	NA	
	CTP	3.26 ± 0.8		
L166S	tRNA-NCC	0.097 ± 0.01	2.08 ± 0.34	2.2 ± 0.15
	ATP	155 ± 29	NA	
	CTP	5.09 ± 1.5		
I326T	tRNA-NCC	0.127 ± 0.032	NA	0.11 ± 0.01
	tRNA-NCC (TCEP)	0.087 ± 0.018	NA	0.12 ± 0.01
	ATP	181 ± 48	NA	
	CTP	13.5 ± 2.1		
K416E	tRNA-NCC	0.426 ± 0.035	2.50 ± 0.47	0.6693 ± 0.04
	tRNA-NCC (TCEP)	0.694 ± 0.10	NA	0.618 ± 0.05
	ATP	150 ± 29	NA	
	CTP	12.04 ± 1.4		

*NA =Not applicable

The potential for an allosteric interaction was suggested by the apparent dimerization described previously (Fig. 5.5, Fig. 5.7). Consistent with this, the data for the I326T variant which showed no apparent dimer (Fig. 5.5, Fig. 5.7), fit best to a hyperbolic initial velocity versus [S] plot suggestive of Michaelis-Menten kinetics (Fig. 5.11). When the kinetic data were collected at one tRNA concentration but with increasing concentrations of ATP or CTP the data fit best to a hyperbolic plot (Fig. 5.11) suggesting that tRNA binding promotes the allosteric transition.

To dissect further how tRNA binding could be linked to an allosteric effect, detailed kinetics analysis were carried out using the mini-tRNA substrate lacking its terminal AMP residue. As with the complete tRNA, the data generated using the mini-tRNA showed a sigmoidal plot and provided a cooperativity coefficient of 1.85 ± 0.3 for the native enzyme (Fig. 5.11, Table 5.1). In addition, the turnover number (0.72 s^{-1} ± 0.19) was in good agreement with that seen when the $tRNA^{Asp}$ was used as substrate although the binding constant (1.14 µM ± 0.36) was about 20 fold higher.

When the experiments were repeated in the presence of a reducing agent (Fig. 5.11), all kinetic data fit best to a hyperbolic plot suggestive of Michaelis-Menten kinetics. This suggests that the cooperativity is a consequence of the homodimer created by disulfide bond formation between monomers. Interestingly, K_M and k_{cat} values changed minimally between reducing and non-reducing conditions (Table 5.1). In either case, the K_M's determined for the $tRNA^{Asp}$ substrate for the native, L166S and I326T enzymes were approximately 100 nM while the K416E variant showed a seven-fold increase in K_M (Table 5.1). In contrast, the k_{cat} values for the native, L166S and K416E enzymes were essentially identical while the I326T variant showed a ten-fold reduction in turnover number (Table 5.1). For all enzymes analyzed, K_M values ranged between approximately 150-230 µM and 3-13 µM for ATP and CTP, respectively (Table 5.1).

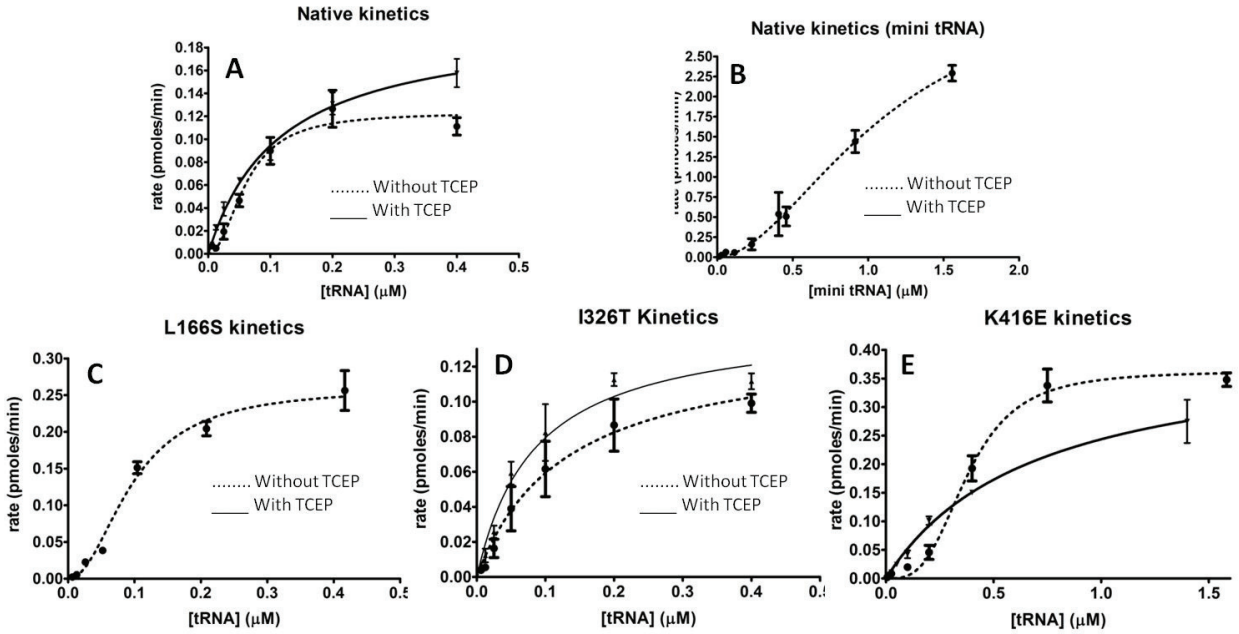


Figure 5.11: Enzyme kinetics of human tRNA nucleotidyltransferases
The rate AMP incorporation was plotted in the presence of increasing amounts of tRNA-NCC template A) native, C) L166S, D) I326T, E) K416E protein or B) native enzyme with mini-tRNA-NCC template. Assays A, D, and E were repeated in the presence of reducing agent. The data were analysed by a least-squares method using the equation for Michaelis–Menten kinetics, v =Vmax [S]/(K_M+[S]) or the Hill equation for cooperativity kinetics v = Vmax $[S]^n$/(K^n+$[S]^n$). The line represents the best fit to the data.

5.6 Discussion

Mitochondrial diseases represent a large family of illnesses showing diverse phenotypes (Tonin and Entelis, 2014). These diseases may be defined by mutations in either the mitochondrial or nuclear genome. Specifically, mitochondrial diseases have been linked to mutations in tRNAs and tRNA modifying enzymes (Abbott *et al.*, 2014). Recently, multiple patients showing congenital sideroblastic anemia with B cell immunodeficiency, period fevers and developmental delay (SIFD) have had the mutations responsible for this disease mapped to the TRNT1 gene (Chakraborty *et al.*, 2014). This gene codes for tRNA nucleotidyltransferase which is responsible for the addition of the 3'-terminal cytidine-cytidine-adenosine sequence to both mitochondrial and cytosolic tRNAs. The human tRNA nucleotidyltransferase is a Class II tRNA nucleotidyltransferase which contains five highly conserved

motifs (A-E) in the amino terminal half of the protein (Li *et al.*, 2002). In contrast to the conservation in the head and neck regions, the remaining body and tail regions of the Class II tRNA nucleotidyltransferases have little primary sequence conservation (Yakunin *et al.*, 2004), although the available crystal structures show these regions might be involved in tRNA binding (Tomita *et al.*, 2004, Xiong and Steitz, 2004).

The seven mutations identified in patients with SIFD (Chakraborty *et al.*, 2014) have been mapped to motifs B (T154I or M158V), C (L166S), to the region between motifs C and D (R190I), to the main region between motifs D and E (I223T), and to the more carboxy-terminal region (I326T or K416E) far from the well characterized catalytic motifs. The mutations mapping to each of these distinct domains will be discussed individually to explore their possible effects on tRNA nucleotidyltransferase structure, stability, activity and localization that may be linked to the disease phenotype.

5.6.1 Mutations found within conserved Motif B

Motif B is responsible for ensuring that ribonucleotides and not deoxyribonucleotides are incorporated into the tRNA substrate (Li *et al.*, 2002). A specific highly conserved RRD sequence has been shown to be required for this selectivity (Li *et al.*, 2002). The central arginine has been shown to form a hydrogen bond with the hydroxyl group of an incoming ribonucleotide resulting in the preference for ribonucleotides over deoxyribonucleotides. An arginine to isoleucine substitution at the central residue of the triplet resulted in the loss of the hydrogen bonding potential and removed the selection of ribonucleotides over deoxyribonucleotides for incorporation (Cho *et al.*, 2007). In motif B of the human enzyme (Fig. 5.12), the residues T154 and M158 are two and six amino acids, respectively from the conserved RRD sequence and the threonine at this position also is absolutely conserved.

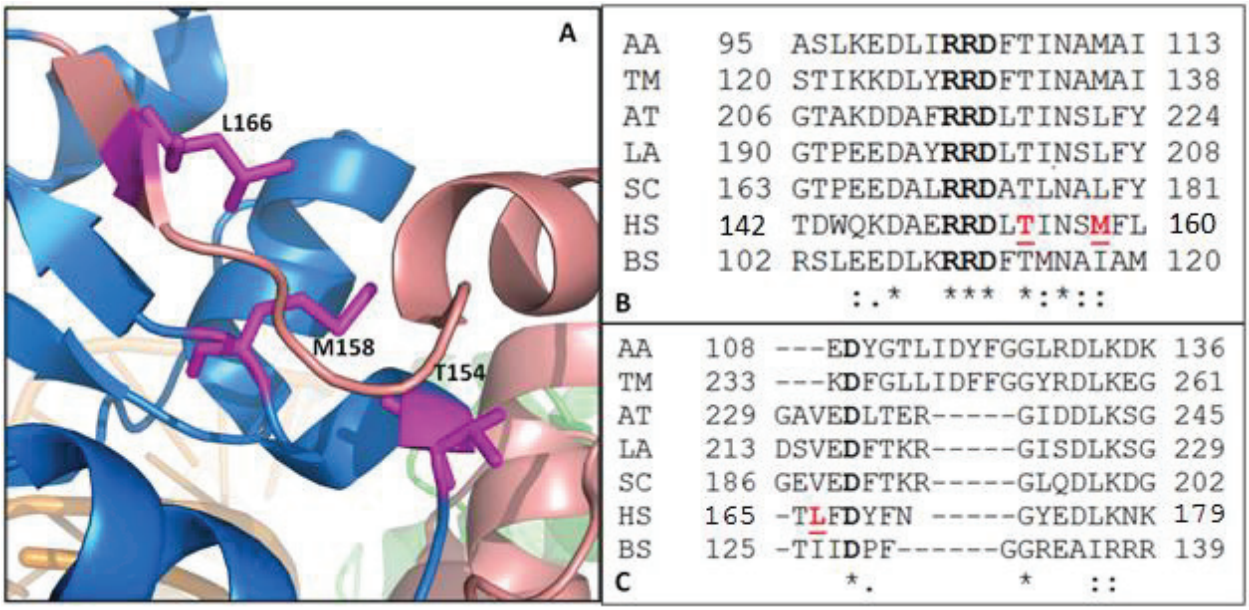


Figure 5.12: Motif B and C of human tRNA nucleotidyltransferase
(A) Conserved motifs B (blue) and C (salmon) are shown in ribbon form. Amino acids of interest are shown as sticks (pink) and are identified by the one letter code. (B) Alignment of Motif B and (C) Motif C. The alignment was performed with ClustalW with Class II tRNA nucleotidyltransferases from *Thermotoga maritima***,** TM (3H39); *Aquifex aeolicus*, AA (1VFG); *Saccharomyces cerevisiae*, SC (NP_011095); *Arabidopsis thaliana*, AT (NP_173680); *Lupinus albus*, LA (AAB03077); *Homo sapiens*, HS (10U5_A); *Bacillus stearothermophilus*, BS (Q7SIB1). Invariant amino acids are noted by (*), conserved amino acids are noted by (:) and similar amino acids are noted by (.). Catalytically important residues are bold (black) and amino acids described here are bold and underlined (red).

Although a methionine is not absolutely conserved at the position corresponding to 158 in the human enzyme, this position is typically a large aliphatic hydrophobic residue. Given the conservation at these two positions in motif B, it seems reasonable to predict that they also may be involved in nucleotide selection. When tested initially for activity in the standard enzyme assay for both AMP addition and the final CMP addition (Fig. 5.8), both variants showed activity consistent with that of the native enzyme although the M158V variant was slightly less active. This is surprising given that a single phenylalanine to alanine substitution in Motif B of the *Aquifex aeolicus* enzyme have been shown to decrease enzyme activity by about 10-fold (Tomita *et al.*, 2004). Perhaps the substitution of a methionine by another aliphatic hydrophobic amino acid retains enough conservation in physical properties that activity is maintained. However, one would anticipate that the conversion of the

conserved threonine to isoleucine (from small and hydrophilic to large and hydrophobic) would have a more dramatic effect on activity. Activity approximately that of the native enzyme was seen when a mini-tRNA was used as the template for AMP addition (Fig. 5.8). This indicated that the disease phenotype likely does not result from a hypomorphic effect due to reduced enzyme activity as we saw with the temperature-sensitive phenotype in yeast (Goring *et al.*, 2013) or as was seen for mutations in TRNT1 which cause retinitis pigmentosa with erythrocytic microcytosis (DeLuca *et al.*, 2016). In the standard *in vitro* assays, the enzymes are provided with specific CTP and ATP substrates but *in vivo,* there may be competition between ATP and dATP or CTP and dCTP for binding at the active site. The native enzyme will select for the ribonucleotide triphosphates but the T154I and M158V variants may alter the nucleotide triphosphate binding region of the enzyme to allow a certain level of misincorporation of dATP or dCTP which would generate a population of tRNAs which would not be aminoacylated and which would affect the efficiency of protein synthesis. To address this possibility the standard *in vitro* assay was carried out with the tRNA-NC template in the presence of the native enzyme or the T154I or M158V variant but with dCTP instead of CTP as substrate, or with the tRNA-NCC template with dATP instead of ATP as substrate. When compared to the native enzyme the M158V variant was even less efficient at incorporating dCTP (Fig. 5.9 A) or dATP (Fig. 5.9 B). In contrast, the T154I variant showed an apparent increase in both dATP and dCTP incorporation (Fig. 5.9). These observations suggest that the change at position 158 from methionine to valine does reduce activity to extent degree when either NTPs or dNTPs are used as substrates so that this decrease in activity may be linked to the phenotype observed. In contrast, the T154I variant showed nucleotide incorporation equal to or better than that of the native enzyme most notably for AMP incorporation (Fig. 5.8) and for dAMP and dCMP incorporation (Fig. 5.9). Perhaps in this case, it is not the reduced incorporation of NTPs that leads to the phenotype but the misincorporation of dNTPs.

