MODIFICATIONS OF TRANSFER RNA ENHANCE SELENOPROTEIN BIOSYNTHESIS

DAMIAN LA ROSA MONTES
Bachelor of Science, Mount Royal University, 2019

A thesis submitted

in partial fulfilment of the requirement for the degree of

MASTER OF SCIENCE

in

BIOCHEMISTRY

Department of Chemistry and Biochemistry
University of Lethbridge
LETHBRIDGE, ALBERTA, CANADA

POST-TRANSCRIPTION MODIFICATIONS OF TRANSFER RNA ENHANCE SELENOPROTEIN BIOSYNTHESIS

DAMIAN LA ROSA MONTES

Date of Defense: April 20th, 2022

Dr. U. Kothe Thesis Supervisor	Professor	Ph.D.
Dr. T. Patel Thesis Co-supervisor	Associate Professor	Ph.D.
Dr. E. Schultz Thesis Examination Committee Member	Associate Professor	Ph.D.
Dr. H.J. Wieden Thesis Examination Committee Member	Professor	Ph.D.
Dr. N. Budisa External Examiner	Professor	Ph.D.

ABSTRACT

Selenocysteine-containing proteins, selenoproteins, are generally favoured over cysteine-containing proteins in redox reactions and therefore of high interest due to their multiple bio-industrial applications. However, the unique features of Selenocysteine biosynthesis and insertion machinery limits the freedom of selenoprotein bioengineering. Whereas engineered Selenocysteine machinery alternatives have been proposed, the role of tRNA post-transcriptional modifications on selenoprotein biosynthesis remains unexplored. This thesis provides insights into the overall positive role tRNA modification on stop codon re-assignment. I have shown that by supplementing coupled transcription-translation cell-free systems with tRNA modifying enzymes, the fluorescence signal of a reporter protein is enhanced compared to protein synthesis in the absence of the modifying enzymes. I conclude that tRNA modifications, in particular N6-isopenthenyladenosine (i^6A_{37}), can potentially improve selenoprotein biosynthesis *in vitro.* Thus, the result of this thesis provides valuable knowledge that could contribute to future optimization of cell-free platforms for selenoprotein bioengineering.

ACKNOWLEDGEMENTS

First and foremost, a heartfelt thank you goes to my supervisor Dr. Ute Kothe for her invaluable advice, patience, and continuous support during the course of my M.Sc. study. She brought me to the world of applied science when she recruited me for the Future Innovators in Technology pilot program between the University of Lethbridge and Mount Royal University. Her immense knowledge, plentiful experience and mentorship have made everything possible. I am eternally grateful.

I would also like to thank my co-supervisor Dr. Trushar Patel and my committee members, Dr. Hans-Joachim Wieden and Dr. Elizabeth Schultz, for providing great insights and advice for my thesis work during the committee meetings.

A special thank you goes to all my lab members for welcoming me with opened hands. In particular, I would like to express my most sincere gratitude to Sarah Schultz for being a great source of knowledge and inspiration, Hope Vienneau for being my first supervisor as a summer intern, Dr. Laura Keffer-Wilkes for coaching me during the RNA Innovation program, and Dr. Julia Guegueniat for providing my first orientation to the lab and continuous support.

A sincere thank you goes to all my family members, who have motivated me to continue thriving for success. A wholehearted thank you goes to my beloving wife for unconditionally supporting me during this time of personal growth. You are my rock.

Lastly, I acknowledge the generous financial support from the University of Lethbridge, Alberta RNA Research and Training institute, NSERC, and Alberta Innovates.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Cys	Cystine
DDT	Dithiothreitol
DMAPP	Dimethylallyl pyrophosphate
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
DsDNA	Double stranded DNA
EDTA	Ethylenediaminetetraacetic acid
EF-Tu	Elongation factor Tu
Et-Br	Ethidium Bromide
GTP	Guanosine-5´-triphosphate
IDT	Integrated DNA Technologies
IPTG	Isopropyl β-D-1-thiogalactopyranoside
K_D	Dissociation constant
LB	Lysogeny broth
MiaA	tRNA dimethylallyltransferase
mRNA	Message RNA
nt	Nucleotide
NTP	Nucleotide triphosphate
OD	Optical density
PAGE	polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PLP	Pyridoxal-5'-phosphate
PMSF	Phenyl-methyl-sulfonyl fluoride
RF1	Release factor 1
RF2	Release factor 2
RF3	Release factor 3
RNA	Ribonucleic acid
SAM	S-Adenosyl-L-Methionine
SDS	Sodium dodecyl sulfate
SEC	Size exclusion chromatography
Sec	Selenocysteine
SECIS	Selenecysteine insertion sequence
SelA	*L*-seryl-tRNA^{Sec} selenium transferase
SelB	Selenocysteine-specific elongation factor
SelD	Selenophosphate synthase
Ser	Serine
SerRS	Serine-tRNA ligase
sfGFP	Super-fold green fluorescence protein
TrmA	tRNA/tmRNA (uracil-C(5))-methyltransferase
tRNA	Transfer RNA
TruB	tRNA pseudouridine synthase B
TXTL	Coupled in-vitro transcription-translation reactions
UTP	Uridine triphosphate
WT	Wildtype