Are there other effects of these two amino acid substitutions that could lead to the disease phenotype? The analysis of the human crystal structure reveals that both T154 and M158 are buried within the protein such that one would anticipate that the resulting hydrophobic substitutions would not dramatically alter protein structure (Fig. 5.12). This appears to be the case as no structural change as measured by circular dichroism spectroscopy (Fig. 5.2), fluorescence spectroscopy (Fig. 5.4) or gel filtration chromatography (Fig. 5.5). However, both variants show reduced thermal stability at nearly 5°C below that of the native enzyme (Fig. 5.3). In good agreement with the thermal denaturation experiments, preincubation of these proteins at 37°C for 20 minutes results in a decrease in enzyme activity by about three and seven-fold for T154I and M158V, respectively as compared to the native enzyme (Fig. 5.10). This suggests either that some population of the protein molecules unfold completely at this temperature and lose activity or that all of the protein molecules unfold to some degree such that they have reduced activity. In human cell lines, Chakraborty *et al.*, (2014) show that cells with the T154I substitution retained about two-thirds of the tRNA nucleotidyltransferase protein as compared to native cells as measured by Western blotting. Recently (Sasarman *et al.*, 2015), an alanine to valine substitution was identified in Motif B two amino acids amino terminal to the conserved RRD sequence. This patient suffered a first crisis of lactic acidosis at 18 days of age, showed the characteristics of SIFD and died of cardiac arrest during an episode of fever and acute acidosis at 21 months of age. Immunoblot analysis revealed about a 10-fold drop in tRNA nucleotidyltransferase in the variant as compared to the native enzyme. Taken together with our results, this suggests that protein stability may be an important parameter in defining the disease phenotype. One can speculate that the disease phenotype results from a combination of changes to structure, stability and activity. In the case of the M154V variant, there is reduced stability and loss of activity and for the T154I variant, there is reduced stability and increased incorporation of dNTPs as compared to the native enzyme.

Taking all of these data into account it is interesting that patients carrying the T154I variant are neither transfusion nor intravenous immunoglobulin–dependent suggesting that the T154I variant is less severe than the M158V variant (Chakraborty *et al.*, 2014) even though the amino acid substitution looks more dramatic (hydrophilic to hydrophobic) and seems to have an effect on the discrimination between ATP and dATP (at least in the *in vitro* assay). To test the hypothesis that the level of enzyme activity is reduced because the M158V variant protein is less stable than native (and not because it is less active than native), it will be useful to check for tRNA nucleotidyltransferase protein levels in cell lines expressing the M158V substitution.

5.6.2 Mutation found within conserved Motif C

Motif C acts as a connection point between the head and neck region of tRNA nucleotidyltransferases (Li *et al.*, 2002) and may act as a flexible spring element modulating the relative orientation of the head and body domains to help accommodate the growing 3' end of the tRNA substrate (Ernst *et al.*, 2015). Studies have shown that changing the conserved aspartate (Fig. 5.12) in Motif C to an alanine resulted in a 15-fold decrease in turnover for the addition of ATP to the tRNA-NCC template, but did not affect the binding of the tRNA substrate (Ernst *et al.*, 2015). A similar observation was made in *Saccharomyces cerevisiae* tRNA nucleotidyltransferase when the glutamate residue (adjacent to the conserved aspartate) was changed to a phenylalanine or lysine resulting in a large decrease in activity leading to a temperature-sensitive phenotype (Shan *et al.*, 2008). The nearly 100-fold decrease in enzyme activity was sufficient to generate a temperature-sensitive phenotype in yeast indicating a hypomorphic effect resulting from this reduction in activity (Goring *et al.*, 2013). Leucine at position 166 represents a conserved large aliphatic hydrophobic amino acid found in Motif C among Class II tRNA nucleotidyltransferases that is two amino acids amino-terminal to the important aspartic acid residue. It is conceivable that a change of the leucine at this position to a serine residue which is smaller and more hydrophilic could be linked to the phenotype. In contrast to what was observed when

amino acids were changed in Motif C of the yeast (Goring *et al.*, 2013) and human (Ernst *et al.*, 2015) enzymes, the ability of the L166S variant to add CTP and ATP to the 3' end of the tRNA lacking the terminal AMP or both CMP and AMP residues appears to be equivalent to that of the native enzyme (Fig. 5.8). Detailed kinetic analyses further support this observation, generating equivalent turnover numbers for both the L166S variant and the native enzyme (Table 5.1) with nearly equivalent binding constants for tRNA, ATP and CTP (Table 5.1). In fact, the turnover number for the L166S variant is two times greater than that of the native enzyme for AMP addition. Previous studies of variants within Motif C showed changes in k_{cat} but no changes in K_M values for tRNA binding (Leibovitch, in preparation, Ernst *et al.*, 2015) for the variant proteins. So here while the change in turnover number is small, it may suggest that the L166S variant is more efficient than the native enzyme. As with the Motif B variants, this amino acid substitution does not suggest that a direct affect on enzyme activity is the cause of the observed phenotype.

To help identify any change to the L166S variant, secondary and tertiary structure analyses using dichroism circular and tryptophan fluorescence spectroscopies, respectively were performed. As in Motif B, negligible changes in secondary (Fig. 5.2) or tertiary (Fig. 5.4) structure were observed as compared to the native enzyme. Gel filtration chromatography (Fig. 5.5) and SDS-PAGE (Fig. 5.7) indicated that quaternary structure also was not affected. However, as observed for the Motif B variants, the stability of the L166S protein appears to be compromised showing about a 7°C decrease in Tm as compared to the native enzyme (Fig. 5.3). Pre-incubation of the L166S variant at 37°C prior to assaying for activity further supports the loss of thermostability. The L166S variant lost 10% of its activity after one minute of pre-incubation and almost half of its activity within five minutes (Fig. 5.10). Based on this experiment, the L166S variant is the most thermolabile as all of the other enzymes retain full activity after one minute and more than 80% of activity at five minutes (Fig. 5.10). Even though L166 has no catalytic role, it may be directly involved in stabilizing the connection between the head and the neck

portions of the protein (Li *et al.*, 2002). In human cell lines (Chakraborty *et al.*, 2014) cells with the L166S substitution in combination with the T154I substitution showed significantly less protein than the native enzyme. This destabilization of the protein may result in reduced levels in the cell. Sasarman *et al.*, (2015) showed that a heterozygous tyrosine to phenylalanine substitution in Motif C in consort with a heterozygous alanine to valine substitution at position 148 resulted in a patient with motor development delay and normal cognitive development. This patient showed nearly normal levels of TRNT1 by immunoblotting and only moderately reduced mitochondrial protein synthesis as compared to the controls (Sasarman *et al.*, 2015). This suggests that the milder form of the disease results from a moderate change in protein structure and function. For the L166S variant it appears that decreased stability leading to reduced activity and not an intrinsic reduction of activity is responsible for the SIFD phenotype.

5.6.3 Mutation found between conserved Motifs C and D

Motif D contains the well conserved (E/D)DxxR sequence that helps to accommodate incoming CTP or ATP substrates (Li *et al.*, 2002). The conserved glutamate, aspartate and arginine residues recognize CTP or ATP by forming Watson-Crick-like base pairs with the nucleotides (Yue *et al.*, 1996, Shi *et al.*, 1998, Cho *et al.*, 2007). In converting from CTP to ATP binding the active site of the enzyme rearranges to allow the appropriate interactions with the incoming ATP or CTP (Li *et al.*, 2002). R190 is three residues amino-terminal of the conserved glutamate but is not generally conserved among Class II tRNA nucleotidyltransferases (Fig. 5.13). Interestingly, the crystal structure of human tRNA nucleotidyltransferase reveals that R190 forms a hydrogen bond with the conserved glutamate of the EDxxR motif in Motif D (Fig. 5.13). As this glutamate hydrogen bonds with the arginine residue of the motif to orient it for ATP or CTP binding during the rearrangement of the active site (Li *et al.*, 2002), R190 also could participate in the rearrangement of the ATP/CTP binding pocket and in the binding of

incoming NTPs. It would be interesting to compare the crystal structures of the native and R190I variants in the presence of each nucleotide triphosphate to see if any differences could be observed.

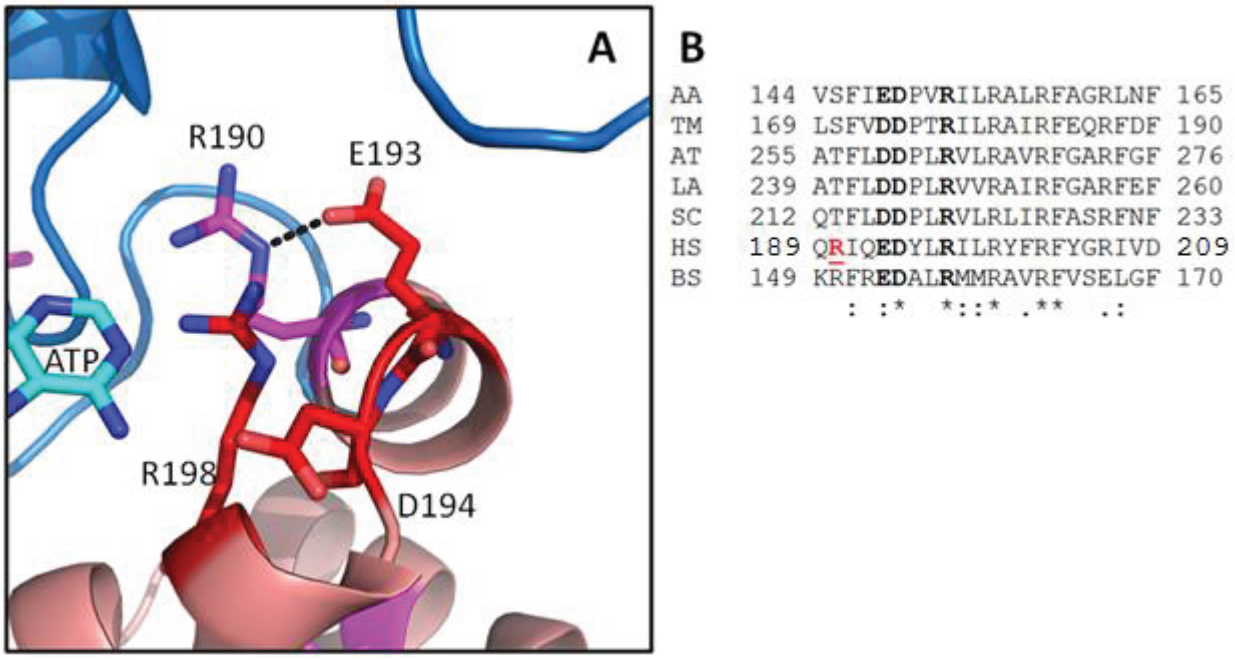


Figure 5.13: Motif D of human tRNA nucleotidyltransferase
(A) Conserved motif D (salmon) is shown as a ribbon. Amino acids of interest are shown in sticks (pink or red) and are identified by the one letter code. (B) Alignment of Motif D. The alignment was performed with ClustalW with Class II tRNA nucleotidyltransferase from *Thermotoga maritima*, TM (3H39); *Aquifex aeolicus*, AA (1VFG); *Saccharomyces cerevisiae*, SC (NP_011095); *Arabidopsis thaliana*, AT (NP_173680); *Lupinus albus*, LA (AAB03077); *Homo sapiens*, HS (10U5_A); *Bacillus stearothermophilus*, BS (Q7SIB1). Invariant amino acids are noted by (*), conserved amino acid are noted by (:) and similar amino acids are noted by (.). Catalytically important residues are bold (black) and amino acids of described in this study are bold and underlined (red).

Surprisingly, replacing the arginine at position 190 with isoleucine does not appear to affect enzyme activity under standard assay conditions except for CMP addition at the lowest enzyme concentration tested (Fig. 5.8). When the activity assays were repeated with a mini-tRNA containing only the aminoacyl stem and the TψC stem and loop, a difference in AMP addition that was not observed with the full length tRNA substrate became evident. While the native and R190I enzymes added AMP at similar rates at the 4 ng enzyme level when using the full length tRNA as a substrate, the R190I variant yielded only about 20% of the activity compared with the native enzyme when the mini-tRNA substrate

was used (Fig. 5.8). It would be interesting to repeat the experiment with a mini-tRNA designed as a substrate for CMP addition to determine if the R190I substitution would reduce CMP incorporation. If both AMP and CMP incorporation are altered, this might suggest that the R190I substitution causes a change in the protein that alters mini-tRNA binding since it has fewer contact points with the enzyme than does the full-length tRNA. A more detailed kinetic analysis of activity with the mini-tRNA substrate may provide insight into this question. While the native enzyme showed an approximately 10-fold increase in K_M for the mini-tRNA as compared to the full-length tRNA it would be interesting to determine the K_M values for the R190I variant. Perhaps there is no apparent change in kinetic parameters for the native and R190I variant enzymes with the full-length tRNA because any changes introduced by the amino acid change are not significant enough to affect tRNA binding and catalysis. In contrast, the changes may lead to decreased binding and catalysis for the mini-tRNA. This may be physiologically relevant as some human mitochondrial tRNAs, *e.g.*, $tRNA^{Ser(AGY)}$ have non-canonical structures. Recently (Sasarman *et al.*, 2015), two patients were identified with mitochondrial encephalomyopathies largely due to impaired mitochondrial translation resulting from defective CCA addition to mitochondrial $tRNA^{Ser(AGY)}$ caused by mutations mapped to the TRNT1 gene coding for tRNA nucleotidyltransferase. Therefore, it is apparent that mutations which make the enzyme less able to use the non-standard mitochondrial tRNAs as substrates can lead to the observed phenotype.

In terms of the biophysical characteristics of the R190I variant, structural analyses revealed negligible changes in secondary (Fig. 5.2), tertiary (Fig. 5.4) or quaternary (Fig. 5.5, Fig. 5.7) structure as compared to the native enzyme, however thermal denaturation shows the a drop of about 6°C in Tm for the variant as compared to the native enzyme (Fig. 5.3). The lower stability is likely a direct result of minor reorganization of the ATP/CTP binding pocket due to a loss of the hydrogen bond between R190 and the catalytic glutamate at position 193. The change in hydrophobicity at position 190 also poses a problem since adjacent residues are polar and R190 is normally water exposed. Burying of the isoleucine

could cause the binding pocket to be moderately altered reducing activity at higher temperatures. A twenty minute incubation at 37°C shows about a 50% drop in activity confirming the loss of structure. The increase in hydrophobicity and loss of the hydrogen bond likely cause the isoleucine at position 190 to be more buried, partially destabilizing the protein. Even though the altered stability is an unlikely explanation for reduced activity, it may participate in reducing CCA activity in the cell as this mutation in human cell lines results in about 30% lower protein concentrations as compared to the native enzyme (Chakraborty *et al.*, 2014). In an effort to change the specificity of the motif D binding pocket, triple or double mutants were created in the EDxxR motif in the *B. stearothermophilus* tRNA nucleotidyltransferase with reduced activity demonstrating the importance of this region in NTP binding and activity (Cho *et al.*, 2007). We hypothesize that the loss of the hydrogen bond to E193 resulting from the conversion of R190 to isoleucine restricts movement in Motif D altering the specificity for NTPs in the binding pocket of the enzyme. We propose that R190 may also directly interact with the tRNA substrate sugar phosphate backbone and help properly anchor the tRNA substrate via electrostatic interactions. Activity assays with the modified mini-tRNA suggest that activity within mitochondria might also be further compromised as mitochondrial tRNA may lack regions which could interact with the enzyme. We propose that the combination of altered specificity, activity and stability for the R190I variant may lead to the SIFD phenotype in humans.

5.6.4 Mutations found within poorly characterized regions of human tRNA nucleotidyltransferase

5.6.4.1 I223T

Isoleucine 223 is situated in an α-helix buried deep within the neck portion of the protein between Motif D and Motif E (Fig. 5.14). Although much is known about the function of Motif D and there is some information as to the role of Motif E, the amino acids between these regions are not conserved among Class II tRNA nucleotidyltransferases. This specific region likely acts as a simple

connection between Motif D and Motif E in the neck portion of the protein. As seen for the four previously characterized variants there were no obvious changes in secondary (Fig. 5.2) or tertiary (Fig. 5.4) structure for the I223T variant as compared to the native enzyme. Intriguingly, although the I223T

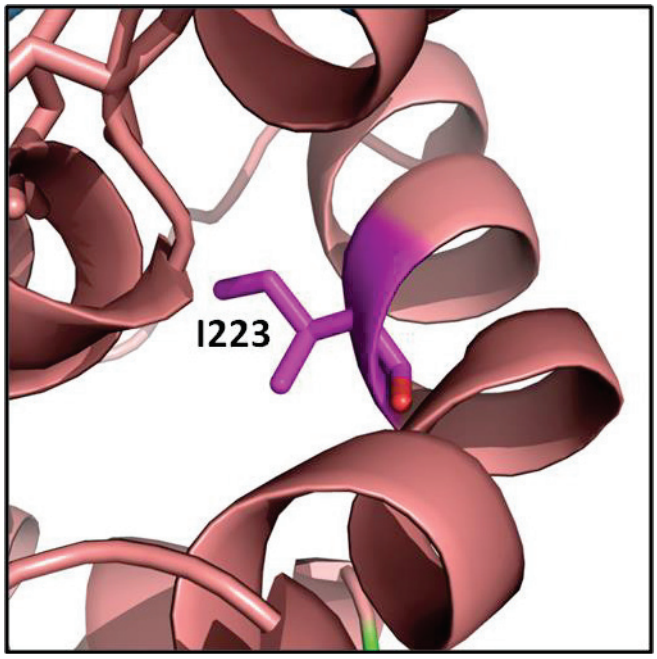


Figure 5.14: Position of residue I223 between motifs D and E of human tRNA nucleotidyltransferase
Human tRNA nucleotidyltransferase polypeptide backbone is shown in ribbon (salmon). Amino acid I223 is shown in sticks (pink).

variant showed no structural changes (except for a minor change in fluorescence maximum), SDS-PAGE showed an anomalous migration distance for this protein. Although mass spectrometry confirmed a mass of 48.5 kDa as expected, the protein migrated differently than all of the other proteins with an apparent molecular mass of about 50 kDa (Fig. 5.7 A). This may indicate that the amino acid substitution does somehow alter how the protein unfolds under the denaturing conditions of SDS-PAGE. Under the non-denaturing conditions of gel filtration chromatography (Fig. 5.5) the protein elutes with apparent molecular masses of 52 kDa and 162 kDa where the monomeric species (52 kDa) corresponds well with the apparent molecular mass of native protein but the dimeric species (162 kDa) seems much larger (~140 kDa for native enzyme). SDS-PAGE under non-reducing conditions also shows I223T with more

intense bands around 116 kDa suggesting that the substitution at position 223 may alter quaternary structure but not the ability to form a dimer. Perhaps this amino acid substitution causes some type of minor conformational change in the protein such that it does not unfold as the other proteins do under reducing and denaturing conditions. We have seen this type of anomalous slow migration of an Arabidopsis tRNA nucleotidyltransferase variant which also showed altered localization and reduced activity (Leibovitch *et al.*, 2013).

Thermal denaturation (Fig. 5.3) revealed that the Tm for this variant was about 5°C lower than the native enzyme in good agreement with what had been observed for the four variants described above. I223T showed activity similar to that of the native enzyme (Fig. 5.8) and only minor destabilization on preincubation at 37°C with retention of fully one third of its activity after 20 minutes at 37°C. Again, while reduced stability may play a role in the onset of SIFD in patients carrying this mutation, a role in altered protein localization cannot be excluded. The Arabidopsis tRNA nucleotidyltransferase showed altered migration on SDS-PAGE and reduced targeting to mitochondria *in vivo* (Leibovitch *et al.*, 2013). Experiments with human cell lines should be able to address these possibilities.

5.6.4.2 I326T and K416E

The I326T and K416E substitutions are found within the body and tail portions of the protein about which not much is known (Yakunin *et al.*, 2004). The body and tail domains of tRNA nucleotidyltransferase are believed to be involved in tRNA recognition and binding (Li *et al.*, 2002). The body and tail regions appear to exclusively interact with the sugar-phosphate backbone of the tRNA substrate as determined by observations using the available crystal structures (Tomita *et al.*, 2004, Xiong and Steitz, 2004). The available crystal structure of the human enzyme indicates that residue I326 is found in an α-helix buried in the body portion while residue K416 is found at the end of an α-helix and

exposed to the aqueous environment (Fig. 5.15). As had been observed for all of the previously characterized variants there were no major changes in secondary (Fig. 5.2) or tertiary (Fig. 5.4) structure for either variant as compared to the native enzyme. Thermal denaturation revealed that the Tm for the I326T variant was about 7°C lower than the native enzyme in good agreement with what had been observed for all of the variants described previously. In contrast, the K416E variant showed a Tm similar to that of the native enzyme (Fig. 5.3).

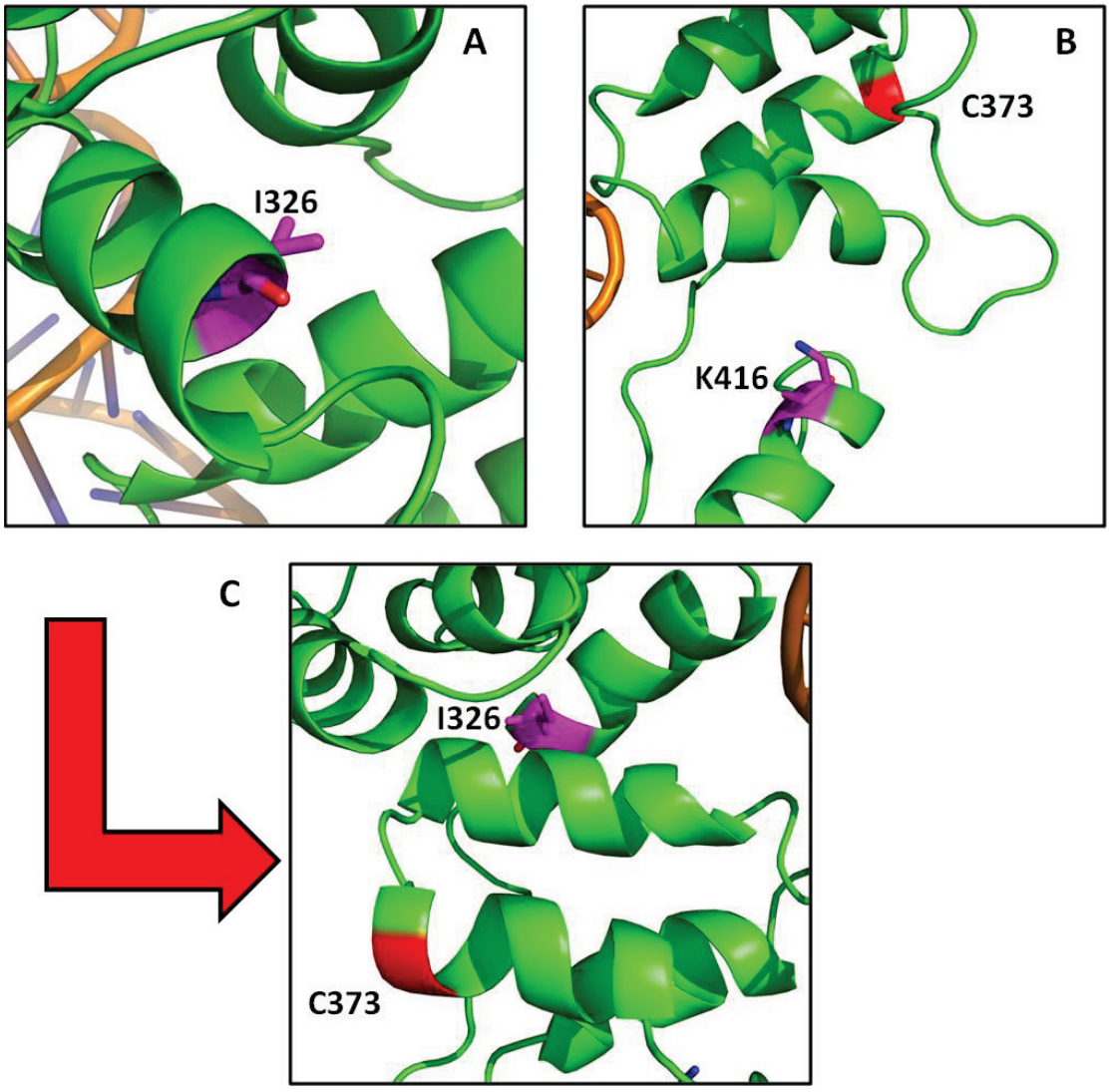


Figure 5.15: Body and tail domains of human tRNA nucleotidyltransferase
Human tRNA nucleotidyltransferase polypeptide backbone is shown in ribbon (green). Amino acids of interest are shown in sticks (pink or red) and are as indicated.

The I326T variant does not show any apparent dimer on gel filtration chromatography (Fig. 5.6) while the K416E variant shows a greater ratio of dimer to monomer than does the native or any other variant enzyme (Fig. 5.6). As the available human crystal structure showed a C373-C373 disulfide bond involved in dimer formation (Augustin *et al.*, 2003, Ernst *et al.*, 2015) the gel filtration experiments were repeated in the presence of a reducing agent and as expected the dimers disappeared (Fig. 5.6). These observations were confirmed by SDS-PAGE in the presence and absence of a reducing agent (Fig. 5.7). Again all proteins except the I326T variant showed a higher molecular weight species in the absence, but not the presence of the reducing agent (Fig. 5.6). Given the position of the disulfide bond in the tail of the protein (Fig. 5.15), it could be argued that amino acid substitutions such as those at positions 326 and 416 could cause a rearrangement of this portion of the enzyme to either decrease (in the case of I326T) or increase (in the case of K416E) disulfide bond formation. Given the predicted role of the tail portion of the protein in tRNA recognition and binding it would seem likely that rearrangements of the tail region also could affect tRNA binding. The I326T variant which does not form a homodimer, binds poorly to tRNA (Fig. 5.6 B) while the K416E variant which forms the homodimer at a much greater ratio than the native enzyme does not appear to bind tRNA at all (Fig. 5.6 C). The apparent K_D values for tRNA binding to the I326T variant and the K416E variant are 40 fold and 125 fold higher than that of the native enzyme, respectively. A similar increase in K_D was observed for a C-terminal variant of the Arabidopsis tRNA nucleotidyltransferase where the final four C-terminal positively charged amino acids were converted to negatively charged amino acids (Leibovitch *et al.*, 2013). Experimental evidence has previously shown that the C-terminus may play a role in tRNA binding as swapping the tail of *E. coli* tRNA nucleotidyltransferase with that of poly (A) polymerase caused tRNA slippage which suggested that the function of the tail is to hold and anchor the tRNA substrate in place (Betat *et al.*, 2004). In the case of the I326T variant, the substitution likely causes adjacent helices to move causing misalignment of the tRNA substrate with positively charged amino acids responsible for recognition and anchoring of the

tRNA substrate (Fig. 5.15). The lysine at position 416 however, may be directly interacting with the negatively charge sugar phosphate backbone of the tRNA substrate and may have a central role in tRNA binding. Therefore, the K416E substitution prevents direct contact with the tRNA substrate by electronegative repulsion via electrostatic interactions.

Treating the native enzyme, or the I326T or K416E variant with RNAseA in the presence or absence of reducing agent clearly shows the relative amounts of dimer and monomer. While the I326T variant forms only a monomer in the presence or absence of tRNA (Fig. 5.6), the native protein is predominantly monomeric (Fig. 5.6) and the K416E variant exists as both a monomer and a dimer (Fig. 5.6). Even though the dimer species is lost under reducing conditions, tRNA binding is not affected (Fig. 5.6). Detailed kinetic analyses of the native and I326T, and K416E variant enzymes support these observations. Under non-reducing conditions, the native and K416E variant enzymes show cooperativity whereas the I326T shows none (Fig. 5.11). Under reducing conditions, apparent cooperativity was lost for all three proteins (Fig. 5.11).

The physiological relevance (if any) of dimer formation remains to be determined. Since the cytosol is a reducing environment (Go *et al.*, 2008), one would not expect dimerization. However, the mitochondrion provides an oxidizing environment where dimerization may occur. It will be useful to examine tRNA nucleotidyltransferase isolated from mitochondria and the cytosol to determine if these proteins exist in monomeric or dimeric conformations. To form this thiol linkage, the monomers must be in close proximity (Trivedi *et al.*, 2009). Perhaps the relative mitochondrial tRNA nucleotidyltransferase concentration is greater than the cytosolic concentration. The relative amounts of dimers and monomers seen under the experimental conditions here suggest that protein dimerization is through weak interactions and is short lived. It is interesting that the K416E variant enzyme which has the greatest amount of dimer is also the variant with the highest Tm (Fig. 5.3) and the greatest remaining activity upon incubation at 37°C (Fig. 5.10). However, the I326T variant which shows the least amount of

dimer has neither the lowest Tm (Fig 5.3) nor the least retained activity after incubation at 37°C (Fig. 5.10). So there is not a direct link between dimerization and protein stability.

A detailed kinetic analysis was performed on the I326T and K416E variants to see if any of the features of tRNA binding or altered quaternary structure affected enzyme activity. The turnover number for the I326T variant is approximately 10-fold that of the native enzyme while the K_M for tRNA binding was similar to the native enzyme (Table 5.1). In contrast, the K_M for tRNA was 5-7 fold higher for the K416E variant as compared to the native enzyme and the turnover number was only reduced by approximately 50% as compared to the native enzyme (Table 5.1). In each case the K_M vlaues for ATP and CTP remained essentially the same among the native and variant enzymes (Table 5.1). So while the I326T replacement dramatically reduces catalysis, it has little effect on tRNA binding at the active site. In contrast, the K416E change does not really alter catalysis but increases the K_M for tRNA binding. These data suggest that in the case of the I326T variant it is not simply the poor binding of the tRNA substrate that is directly responsible for the 10-fold drop in turnover. There must be some other factor at play.

As with the other variants it seems that reduced stability of the I326T variant may be linked to SIFD. In contrast, the K416E variant represents the only variant that has structural stability and retained activity comparable to the native enzyme. In this example, it is not simply a loss of stability that leads to the phenotype, but perhaps may reflect poor substrate binding. Again, experiments to look at the levels and localization of these variants may be informative as to defining how SIFD results. We previously have shown that altering or removing the C-terminal eight amino acids of the Arabidopsis tRNA nucleotidyltransferase allows the enzyme to maintain activity but alters tRNA binding by about nine-fold and changes the intracellular distribution of the resulting protein (Leibovitch *et al.*, 2013). Recently (DeLuca *et al.*, 2016), it has been shown that a patient suffering from retinitis pigmentosa with erythrocytic microcystis carries a frame shift mutation in the TRNT1 gene that removes the last seven amino acids of the human tRNA nucleotidyltransferase and replaces them with ten other amino acids.

These data point to some role for the carboxy terminus of tRNA nucleotidyltransferase in enzyme structure or function. Further studies are required to see if SIFD results from a combination of a change in activity and/or localization.

5.7 Conclusion

In this chapter, we suggest that the SIFD phenotype is linked to poor stability of the T154I, L166S and I223T variant proteins, and to a combination of both reduced stability and catalytic function in both the M158V and R190I variants. However, the phenotype created by the I326T variant may be related to a multitude of problems which include reduced stability, low catalytic activity, poor affinity for the tRNA substrate and inability to dimerize. In contrast, the poor affinity of the K416E variant for tRNA may lead to the SIFD phenotype. Localization studies *in vivo* are necessary to fully explore whether changes in higher order structure (I223T and I326T) or tRNA binding (I326T and K416E) have any influence on the cellular distribution of tRNA nucleotidyltransferase in humans.

5. 8 Supplementary figures

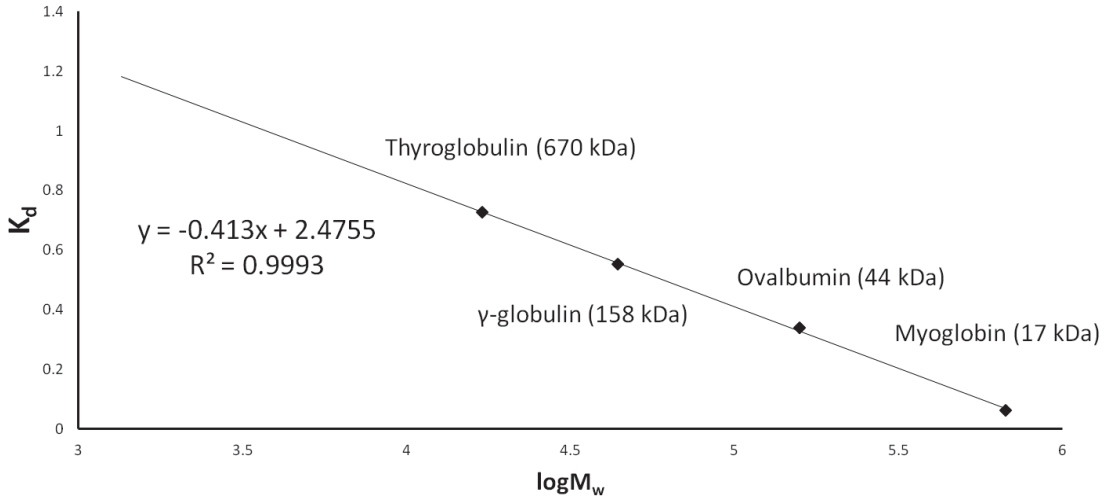


Figure S5.1: Molecular weight standard curve
Bio-Rad molecular weight standard (151-1901) was used to generate a calibration curve for gel filtration. Standards were run as described in 5.4.3.4. The distribution constants (K_d) for the standards and samples was calculated as follows: $K_d = (V_e - V_0)/V_s$ where the V_e is the eluted volume of sample, V_0 is the void volume and V_s is the stationary phase volume. The equation of the line was used to assign molecular weight.

6.0 Conclusions and future work

In this study, I set out to explore how localization, activity, structure and stability all are intertwined in defining a protein's function. I used tRNA nucleotidyltransferase which catalyzes an interesting reaction (requiring three different substrates and incorporating nucleotides without a nucleic acid template) and is targeted to multiple intracellular destinations (cytosol, nucleus, mitochondria and plastids) as my model protein. Using the Arabidopsis, yeast and human tRNA nucleotidyltransferases, I was able to show some common features that help to establish the interplay of all of these factors in the function of this enzyme.

In Chapters 2 and 3, we showed that stability and/or activity may play a role in defining the distribution of this enzyme in the cell. Modifications in the mature domain of the protein in the body (Mat-Mu1) and tail domains (Mat-Mu2) both altered intracellular localization while Mat-Mu1 also eliminated activity and altered stability. We predict that the Mat-Mu1 variant causes the protein to fold more tightly making mitochondrial import less energetically favorable. To explore more precisely the role of Mat-Mu1 we fine-structure mapped the functions of specific amino acids in this region. Three amino acid substitutions in Mat-Mu1A (the first half of Mat-Mu1) appear to alter the folding of the protein and suggest that these amino acid substitutions result in the more tightly folded protein which is favored for plastid rather than mitochondrial import. In contrast, amino acid substitutions in Mat-Mu1B (the second half of Mat-Mu1) reduced enzyme activity but had little effect on structure or stability. In fact, we were able to define a single amino acid substitution, K418 that likely plays a direct role in tRNA binding. We now are in a position to see if it is the change in folding (Mat-Mu1A) or in activity (Mat-Mu1B) that is linked to the altered distribution of this protein. Import experiments with all of these variant proteins are underway in the laboratory of our collaborators and these results will be important to distinguish between these two possibilities. In terms of defining the specific function of this region of the protein in enzyme activity it would be interesting to explore more closely the role of the lysine

residue at position 420. Converting this lysine residue to glutamate resulted in a 50% increase in turnover number. Studies such as described in Chapters 2 and 3 of this thesis could be used to define the contribution to enzyme activity of this amino acid. For example, what effect would a K420A or K418EK420E double variant have on K_M and k_{cat}.

Although changes in motif Mat-Mu2 altered intracellular distribution of the protein, dramatically reducing the number of cells showing mitochondrial localization, it showed no apparent changes in enzyme activity or physicochemical properties. Changing the charge of Mat-Mu2 or removing it completely resulted in an approximately 9-fold change in dissociation constant although the catalytic constants for these and the native enzyme were unchanged. Given the increased importance placed on the carboxy-terminal portion of tRNA nucleotidyltransferases as seen here and with the human enzyme (K416E substitution, for example), a more detailed understanding of what is happening in this region of the protein is required. We can carry out similar fine structure mapping experiments with this carboxy-terminal motif as we have done with Mat-Mu1. Given that tRNA nucleotidyltransferase is a sorting isozyme in most eukaryotes, it will be interesting to extrapolate from what we have seen in Arabidopsis to other organisms which contain both mitochondria and plastids to see if a similar mechanism of discrimination between organellar targeting is conserved.

In Chapter 4, we demonstrate that a single amino acid substitution (E189F) in motif C generates a temperature-sensitive phenotype not because of a loss of stability but due to a decrease in enzyme activity. This hypomorphic effect was suppressed by overexpression of the variant enzyme or by an intragenic mutation changing arginine at position 64 to tryptophan (which increased the activity of the E189F variant 10 fold but did not restore its stability). We show that by increasing or decreasing the amount of variant tRNA nucleotidyltransferases in yeast, we can modulate the hypomorphic effect and where the temperature-sensitive phenotype manifests itself (Fig. 4.6). It would be instructive to explore

the proteomes in these cells to see how yeast adapts to growth under these conditions of reduced tRNA nucleotidyltransferase activity given the large number of functions attributed to this enzyme.

In addition, we have begun to define more precisely the roles of specific amino acids in enzyme structure, stability and activity. Based on these experiments, we were the first to suggest that motif C played a role in enzyme activity (other than acting as a linker between the head and neck region of the protein) and that residue E189 played a function in defining this role. Our hypothesis has been supported by recent work suggesting that motif C plays a role in the orientation of the enzyme's head and body domains to accommodate the growing 3'-end of the tRNA (Ernst *et al.*, 2015). Our suppressor data further suggest that there is an interaction between motif A (R64) and motif C in enzyme activity as the R64W change suppresses the temperature-sensitive phenotype resulting from the E189F or E189K change. A more detailed biochemical and physicochemical analysis should be performed to explore in more depth the roles of amino acids E189 and R64. For example, it would be useful to have crystal structures of these yeast variant enzymes. Failing generation of crystals, it would be informative to carry out FTIR spectroscopy on the native and variant enzymes. As the available crystal structures indicate that tRNA nucleotidyltransferase is primarily α-helical with a single β-sheet in the head domain (Fig. 4.1) and R64 is found at the end of one of the β-strands of this sheet and E189 is found at a loop connecting another β-strand with an α-helix, FTIR analysis (which can more specifically separate the β-sheet signal from that of an α-helix) would allow us to more specifically identify if changing R64 or E189 alters the orientation or conformation of this β-sheet.

It is interesting that the R64W change suppresses both the E189F and E189K substitutions, but that the R64E substitution (which is viable at the restrictive temperature) will not suppress an E189F, E189K or E189R substitution. I suggest combining R64W with other substitutions which impair catalytic function to assess if the enhanced activity observed is specific to changes to motif C. In this light, it would be interesting to define the kinetic parameters of all of these variants (for example, to determine

the K_M for ATP, CTP or tRNA or the turnover number). This may allow us to define the role of R64 in motif A where the catalytic magnesium ions are bound. It may be useful that an arginine at this position is conserved in the yeast and human enzymes, but not in the bacterial enzymes for which crystal structures have been solved (Fig. 4.1). We can compare the effects of the changes in motifs A and C in the yeast enzyme with the corresponding changes in the human enzyme (and *vice versa*).

In a traditional sense, there has always been a direct connection between protein stability and activity. Here we show that is not that simple. For example, in Chapter 2 we show how changes in structure lead to altered stability but not to changes in activity. The fine structure mapping of Mat-Mu1 in Chapter 3 further revealed how altered stability could be separated from altered activity. In Chapter 4 where a yeast temperature-sensitive phenotype was explored, our data suggest that the phenotype is the result of a hypomorphic effecting loss of activity but not loss of stability. A suppressor mutation was found that restored activity but not stability. All of these observations point out how complicated it is to understand the properties that define a protein's function. Given that human diseases (Chakraborty *et al.*, 2014, Sasarman *et al.*, 2015, DeLuca *et al.*, 2016) have now been linked to mutations in the TRNT1 gene coding for tRNA nucleotidyltransferase it has become more important to develop a detailed understanding of how this protein functions.

Our preliminary characterization of seven variant proteins first identified in patients suffering from SIFD (Chakraborty *et al.*, 2014) in Chapter 5 revealed that they segregated into two groups: those with reduced thermostability and those with reduced activity. As we showed previously that reduced activity can lead to a temperature-sensitive phenotype in yeast, one can imagine that reduced activity could lead to a disease phenotype in humans. One also could imagine that reduced stability could either reduce protein function or eliminate protein molecules due to improper folding such that activity is reduced. Simply put in each of these cases, there is not sufficient tRNA nucleotidyltransferase activity to meet the needs of the cell and the disease phenotype arises. This is particularly evident in cells that

depend heavily on mitochondrial respiration and could also reflect improper localization of the enzyme which also may depend on structure or stability as we already have seen in Arabidopsis. To address these possibilities, we could initially correlate thermostability *in vitro* with protein levels *in vivo*. For example, immunoblotting as has been done with the A148V (Sasarman *et al.*, 2015), T154I, L166S, R190I and K416E variants (Chakraborty *et al.*, 2014) in human cell lines should be performed to check for tRNA nucleotidyltransferase levels in the remaining variants. At the same time intracellular localization could be explored.

In the context of structure, we showed that the human enzyme can adopt a dimeric quaternary structure (Fig. 5.6). I would like to further explore how this quaternary structure affects stability, activity and localization. I showed that the I326T variant which does not dimerize, has both reduced activity (turnover number ~nine fold less than native) and a reduced melting temperature (Fig. 5.3). I propose eliminating the disulfide bond from the native enzyme to see if activity and stability are similarly reduced. I also will explore whether changes in quaternary structure or tRNA binding have any influence on the cellular distribution of tRNA nucleotidyltransferase in humans.

While a better understanding how structure, stability, activity and localization define a protein's function will help us to better understand and treat human disease, there is still a big leap from understanding the changes in the biochemical and physicochemical properties of tRNA nucleotidyltransferase to the changes that lead to a disease phenotype. Our detailed understanding of the yeast and Arabidopsis tRNA nucleotidyltransferase have provided a solid foundation to explore the human enzyme but much work remains to be done to explore what is happening in human cells to develop the disease.

7.0 References

ABBOTT, J. A., FRANCKLYN, C. S. & ROBEY-BOND, S. M. 2014. Transfer RNA and human disease. *Front Genet,* 5**,** 158.

ABE, T., INOKUCHI, H., YAMADA, Y., MUTO, A., IWASAKI, Y. & IKEMURA, T. 2014. tRNADB-CE: tRNA gene database well-timed in the era of big sequence data. *Front Genet,* 5**,** 114.

AEBI, M., KIRCHNER, G., CHEN, J. Y., VIJAYRAGHAVAN, U., JACOBSON, A., MARTIN, N. C. & ABELSON, J. 1990. Isolation of a temperature-sensitive mutant with an altered tRNA nucleotidyltransferase and cloning of the gene encoding tRNA nucleotidyltransferase in the yeast Saccharomyces cerevisiae. *Journal of Biological Chemistry,* 265**,** 16216-16220.

AGIUS, L. 2015. Role of glycogen phosphorylase in liver glycogen metabolism. *Mol Aspects Med,* 46**,** 34-45.

ALTSCHUL, S. F., MADDEN, T. L., SCHAFFER, A. A., ZHANG, J., ZHANG, Z., MILLER, W. & LIPMAN, D. J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res,* 25**,** 3389-402.

ARAVIND, L. & KOONIN, E. V. 1999. DNA polymerase beta-like nucleotidyltransferase superfamily: identification of three new families, classification and evolutionary history. *Nucleic Acids Research,* 27**,** 1609-1618.

AUGUSTIN, M. A., REICHERT, A. S., BETAT, H., HUBER, R., MÖRL, M. & STEEGBORN, C. 2003. Crystal Structure of the Human CCA-adding Enzyme: Insights into Template-independent Polymerization. *Journal of Molecular Biology,* 328**,** 985-994.

AUSUBEL, F. M. 1989. *Short protocols in molecular biology: A compendium of methods from Current protocols in molecular biology*, Greene Pub. Associates and Wiley-Interscience.

BACHMAIR, A., FINLEY, D. & VARSHAVSKY, A. 1986. *In vivo* half-life of a protein is a function of its amino-terminal residue. *Science,* 234**,** 179-86.

BAUER, N. C., DOETSCH, P. W. & CORBETT, A. H. 2015. Mechanisms Regulating Protein Localization. *Traffic,* 16**,** 1039-61.

BEN WEBB, M. C. *MODELLER v.9.13* [Online]. Available: https://salilab.org/modeller/.

BETAT, H., RAMMELT, C., MARTIN, G. & MÖRL, M. 2004. Exchange of Regions between Bacterial Poly(A) Polymerase and the CCA-Adding Enzyme Generates Altered Specificities. *Molecular Cell,* 15**,** 389-398.

BETAT, H., RAMMELT, C. & MÖRL, M. 2010. tRNA nucleotidyltransferases: ancient catalysts with an unusual mechanism of polymerization. *Cellular and Molecular Life Sciences,* 67**,** 1447-1463.

BHUSHAN, S., LEFEBVRE, B., STAHL, A., WRIGHT, S. J., BRUCE, B. D., BOUTRY, M. & GLASER, E. 2003. Dual targeting and function of a protease in mitochondria and chloroplasts. *EMBO Rep,* 4**,** 1073-8.

BIONDA, T., TILLMANN, B., SIMM, S., BEILSTEIN, K., RUPRECHT, M. & SCHLEIFF, E. 2010. Chloroplast import signals: the length requirement for translocation in vitro and in vivo. *J Mol Biol,* 402**,** 510-23.

BLOOM, J. & CROSS, F. R. 2007. Multiple levels of cyclin specificity in cell-cycle control. *Nat Rev Mol Cell Biol,* 8**,** 149-60.

BOEKE, J. D., LACROUTE, F. & FINK, G. R. 1984. A positive selection for mutants lacking orotidine-5'-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. *Mol Gen Genet,* 197**,** 345-6.

BOLTE, K., BULLMANN, L., HEMPEL, F., BOZARTH, A., ZAUNER, S. & MAIER, U. G. 2009. Protein Targeting into Secondary Plastids. *Journal of Eukaryotic Microbiology,* 56**,** 9-15.

BOWLER, C., ALLIOTTE, T., VAN DEN BULCKE, M., BAUW, G., VANDEKERCKHOVE, J., VAN MONTAGU, M. & INZÉ, D. 1989. A plant manganese superoxide dismutase is efficiently imported and correctly

processed by yeast mitochondria. *Proceedings of the National Academy of Sciences,* 86**,** 3237-3241.

BRACHMANN, C. B., DAVIES, A., COST, G. J., CAPUTO, E., LI, J., HIETER, P. & BOEKE, J. D. 1998. Designer deletion strains derived from Saccharomyces cerevisiae S288C: A useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast,* 14**,** 115-132.

BRADFORD, M. M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem,* 72**,** 248-54.

CAPLAN, A. J. & DOUGLAS, M. G. 1991. Characterization of YDJ1: a yeast homologue of the bacterial dnaJ protein. *J Cell Biol,* 114**,** 609-21.

CARRIE, C., GIRAUD, E. & WHELAN, J. 2009. Protein transport in organelles: Dual targeting of proteins to mitochondria and chloroplasts. *Febs j,* 276**,** 1187-95.

CASTELLS-ROCA, L., GARCÍA-MARTÍNEZ, J., MORENO, J., HERRERO, E., BELLÍ, G. & PÉREZ-ORTÍN, J. E. 2011. Heat Shock Response in Yeast Involves Changes in Both Transcription Rates and mRNA Stabilities. *PLoS ONE,* 6**,** e17272.

CHAKRABORTY, P. K., SCHMITZ-ABE, K., KENNEDY, E. K., MAMADY, H., NAAS, T., DURIE, D., CAMPAGNA, D. R., LAU, A., SENDAMARAI, A. K., WISEMAN, D. H., MAY, A., JOLLES, S., CONNOR, P., POWELL, C., HEENEY, M. M., GIARDINA, P. J., KLAASSEN, R. J., KANNENGIESSER, C., THURET, I., THOMPSON, A. A., MARQUES, L., HUGHES, S., BONNEY, D. K., BOTTOMLEY, S. S., WYNN, R. F., LAXER, R. M., MINNITI, C. P., MOPPETT, J., BORDON, V., GERAGHTY, M., JOYCE, P. B., MARKIANOS, K., RUDNER, A. D., HOLCIK, M. & FLEMING, M. D. 2014. Mutations in TRNT1 cause congenital sideroblastic anemia with immunodeficiency, fevers, and developmental delay (SIFD). *Blood,* 124**,** 2867-71.

CHAKSHUSMATHI, G., MONDAL, K., LAKSHMI, G. S., SINGH, G., ROY, A., CH., R. B., MADHUSUDHANAN, S. & VARADARAJAN, R. 2004. Design of temperature-sensitive mutants solely from amino acid sequence. *Proceedings of the National Academy of Sciences of the United States of America,* 101**,** 7925-7930.

CHAUMONT, F., O'RIORDAN, V. & BOUTRY, M. 1990. Protein transport into mitochondria is conserved between plant and yeast species. *Journal of Biological Chemistry,* 265**,** 16856-16862.

CHEN, J. Y., JOYCE, P. B., WOLFE, C. L., STEFFEN, M. C. & MARTIN, N. C. 1992. Cytoplasmic and mitochondrial tRNA nucleotidyltransferase activities are derived from the same gene in the yeast Saccharomyces cerevisiae. *Journal of Biological Chemistry,* 267**,** 14879-83.

CHO, H. D., CHEN, Y., VARANI, G. & WEINER, A. M. 2006. A Model for C74 Addition by CCA-adding Enzymes: C74 ADDITION, LIKE C75 AND A76 ADDITION, DOES NOT INVOLVE tRNA TRANSLOCATION. *Journal of Biological Chemistry,* 281**,** 9801-9811.

CHO, H. D., OYELERE, A. K., STROBEL, S. A. & WEINER, A. M. 2003. Use of nucleotide analogs by class I and class II CCA-adding enzymes (tRNA nucleotidyltransferase): deciphering the basis for nucleotide selection. *Rna,* 9**,** 970-81.

CHO, H. D., TOMITA, K., SUZUKI, T. & WEINER, A. M. 2002. U2 small nuclear RNA is a substrate for the CCA-adding enzyme (tRNA nucleotidyltransferase). *J Biol Chem,* 277**,** 3447-55.

CHO, H. D., VERLINDE, C. L. M. J. & WEINER, A. M. 2007. Reengineering CCA-adding enzymes to function as (U,G)- or dCdCdA-adding enzymes or poly(C,A) and poly(U,G) polymerases. *Proceedings of the National Academy of Sciences of the United States of America,* 104**,** 54-59.

CHRISTIE, M., CHANG, C. W., RONA, G., SMITH, K. M., STEWART, A. G., TAKEDA, A. A., FONTES, M. R., STEWART, M., VERTESSY, B. G., FORWOOD, J. K. & KOBE, B. 2015. Structural Biology and Regulation of Protein Import into the Nucleus. *J Mol Biol*.

CHRISTOPHE, D., CHRISTOPHE-HOBERTUS, C. & PICHON, B. 2000. Nuclear targeting of proteins: how many different signals? *Cellular Signalling,* 12**,** 337-341.

COKOL, M., NAIR, R. & ROST, B. 2000. Finding nuclear localization signals. *EMBO Rep,* 1**,** 411-5.

COOK, A. G., FUKUHARA, N., JINEK, M. & CONTI, E. 2009. Structures of the tRNA export factor in the nuclear and cytosolic states. *Nature,* 461**,** 60-5.
CRICK, F. 1970. Central dogma of molecular biology. *Nature,* 227**,** 561-3.
CUDNY, H., PIETRZAK, M. & KACZKOWSKI, J. 1978. Plant tRNA nucleotidyltransferase. *Planta,* 142**,** 23-27.
CZECH, A., WENDE, S., MORL, M., PAN, T. & IGNATOVA, Z. 2013. Reversible and rapid transfer-RNA deactivation as a mechanism of translational repression in stress. *PLoS Genet,* 9**,** e1003767.
DARZYNKIEWICZ, Z., ZHAO, H., ZHANG, S., LEE, M. Y., LEE, E. Y. & ZHANG, Z. 2015. Initiation and termination of DNA replication during S phase in relation to cyclins D1, E and A, p21WAF1, Cdt1 and the p12 subunit of DNA polymerase delta revealed in individual cells by cytometry. *Oncotarget,* 6**,** 11735-50.
DELUCA, A. P., WHITMORE, S. S., BARNES, J., SHARMA, T. P., WESTFALL, T. A., SCOTT, C. A., WEED, M. C., WILEY, J. S., WILEY, L. A., JOHNSTON, R. M., SCHNIEDERS, M. J., LENTZ, S. R., TUCKER, B. A., MULLINS, R. F., SCHEETZ, T. E., STONE, E. M. & SLUSARSKI, D. C. 2016. Hypomorphic mutations in TRNT1 cause retinitis pigmentosa with erythrocytic microcytosis. *Hum Mol Genet,* 25**,** 44-56.
DENG, X. Y., HANIC-JOYCE, P. J. & JOYCE, P. B. 2000. Isolation and nucleotide sequence of a gene encoding tRNA nucleotidyltransferase from Kluyveromyces lactis. *Yeast,* 16**,** 945-52.
DESHAIES, R. J., KOCH, B. D., WERNER-WASHBURNE, M., CRAIG, E. A. & SCHEKMAN, R. 1988. A subfamily of stress proteins facilitates translocation of secretory and mitochondrial precursor polypeptides. *Nature,* 332**,** 800-5.
DEUTSCHER, M. P. 1990. Ribonucleases, tRNA nucleotidyltransferase, and the 3' processing of tRNA. *Prog Nucleic Acid Res Mol Biol,* 39**,** 209-40.
DEUTSCHER, M. P. & EVANS, J. A. 1977. Transfer RNA nucleotidyltransferase repairs all transfer RNAs randomly. *J Mol Biol,* 109**,** 593-7.
DIACK, A. B., HEAD, M. W., MCCUTCHEON, S., BOYLE, A., KNIGHT, R., IRONSIDE, J. W., MANSON, J. C. & WILL, R. G. 2014. Variant CJD. 18 years of research and surveillance. *Prion,* 8**,** 286-95.
DUCHÊNE, A.-M., GIRITCH, A., HOFFMANN, B., COGNAT, V., LANCELIN, D., PEETERS, N. M., ZAEPFEL, M., MARÉCHAL-DROUARD, L. & SMALL, I. D. 2005. Dual targeting is the rule for organellar aminoacyl-tRNA synthetases in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences of the United States of America,* 102**,** 16484-16489.
DUDEK, J., REHLING, P. & VAN DER LAAN, M. 2013. Mitochondrial protein import: common principles and physiological networks. *Biochim Biophys Acta,* 1833**,** 274-85.
EMMANOUILIDIS, L., GOPALSWAMY, M., PASSON, D. M., WILMANNS, M. & SATTLER, M. 2015. Structural biology of the import pathways of peroxisomal matrix proteins. *Biochim Biophys Acta*.
ENTELIS, N., BRANDINA, I., KAMENSKI, P., KRASHENINNIKOV, I. A., MARTIN, R. P. & TARASSOV, I. 2006. A glycolytic enzyme, enolase, is recruited as a cofactor of tRNA targeting toward mitochondria in Saccharomyces cerevisiae. *Genes Dev,* 20**,** 1609-20.
ERNST, F. G., RICKERT, C., BLUSCHKE, A., BETAT, H., STEINHOFF, H. J. & MORL, M. 2015. Domain movements during CCA-addition: a new function for motif C in the catalytic core of the human tRNA nucleotidyltransferases. *RNA Biol,* 12**,** 435-46.
EVANS, J. A. & DEUTSCHER, M. P. 1978. Kinetic analysis of rabbit liver tRNA nucleotidyltransferase. *Journal of Biological Chemistry,* 253**,** 7276-81.
FENG, W. & HOPPER, A. K. 2002. A Los1p-independent pathway for nuclear export of intronless tRNAs in Saccharomyces cerevisiae. *Proceedings of the National Academy of Sciences,* 99**,** 5412-5417.
FLEMING, M. D. 2011. Congenital sideroblastic anemias: iron and heme lost in mitochondrial translation. *Hematology Am Soc Hematol Educ Program,* 2011**,** 525-31.
GASTEIGER, J. 2006. Of molecules and humans. *J Med Chem,* 49**,** 6429-34.
GHAVIDEL, A., KISLINGER, T., POGOUTSE, O., SOPKO, R., JURISICA, I. & EMILI, A. 2007. Impaired tRNA nuclear export links DNA damage and cell-cycle checkpoint. *Cell,* 131**,** 915-26.

GILLMAN, E. C., SLUSHER, L. B., MARTIN, N. C. & HOPPER, A. K. 1991. MOD5 translation initiation sites determine N6-isopentenyladenosine modification of mitochondrial and cytoplasmic tRNA. *Molecular and Cellular Biology,* 11**,** 2382-2390.

GORING, M. E., LEIBOVITCH, M., GEA-MALLORQUI, E., KARLS, S., RICHARD, F., HANIC-JOYCE, P. J. & JOYCE, P. B. M. 2013. The ability of an arginine to tryptophan substitution in Saccharomyces cerevisiae tRNA nucleotidyltransferase to alleviate a temperature-sensitive phenotype suggests a role for motif C in active site organization. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics,* 1834**,** 2097-2106.

GREBENOK, R. J., PIERSON, E., LAMBERT, G. M., GONG, F. C., AFONSO, C. L., HALDEMANCAHILL, R., CARRINGTON, J. C. & GALBRAITH, D. W. 1997. Green-fluorescent protein fusions for efficient characterization of nuclear targeting. *Plant Journal,* 11**,** 573-586.

GREENFIELD N. & FASMAN, G. D. 1969. COMPUTED CIRCULAR DICHROISM SPECTRA FOR EVALUATION OF PROTEIN CONFORMATION. *Biochemistry,* 8**,** 4108-&.

GRUNWALD, D. & SINGER, R. H. 2012. Multiscale dynamics in nucleocytoplasmic transport. *Curr Opin Cell Biol,* 24**,** 100-6.

HACHIYA, N., ALAM, R., SAKASEGAWA, Y., SAKAGUCHI, M., MIHARA, K. & OMURA, T. 1993. A mitochondrial import factor purified from rat liver cytosol is an ATP-dependent conformational modulator for precursor proteins. *Embo j,* 12**,** 1579-86.

HANIC-JOYCE, P. J. 1985. Mapping CDC mutations in the yeast S. cerevisiae by rad52-mediated chromosome loss. *Genetics,* 110**,** 591-607.

HANIC-JOYCE, P. J. & JOYCE, P. B. M. 2002. Characterization of a gene encoding tRNA nucleotidyltransferase from Candida glabrata. *Yeast,* 19**,** 1399-1411.

HARTWELL, L. H. 1967. Macromolecule synthesis in temperature-sensitive mutants of yeast. *Journal of Bacteriology,* 93**,** 1662-1670.

HOFFMEIER, A., BETAT, H., BLUSCHKE, A., GÜNTHER, R., JUNGHANNS, S., HOFMANN, H.-J. & MÖRL, M. 2010. Unusual evolution of a catalytic core element in CCA-adding enzymes. *Nucleic Acids Research,* 38**,** 4436-4447.

HOLM, L. & SANDER, C. 1995. DNA polymerase β belongs to an ancient nucleotidyltransferase superfamily. *Trends in Biochemical Sciences,* 20**,** 345-347.

HOPPER, A. K. 2013. Transfer RNA post-transcriptional processing, turnover, and subcellular dynamics in the yeast Saccharomyces cerevisiae. *Genetics,* 194**,** 43-67.

HURTO, R. L., TONG, A. H., BOONE, C. & HOPPER, A. K. 2007. Inorganic phosphate deprivation causes tRNA nuclear accumulation via retrograde transport in Saccharomyces cerevisiae. *Genetics,* 176**,** 841-52.

IOLASCON, A. 2014. Transfer RNA and syndromic sideroblastic anemia. *Blood,* 124**,** 2763-4.

JAEDICKE, K., ROSLER, J., GANS, T. & HUGHES, J. 2011. Bellis perennis: a useful tool for protein localization studies. *Planta,* 234**,** 759-68.

JEFFREY, G. A. 1997. *An introduction to hydrogen bonding*, Oxford University Press.

JIMENO, S., RONDÓN, A. G., LUNA, R. & AGUILERA, A. 2002. The yeast THO complex and mRNA export factors link RNA metabolism with transcription and genome instability. *The EMBO Journal,* 21**,** 3526-3535.

JOYCE, C. M. & STEITZ, T. A. 1994. Function and structure relationships in DNA polymerases. *Annu Rev Biochem,* 63**,** 777-822.

JÜHLING, F., MÖRL, M., HARTMANN, R. K., SPRINZL, M., STADLER, P. F. & PÜTZ, J. 2009. tRNAdb 2009: compilation of tRNA sequences and tRNA genes. *Nucleic Acids Research,* 37**,** D159-D162.

JUNG, J. & LEE, B. 2000. Protein structure alignment using environmental profiles. *Protein Eng,* 13**,** 535-43.

JUST, A., BUTTER, F., TRENKMANN, M., HEITKAM, T., MÖRL, M. & BETAT, H. 2008. A comparative analysis of two conserved motifs in bacterial poly(A) polymerase and CCA-adding enzyme. *Nucleic Acids Research,* 36**,** 5212-5220.

KAISER, P., MONCOLLIN, V., CLARKE, D. J., WATSON, M. H., BERTOLAET, B. L., REED, S. I. & BAILLY, E. 1999. Cyclin-dependent kinase and Cks/Suc1 interact with the proteasome in yeast to control proteolysis of M-phase targets. *Genes Dev,* 13**,** 1190-202.

KATOH, K., KUMA, K., TOH, H. & MIYATA, T. 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res,* 33**,** 511-8.

KATOH, K. & TOH, H. 2008. Recent developments in the MAFFT multiple sequence alignment program. *Brief Bioinform,* 9**,** 286-98.

KATOH, K. & TOH, H. 2010. Parallelization of the MAFFT multiple sequence alignment program. *Bioinformatics,* 26**,** 1899-900.

KEADY, B. T., ATTFIELD, K. R. & HAKE, L. E. 2002. Differential processing of the Xenopus ATP(CTP):tRNA nucleotidyltransferase mRNA. *Biochem Biophys Res Commun,* 297**,** 573-80.

KIMOTO, C., MORIYAMA, T., TSUJII, A., IGARASHI, Y., OBUSE, C., MIYAMOTO, Y., OKA, M. & YONEDA, Y. 2015. Functional characterization of importin alpha8 as a classical nuclear localization signal receptor. *Biochim Biophys Acta,* 1853**,** 2676-83.

KRIECHBAUMER, V., WANG, P., HAWES, C. & ABELL, B. M. 2012. Alternative splicing of the auxin biosynthesis gene YUCCA4 determines its subcellular compartmentation. *Plant J,* 70**,** 292-302.

KUHN, C.-D., WILUSZ, JEREMY E., ZHENG, Y., BEAL, PETER A. & JOSHUA-TOR, L. 2015. On-Enzyme Refolding Permits Small RNA and tRNA Surveillance by the CCA-Adding Enzyme. *Cell,* 160**,** 644-658.

KUNZE, M. & BERGER, J. 2015. The similarity between N-terminal targeting signals for protein import into different organelles and its evolutionary relevance. *Front Physiol,* 6**,** 259.

LAKOWICZ, J. R. 2006. *Principles of Fluorescence Spectroscopy*, Springer Science+ Business Media.

LAMBERTI, G., DRUREY, C., SOLL, J. & SCHWENKERT, S. 2011. The phosphorylation state of chloroplast transit peptides regulates preprotein import. *Plant Signaling & Behavior,* 6**,** 1918-1920.

LANGE, A., MILLS, R. E., LANGE, C. J., STEWART, M., DEVINE, S. E. & CORBETT, A. H. 2007. Classical nuclear localization signals: definition, function, and interaction with importin alpha. *J Biol Chem,* 282**,** 5101-5.

LAWRENCE, C. W. 1991. Classical mutagenesis techniques. *Methods in Enzymology.* Academic Press.

LEFEBVRE-LEGENDRE, L., VAILLIER, J., BENABDELHAK, H., VELOURS, J., SLONIMSKI, P. P. & DI RAGO, J.-P. 2001. Identification of a Nuclear Gene (FMC1) Required for the Assembly/Stability of Yeast Mitochondrial F1-ATPase in Heat Stress Conditions. *Journal of Biological Chemistry,* 276**,** 6789-6796.

LEIBOVITCH, M., BUBLAK, D., HANIC-JOYCE, PAMELA J., TILLMANN, B., FLINNER, N., AMSEL, D., SCHARF, K.-D., MIRUS, O., JOYCE, PAUL B. M. & SCHLEIFF, E. 2013. The folding capacity of the mature domain of the dual-targeted plant tRNA nucleotidyltransferase influences organelle selection. *Biochemical Journal,* 453**,** 401-412.

LESLIE, M. 2005. Lost in translation: the signal hypothesis. *The Journal of Cell Biology,* 170**,** 338.

LEWIS, M. 2013. Allostery and the lac Operon. *J Mol Biol,* 425**,** 2309-16.

LI, F., XIONG, Y., WANG, J., CHO, H. D., TOMITA, K., WEINER, A. M. & STEITZ, T. A. 2002. Crystal Structures of the Bacillus stearothermophilus CCA-Adding Enzyme and Its Complexes with ATP or CTP. *Cell,* 111**,** 815-824.

LI, W. & GODZIK, A. 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics,* 22**,** 1658-9.

LIZANO, E., SCHEIBE, M., RAMMELT, C., BETAT, H. & MÖRL, M. 2008. A comparative analysis of CCA-adding enzymes from human and E. coli: Differences in CCA addition and tRNA 3′-end repair. *Biochimie,* 90**,** 762-772.

MARTIN, G. & KELLER, W. 1996. Mutational analysis of mammalian poly(A) polymerase identifies a region for primer binding and a catalytic domain, homologous to the family X polymerases, and to other nucleotidyltransferases. *EMBO Journal,* 15**,** 2593-2603.

MARTIN, N. C. & HOPPER, A. K. 1994. How single genes provide tRNA processing enzymes to mitochondria, nuclei and the cytosol. *Biochimie,* 76**,** 1161-1167.

MIAO, M., RYAN, K. J. & WENTE, S. R. 2006. The integral membrane protein Pom34p functionally links nucleoporin subcomplexes. *Genetics,* 172**,** 1441-1457.

MISHRA, S. K., TRIPP, J., WINKELHAUS, S., TSCHIERSCH, B., THERES, K., NOVER, L. & SCHARF, K. D. 2002. In the complex family of heat stress transcription factors, HsfA1 has a unique role as master regulator of thermotolerance in tomato. *Genes Dev,* 16**,** 1555-67.

MITCHELL, P., PETFALSKI, E., HOUALLA, R., PODTELEJNIKOV, A., MANN, M. & TOLLERVEY, D. 2003. Rrp47p Is an Exosome-Associated Protein Required for the 3′ Processing of Stable RNAs. *Molecular and Cellular Biology,* 23**,** 6982-6992.

MOLL, T., TEBB, G., SURANA, U., ROBITSCH, H. & NASMYTH, K. 1991. The role of phosphorylation and the CDC28 protein kinase in cell cycle-regulated nuclear import of the S. cerevisiae transcription factor SWI5. *Cell,* 66**,** 743-58.

MUMBERG, D., MÜLLER, R. & FUNK, M. 1995. Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene,* 156**,** 119-122.

MURTHI, A., SHAHEEN, H. H., HUANG, H. Y., PRESTON, M. A., LAI, T. P., PHIZICKY, E. M. & HOPPER, A. K. 2010. Regulation of tRNA bidirectional nuclear-cytoplasmic trafficking in Saccharomyces cerevisiae. *Mol Biol Cell,* 21**,** 639-49.

NAGAIKE, T., SUZUKI, T., TOMARI, Y., TAKEMOTO-HORI, C., NEGAYAMA, F., WATANABE, K. & UEDA, T. 2001. Identification and Characterization of Mammalian Mitochondrial tRNA nucleotidyltransferases. *Journal of Biological Chemistry,* 276**,** 40041-40049.

NEMODA, Z. & SAHIN-TOTH, M. 2006. Chymotrypsin C (caldecrin) stimulates autoactivation of human cationic trypsinogen. *J Biol Chem,* 281**,** 11879-86.

NEUENFELDT, A., JUST, A., BETAT, H. & MÖRL, M. 2008. Evolution of tRNA nucleotidyltransferases: A small deletion generated CC-adding enzymes. *Proceedings of the National Academy of Sciences of the United States of America,* 105**,** 7953-7958.

OH, B. K. & PACE, N. R. 1994. Interaction of the 3'-end of tRNA with ribonuclease P RNA. *Nucleic Acids Research,* 22**,** 4087-4094.

OKABE, M., TOMITA, K., ISHITANI, R., ISHII, R., TAKEUCHI, N., ARISAKA, F., NUREKI, O. & YOKOYAMA, S. 2003. Divergent evolutions of trinucleotide polymerization revealed by an archaeal CCA-adding enzyme structure. *The EMBO Journal,* 22**,** 5918-5927.

PAN, B., XIONG, Y. & STEITZ, T. A. 2010. How the CCA-adding enzyme selects adenine over cytosine at position 76 of tRNA. *Science,* 330**,** 937-40.

PASCHINI, M., TORO, T. B., LUBIN, J. W., BRAUNSTEIN-BALLEW, B., MORRIS, D. K. & LUNDBLAD, V. 2012. A naturally thermolabile activity compromises genetic analysis of telomere function in Saccharomyces cerevisiae. *Genetics,* 191**,** 79-93.

PEATTIE, D. A. 1979. Direct chemical method for sequencing RNA. *Proceedings of the National Academy of Sciences,* 76**,** 1760-1764.

PELTZ, S. W., DONAHUE, J. L. & JACOBSON, A. 1992. A mutation in the tRNA nucleotidyltransferase gene promotes stabilization of mRNAs in Saccharomyces cerevisiae. *Molecular and Cellular Biology,* 12**,** 5778-5784.

PHIZICKY, E. M. & HOPPER, A. K. 2010. tRNA biology charges to the front. *Genes Dev,* 24**,** 1832-60.

PIERCE, J. B., ESWARA, M. B. & MANGROO, D. 2010. The ins and outs of nuclear re-export of retrogradely transported tRNAs in Saccharomyces cerevisiae. *Nucleus,* 1**,** 224-30.
PORTER, B. W., YUEN, C. Y. & CHRISTOPHER, D. A. 2015. Dual protein trafficking to secretory and non-secretory cell compartments: clear or double vision? *Plant Sci,* 234**,** 174-9.
POULTNEY, C. S., BUTTERFOSS, G. L., GUTWEIN, M. R., DREW, K., GRESHAM, D., GUNSALUS, K. C., SHASHA, D. E. & BONNEAU, R. 2011. Rational Design of Temperature-Sensitive Alleles Using Computational Structure Prediction. *PLoS ONE,* 6**,** e23947.
PUNTA, M., COGGILL, P. C., EBERHARDT, R. Y., MISTRY, J., TATE, J., BOURSNELL, C., PANG, N., FORSLUND, K., CERIC, G., CLEMENTS, J., HEGER, A., HOLM, L., SONNHAMMER, E. L., EDDY, S. R., BATEMAN, A. & FINN, R. D. 2012. The Pfam protein families database. *Nucleic Acids Res,* 40**,** D290-301.
RADONS, J. 2016. The human HSP70 family of chaperones: where do we stand? *Cell Stress Chaperones*.
RAMIREZ-ALVARADO, M. 2008. Principles of protein misfolding. *Prog Mol Biol Transl Sci,* 84**,** 115-60.
REICHERT, A. S., THURLOW, D. L. & MORL, M. 2001. A eubacterial origin for the human tRNA nucleotidyltransferase? *Biol Chem,* 382**,** 1431-8.
RICHARDS, F. M. 1974. The interpretation of protein structures: Total volume, group volume distributions and packing density. *Journal of Molecular Biology,* 82**,** 1-14.
ROSSET, R. & MONIER, R. 1965. Instability of the terminal 3'-hydroxy sequence of transfer RNA in microorganisms. I. Turnover of terminal AMP in Saccharomyces cerevisiae. *Biochim Biophys Acta,* 108**,** 376-84.
ROUT, M. P., AITCHISON, J. D., SUPRAPTO, A., HJERTAAS, K., ZHAO, Y. & CHAIT, B. T. 2000. The yeast nuclear pore complex: Composition, architecture, transport mechanism. *Journal of Cell Biology,* 148**,** 635-651.
RUBIN, G. M. 1973. The Nucleotide Sequence of Saccharomyces cerevisiae 5.8 S Ribosomal Ribonucleic Acid. *Journal of Biological Chemistry,* 248**,** 3860-3875.
RUDHE, C., CLIFTON, R., WHELAN, J. & GLASER, E. 2002. N-terminal domain of the dual-targeted pea glutathione reductase signal peptide controls organellar targeting efficiency. *Journal of Molecular Biology,* 324**,** 577-585.
RUPRECHT, M., BIONDA, T., SATO, T., SOMMER, M. S., ENDO, T. & SCHLEIFF, E. 2010. On the Impact of Precursor Unfolding during Protein Import into Chloroplasts. *Molecular Plant,* 3**,** 499-508.
SAITOH, T., IGURA, M., OBITA, T., OSE, T., KOJIMA, R., MAENAKA, K., ENDO, T. & KOHDA, D. 2007. Tom20 recognizes mitochondrial presequences through dynamic equilibrium among multiple bound states. *Embo j,* 26**,** 4777-87.
SARKAR, S., AZAD, A. K. & HOPPER, A. K. 1999. Nuclear tRNA aminoacylation and its role in nuclear export of endogenous tRNAs in Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A,* 96**,** 14366-71.
SASARMAN, F., THIFFAULT, I., WERAARPACHAI, W., SALOMON, S., MAFTEI, C., GAUTHIER, J., ELLAZAM, B., WEBB, N., ANTONICKA, H., JANER, A., BRUNEL-GUITTON, C., ELPELEG, O., MITCHELL, G. & SHOUBRIDGE, E. A. 2015. The 3' addition of CCA to mitochondrial tRNASer(AGY) is specifically impaired in patients with mutations in the tRNA nucleotidyl transferase TRNT1. *Hum Mol Genet,* 24**,** 2841-7.
SASS, E., KARNIELY, S. & PINES, O. 2003. Folding of Fumarase during Mitochondrial Import Determines its Dual Targeting in Yeast. *Journal of Biological Chemistry,* 278**,** 45109-45116.
SBIA, M., PARNELL, E. J., YU, Y., OLSEN, A. E., KRETSCHMANN, K. L., VOTH, W. P. & STILLMAN, D. J. 2008. Regulation of the yeast Ace2 transcription factor during the cell cycle. *J Biol Chem,* 283**,** 11135-45.
SCHIESTL, R. H. & GIETZ, R. D. 1989. High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier. *Curr Genet,* 16**,** 339-46.

SCHMIDT VON BRAUN, S., SABETTI, A., HANIC-JOYCE, P. J., GU, J., SCHLEIFF, E. & JOYCE, P. B. M. 2007. Dual targeting of the tRNA nucleotidyltransferase in plants: not just the signal. *Journal of Experimental Botany,* 58**,** 4083-4093.

SCHRÖDINGER The PyMol Molecular Graphics System. 1.8 ed.: LLC.

SHAHEEN, H. H. & HOPPER, A. K. 2005. Retrograde movement of tRNAs from the cytoplasm to the nucleus in Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A,* 102**,** 11290-5.

SHAN, X., RUSSELL, T. A., PAUL, S. M., KUSHNER, D. B. & JOYCE, P. B. M. 2008. Characterization of a temperature-sensitive mutation that impairs the function of yeast tRNA nucleotidyltransferase. *Yeast,* 25**,** 219-233.

SHANMUGAM, K., HANIC-JOYCE, P. & JOYCE, P. M. 1996. Purification and characterization of a tRNA nucleotidyltransferase from Lupinus albus and functional complementation of a yeast mutation by the corresponding cDNA. *Plant Molecular Biology,* 30**,** 281-295.

SHERMAN, F. 2002. Getting started with yeast. *In:* CHRISTINE, G. & GERALD, R. F. (eds.) *Methods in Enzymology.* Academic Press.

SHI, P. Y., MAIZELS, N. & WEINER, A. M. 1998. CCA addition by tRNA nucleotidyltransferase: polymerization without translocation? *Embo j,* 17**,** 3197-206.

SIKORSKI, R. S. & HIETER, P. 1989. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. *Genetics,* 122**,** 19-27.

SMALDINO, P. J., READ, D. F., PRATT-HYATT, M., HOPPER, A. K. & ENGELKE, D. R. 2015. The cytoplasmic and nuclear populations of the eukaryote tRNA-isopentenyl transferase have distinct functions with implications in human cancer. *Gene,* 556**,** 13-8.

SODING, J., BIEGERT, A. & LUPAS, A. N. 2005. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res,* 33**,** W244-8.

STEINER-MOSONYI, M. & MANGROO, D. 2004. The nuclear tRNA aminoacylation-dependent pathway may be the principal route used to export tRNA from the nucleus in Saccharomyces cerevisiae. *Biochem J,* 378**,** 809-16.

STEITZ, T. A. 1998. Structural biology: A mechanism for all polymerases. *Nature,* 391**,** 231-232.

TACHIBANA, C., BIDDICK, R., LAW, G. L. & YOUNG, E. T. 2007. A Poised Initiation Complex Is Activated by SNF1. *Journal of Biological Chemistry,* 282**,** 37308-37315.

TAKEUCHI, J. & TOH-E, A. 2001. Genetic dissection of the yeast 26S proteasome: Cell cycledefects caused by the Δrpn9 mutation. *Biochimie,* 83**,** 333-340.

TAN, G., CHEN, M., FOOTE, C. & TAN, C. 2009. Temperature-Sensitive Mutations Made Easy: Generating Conditional Mutations by Using Temperature-Sensitive Inteins That Function Within Different Temperature Ranges. *Genetics,* 183**,** 13-22.

THOM, C. S., DICKSON, C. F., GELL, D. A. & WEISS, M. J. 2013. Hemoglobin variants: biochemical properties and clinical correlates. *Cold Spring Harb Perspect Med,* 3**,** a011858.

THOMPSON, D. M. & PARKER, R. 2009. Stressing Out over tRNA Cleavage. *Cell,* 138**,** 215-219.

TOH, Y., TAKESHITA, D., NUMATA, T., FUKAI, S., NUREKI, O. & TOMITA, K. 2009. Mechanism for the definition of elongation and termination by the class II CCA-adding enzyme. *The EMBO Journal,* 28**,** 3353-3365.

TOLERICO, L. H., BENKO, A. L., ARIS, J. P., STANFORD, D. R., MARTIN, N. C. & HOPPER, A. K. 1999. Saccharomyces cerevisiae Mod5p-II contains sequences antagonistic for nuclear and cytosolic locations. *Genetics,* 151**,** 57-75.

TOMITA, K., FUKAI, S., ISHITANI, R., UEDA, T., TAKEUCHI, N., VASSYLYEV, D. G. & NUREKI, O. 2004. Structural basis for template-independent RNA polymerization. *Nature,* 430**,** 700-704.

TOMITA, K., ISHITANI, R., FUKAI, S. & NUREKI, O. 2006. Complete crystallographic analysis of the dynamics of CCA sequence addition. *Nature,* 443**,** 956-960.

TOMITA, K. & WEINER, A. M. 2001. Collaboration Between CC- and A-Adding Enzymes to Build and Repair the 3'-Terminal CCA of tRNA in Aquifex aeolicus. *Science,* 294**,** 1334-1336.
TONIN, Y. & ENTELIS, N. 2014. [Mitochondrial DNA diseases and therapeutic strategies]. *Med Sci (Paris),* 30**,** 1101-9.
TORRES, E. M., SOKOLSKY, T., TUCKER, C. M., CHAN, L. Y., BOSELLI, M., DUNHAM, M. J. & AMON, A. 2007. Effects of Aneuploidy on Cellular Physiology and Cell Division in Haploid Yeast. *Science,* 317**,** 916-924.
TORRES, E. M., WILLIAMS, B. R. & AMON, A. 2008. Aneuploidy: cells losing their balance. *Genetics,* 179**,** 737-46.
TORRES, E. M., WILLIAMS, B. R., TANG, Y.-C. & AMON, A. 2010. Thoughts on Aneuploidy. *Cold Spring Harbor symposia on quantitative biology,* 75**,** 445-451.
TRETBAR, S., NEUENFELDT, A., BETAT, H. & MÖRL, M. 2011. An inhibitory C-terminal region dictates the specificity of A-adding enzymes. *Proceedings of the National Academy of Sciences,* 108**,** 21040-21045.
TRIVEDI, M. V., LAURENCE, J. S. & SIAHAAN, T. J. 2009. The role of thiols and disulfides in protein chemical and physical stability. *Current protein & peptide science,* 10**,** 614-625.
VESTWEBER, D. & SCHATZ, G. 1988. Mitochondria can import artificial precursor proteins containing a branched polypeptide chain or a carboxy-terminal stilbene disulfonate. *J Cell Biol,* 107**,** 2045-9.
VÖRTLER, S. & MÖRL, M. 2010. tRNA-nucleotidyltransferases: Highly unusual RNA polymerases with vital functions. *FEBS Letters,* 584**,** 297-302.
WALSH, C. T., GARNEAU-TSODIKOVA, S. & GATTO, G. J., JR. 2005. Protein posttranslational modifications: the chemistry of proteome diversifications. *Angew Chem Int Ed Engl,* 44**,** 7342-72.
WANG, J. Y., WANG, X. Y., SUN, H. Y., LIU, D. M., ZHANG, W., JIN, C. X., LI, L. & ZHU, L. Q. 2014. [Correlation between pulmonary function impairment and levels of alpha1-antitrypsin in serum and colon of ulcerative colitis patients: a clinical research]. *Zhongguo Zhong Xi Yi Jie He Za Zhi,* 34**,** 20-6.
WHISSTOCK, J. C. & LESK, A. M. 2003. Prediction of protein function from protein sequence and structure. *Q Rev Biophys,* 36**,** 307-40.
WHITNEY, M. L., HURTO, R. L., SHAHEEN, H. H. & HOPPER, A. K. 2007. Rapid and reversible nuclear accumulation of cytoplasmic tRNA in response to nutrient availability. *Mol Biol Cell,* 18**,** 2678-86.
WILUSZ, J. E., WHIPPLE, J. M., PHIZICKY, E. M. & SHARP, P. A. 2011. tRNAs marked with CCACCA are targeted for degradation. *Science,* 334**,** 817-21.
WISEMAN, D. H., MAY, A., JOLLES, S., CONNOR, P., POWELL, C., HEENEY, M. M., GIARDINA, P. J., KLAASSEN, R. J., CHAKRABORTY, P., GERAGHTY, M. T., MAJOR-COOK, N., KANNENGIESSER, C., THURET, I., THOMPSON, A. A., MARQUES, L., HUGHES, S., BONNEY, D. K., BOTTOMLEY, S. S., FLEMING, M. D. & WYNN, R. F. 2013. A novel syndrome of congenital sideroblastic anemia, B-cell immunodeficiency, periodic fevers, and developmental delay (SIFD). *Blood,* 122**,** 112-23.
WOLFE, C. L., HOPPER, A. K. & MARTIN, N. C. 1996. Mechanisms leading to and the consequences of altering the normal distribution of ATP(CTP):tRNA nucleotidyltransferase in yeast. *J Biol Chem,* 271**,** 4679-86.
WOLFE, C. L., LOU, Y. C., HOPPER, A. K. & MARTIN, N. C. 1994. Interplay of heterogeneous transcriptional start sites and translational selection of AUGs dictate the production of mitochondrial and cytosolic/nuclear tRNA nucleotidyltransferase from the same gene in yeast. *J Biol Chem,* 269**,** 13361-6.
XIONG, Y., LI, F., WANG, J., WEINER, A. M. & STEITZ, T. A. 2003. Crystal Structures of an Archaeal Class I CCA-Adding Enzyme and Its Nucleotide Complexes. *Molecular Cell,* 12**,** 1165-1172.
XIONG, Y. & STEITZ, T. A. 2004. Mechanism of transfer RNA maturation by CCA-adding enzyme without using an oligonucleotide template. *Nature,* 430**,** 640-5.

XIONG, Y. & STEITZ, T. A. 2006. A story with a good ending: tRNA 3'-end maturation by CCA-adding enzymes. *Current Opinion in Structural Biology,* 16**,** 12-17.

YAKUNIN, A. F., PROUDFOOT, M., KUZNETSOVA, E., SAVCHENKO, A., BROWN, G., ARROWSMITH, C. H. & EDWARDS, A. M. 2004. The HD domain of the Escherichia coli tRNA nucleotidyltransferase has 2',3'-cyclic phosphodiesterase, 2'-nucleotidase, and phosphatase activities. *J Biol Chem,* 279**,** 36819-27.

YAMASHITA, S., TAKESHITA, D. & TOMITA, K. 2014. Translocation and rotation of tRNA during template-independent RNA polymerization by tRNA nucleotidyltransferase. *Structure,* 22**,** 315-25.

YANG, N. 2008. *ATP:CTP nucleotidyltransferase: interaction with tRNA and functional roles of conserved arginine residues, the C-terminus, the tRNA T-loop, and metal ions.* Ph.D. , Clark University.

YASHIRODA, H., OGUCHI, T., YASUDA, Y., TOH-E, A. & KIKUCHI, Y. 1996. Bul1, a new protein that binds to the Rsp5 ubiquitin ligase in Saccharomyces cerevisiae. *Molecular and Cellular Biology,* 16**,** 3255-63.

YOGEV, O., NAAMATI, A. & PINES, O. 2011. Fumarase: a paradigm of dual targeting and dual localized functions. *Febs j,* 278**,** 4230-42.

YUE, D., MAIZELS, N. & WEINER, A. M. 1996. CCA-adding enzymes and poly(A) polymerases are all members of the same nucleotidyltransferase superfamily: characterization of the CCA-adding enzyme from the archaeal hyperthermophile Sulfolobus shibatae. *RNA,* 2**,** 895-908.

ZASLOFF, M. 1983. tRNA transport from the nucleus in a eukaryotic cell: carrier-mediated translocation process. *Proceedings of the National Academy of Sciences,* 80**,** 6436-6440.

ZASLOFF, M., ROSENBERG, M. & SANTOS, T. 1982. Impaired nuclear transport of a human variant tRNAiMet. *Nature,* 300**,** 81-4.

ZEHORAI, E., YAO, Z., PLOTNIKOV, A. & SEGER, R. 2010. The subcellular localization of MEK and ERK--a novel nuclear translocation signal (NTS) paves a way to the nucleus. *Mol Cell Endocrinol,* 314**,** 213-20.

ZHU, L. & DEUTSCHER, M. P. 1987. tRNA nucleotidyltransferase is not essential for Escherichia coli viability. *Embo j,* 6**,** 2473-7.

ZHU, L. Q., CUDNY, H. & DEUTSCHER, M. P. 1986. A mutation in Escherichia coli tRNA nucleotidyltransferase that affects only AMP incorporation is in a sequence often associated with nucleotide-binding proteins. *J Biol Chem,* 261**,** 14875-7.

ZIMMER, S. L., SCHEIN, A., ZIPOR, G., STERN, D. B. & SCHUSTER, G. 2009. Polyadenylation in Arabidopsis and Chlamydomonas organelles: the input of nucleotidyltransferases, poly(A) polymerases and polynucleotide phosphorylase. *Plant J,* 59**,** 88-99.

ZOLADEK, T., VADUVA, G., HUNTER, L. A., BOGUTA, M., GO, B. D., MARTIN, N. C. & HOPPER, A. K. 1995. Mutations altering the mitochondrial-cytoplasmic distribution of Mod5p implicate the actin cytoskeleton and mRNA 3' ends and/or protein synthesis in mitochondrial delivery. *Mol Cell Biol,* 15**,** 6884-94.

Appendix: Published manuscripts

Chapter 2

Biochem. J. (2013) 453 (401–412) (Printed in Great Britain) doi:10.1042/BJ20121577

The folding capacity of the mature domain of the dual-targeted plant tRNA nucleotidyltransferase influences organelle selection

Matthew Leibovitch*[1], Daniela Bublak†[1], Pamela J. Hanic-Joyce*, Bodo Tillmann†, Nadine Flinner†, Daniel Amsel†, Klaus-Dieter Scharf†, Oliver Mirus†, Paul B. M. Joyce*[2] and Enrico Schleiff†‡§[2]

*Department of Chemistry and Biochemistry and Centre for Structural and Functional Genomics, Concordia University, 7141 Sherbrooke St. W., Montréal, Québec, Canada H4B 1R6, †Department of Biosciences, Goethe University, Max von Laue Str. 9, 60438 Frankfurt/Main, Germany, ‡Cluster of Excellence Frankfurt, Goethe University, Max von Laue Str. 9, 60438 Frankfurt/Main, Germany, and §Center of Membrane Proteomics, Goethe University, Max von Laue Str. 9, 60438 Frankfurt/Main, Germany

tRNA-NTs (tRNA nucleotidyltransferases) are required for the maturation or repair of tRNAs by ensuring that they have an intact cytidine-cytidine-adenosine sequence at their 3′-termini. Therefore this enzymatic activity is found in all cellular compartments, namely the nucleus, cytoplasm, plastids and mitochondria, in which tRNA synthesis or translation occurs. A single gene codes for tRNA-NT in plants, suggesting a complex targeting mechanism. Consistent with this, distinct signals have been proposed for plastidic, mitochondrial and nuclear targeting. Our previous research has shown that in addition to N-terminal targeting information, the mature domain of the protein itself modifies targeting to mitochondria and plastids. This suggests the existence of an as yet unknown determinate for the distribution of dual-targeted proteins between these two organelles. In the present study, we explore the enzymatic and physicochemical properties of tRNA-NT variants to correlate the properties of the enzyme with the intracellular distribution of the protein. We show that alteration of tRNA-NT stability influences its intracellular distribution due to variations in organelle import capacities. Hence the fate of the protein is determined not only by the transit peptide sequence, but also by the physicochemical properties of the mature protein.

Key words: *Arabidopsis thaliana*, *in vivo* protein translocation, protein stability, protein targeting, tRNA nucleotidyltransferase.

Abbreviations used: CCA, cytidine-cytidine-adenosine; Mat-Mu, mutant of the mature domain of tRNA-NT; tRNA-NT, tRNA nucleotidyltransferase.

Chapter 4

Contents lists available at SciVerse ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbapap

The ability of an arginine to tryptophan substitution in *Saccharomyces cerevisiae* tRNA nucleotidyltransferase to alleviate a temperature-sensitive phenotype suggests a role for motif C in active site organization

Mark E. Goring [a], Matthew Leibovitch [b], Ester Gea-Mallorqui [b], Shawn Karls [a], Francis Richard [b], Pamela J. Hanic-Joyce [b], Paul B.M. Joyce [a,b,c,*]

[a] Department of Biology, Concordia University, Montréal, H4B 1R6, Canada
[b] Department of Chemistry and Biochemistry, Concordia University, Montréal, Québec H4B 1R6, Canada
[c] Centre for Structural and Functional Genomics, Concordia University, Montréal, Québec H4B 1R6, Canada

ARTICLE INFO

Article history:
Received 18 January 2013
Received in revised form 8 July 2013
Accepted 10 July 2013
Available online xxxx

Keywords:
Yeast
tRNA nucleotidyltransferase
Temperature-sensitive
Motif C
Hypomorphic
Suppressor

ABSTRACT

We report that the temperature-sensitive (ts) phenotype in *Saccharomyces cerevisiae* associated with a variant tRNA nucleotidyltransferase containing an amino acid substitution at position 189 results from a reduced ability to incorporate AMP and CMP into tRNAs. We show that this defect can be compensated for by a second-site suppressor converting residue arginine 64 to tryptophan. The R64W substitution does not alter the structure or thermal stability of the enzyme dramatically but restores catalytic activity *in vitro* and suppresses the ts phenotype *in vivo*. R64 is found in motif A known to be involved in catalysis and nucleotide triphosphate binding while E189 lies within motif C previously thought only to connect the head and neck domains of the protein. Although mutagenesis experiments indicate that residues R64 and E189 do not interact directly, our data suggest a critical role for residue E189 in enzyme structure and function. Both R64 and E189 may contribute to the organization of the catalytic domain of the enzyme. These results, along with overexpression and deletion analyses, show that the ts phenotype of *cca1-E189F* does not arise from thermal instability of the variant tRNA nucleotidyltransferase but instead from the inability of a partially active enzyme to support growth only at higher temperatures.

1. Introduction

All mature transfer RNA (tRNA) molecules contain an invariant cytidine–cytidine–adenosine (CCA) sequence at their 3′-ends that is required for amino acid attachment and protein synthesis. In many organisms this CCA sequence is not encoded by the tRNA genes so addition of CMP and AMP residues is catalyzed by the enzyme ATP(CTP):tRNA nucleotidyltransferase (EC 2.7.7.25). The CCA-adding enzyme has been the subject of considerable study because the stepwise addition of specific nucleotides occurs without a nucleic acid template [see 1,2 for recent reviews]. CCA-adding enzymes belong to the nucleotidyltransferase (NT) superfamily [3,4] and share a similar structure with four domains (head, neck, body and tail indicated in Fig. 1A) of similar dimensions [5–8] and a conserved active site signature [3]. Based on sequence similarity, the nucleotidyltransferase superfamily can be divided further into two classes [9]. Class I enzymes (including the archaeal CCA-adding enzymes, eukaryotic poly (A) polymerases, DNA polymerase β and kanamycin nucleotidyltransferase) do not show significant overall similarity to each other outside of the shared active site signature. In contrast, class II enzymes (representing eubacterial and eukaryotic CCA-adding enzymes) share a conserved N-terminal region containing several conserved motifs (color-coded in Fig. 1B and designated A to E in Fig. 1C) with, in many cases, defined roles in nucleotide specificity and catalysis [5].

Examination of solved crystal structures [6] and co-crystal structures of class II tRNA nucleotidyltransferases complexed with one [5,10] or more [11,12] substrates revealed roles for residues in these conserved motifs. Motif A (found in all nucleotidyltransferases) includes a highly conserved GGxVRD sequence (where x is any amino acid) and the two metal-binding carboxylates, DxD, involved in catalysis and the binding of incoming nucleotide triphosphates [1,5,13]. This conserved motif is found as part of a consensus five-stranded antiparallel β-sheet flanked by two α-helices common to all tRNA nucleotidyltransferases [1,5,7]. Motif B plays a critical role in ribose recognition with the conserved RRD sequence recognizing the 2′ hydroxyl group of the ribose in each incoming ribonucleotide triphosphate [5,14]. Motif A and motif B are separated by a sequence which is thought to form a highly

* Corresponding author at: Department of Chemistry and Biochemistry, Concordia University, 7141 Sherbrooke St. W., Montreal, QC H4B 1R6, Canada. Tel.: +1 514 848 2424x3351; fax: +1 514 848 2868.
E-mail address: Paul.Joyce@concordia.ca (P.B.M. Joyce).

1570-9639/$ – see front matter © 2013 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.bbapap.2013.07.003