Chapter 1: INTRODUCTION

1.1. Selenium, dose-dependent but essential trace element

Selenium (Se) is an element belonging to the group VIA of the periodic table which is essential to biological systems. Naturally, it is present in four inorganic oxidation states: selenate (SeO_4^{2-}), selenite (SeO_3^{2-}), selenide (Se^{2-}) and elemental selenium (Se^0) [1]. Selenate and selenite are reported to be highly reactive, inducing the production of reactive oxygen species (ROS) that leads to a mixture of negative and positive effects, ranging from drastic reduction of biomass accumulation in plants [2] to anticancer roles in a dose-dependant matter in human cell lines [3].

Selenite and selenide are both easily oxidized and converted to a most stable form, elemental selenium [1]. The presence of organic selenium has been identified in hypermodifications (mnm^5Se^2U and $cmnm^5Se^2U$) of position 34 (wobble position) of certain transfer RNAs (tRNA) in bacteria ($tRNA^{Gly}$, $tRNA^{Gln}$, and $tRNA^{Lys}$), and is hypothesized to be carried out by tRNA 2-selenouridine synthase (SelU) in two consecutive reactions [4]. Although the exact function of these modifications is unknown, they have been proposed to be involved in codon-anticodon interactions, contributing to base pair discrimination at the wobble position and providing higher translation efficiency [5].

Two of the most abundant organic forms of selenium are selenomethionine (SeMet) and selenocysteine (Sec), which are incorporated into proteins. Currently, Se-containing proteins, selenoproteins, can be divided in three categories: specific enzymatic proteins, proteins

containing non-specifically incorporated Se and Se-binding proteins. A great example of selenoproteins having Se non-specifically incorporated are those containing the SeMet amino acids. SeMet is mostly synthesized in plants and some fungi and acquired via nutrition. Although the exact mechanism for SeMet incorporation into proteins is unclear, it is a widely accepted hypothesis that such events occur randomly as many organisms are not able to properly discriminate between SeMet and its sulfur analogue methionine (Met) [6].

On the other hand, specific enzymatic selenoproteins are mainly represented by those containing Sec amino acids, whose biosynthesis and insertional mechanism has been well studied across all domains of life. The human selenoproteome is composed of roughly 25 Sec-containing proteins, having the Sec residue mostly within the catalytic site. Currently, most human selenoproteins are broadly classified as oxidoreductase enzymes, mitigating the damage caused by ROS. Often, defects in selenoprotein malfunction have been associated with numerous diseases such prostate and gastritic cancer [7, 8]. However, it is worth highlighting that the overall effect of Se on human health is dose-dependent and specific to the Se form [9].

1.2. Special canonical amino acid: Selenocysteine and bacterial selenoproteins

Selenocysteine (Sec) is the 21st amino acid naturally encoded by the genetic code and incorporated into the nascent polypeptide chain during translation via stop codon (UGA) suppression (Figure 1.1) [10]. Unlike other proteinogenic amino acids, Sec is synthesized on its cognate tRNA ($tRNA^{Sec}_{UCA}$). First, $tRNA^{Sec}_{UCA}$ is misacylated with serine (Ser) by the

canonical seryl-tRNA synthetase (SerRS) to form Seryl-tRNA$^{Sec}_{UCA}$ (Ser-tRNA^{Sec}), which is then converted to Selenocysteyl- tRNA$^{Sec}_{UCA}$ (Sec-tRNA^{Sec}) by selenocysteine synthase SelA.

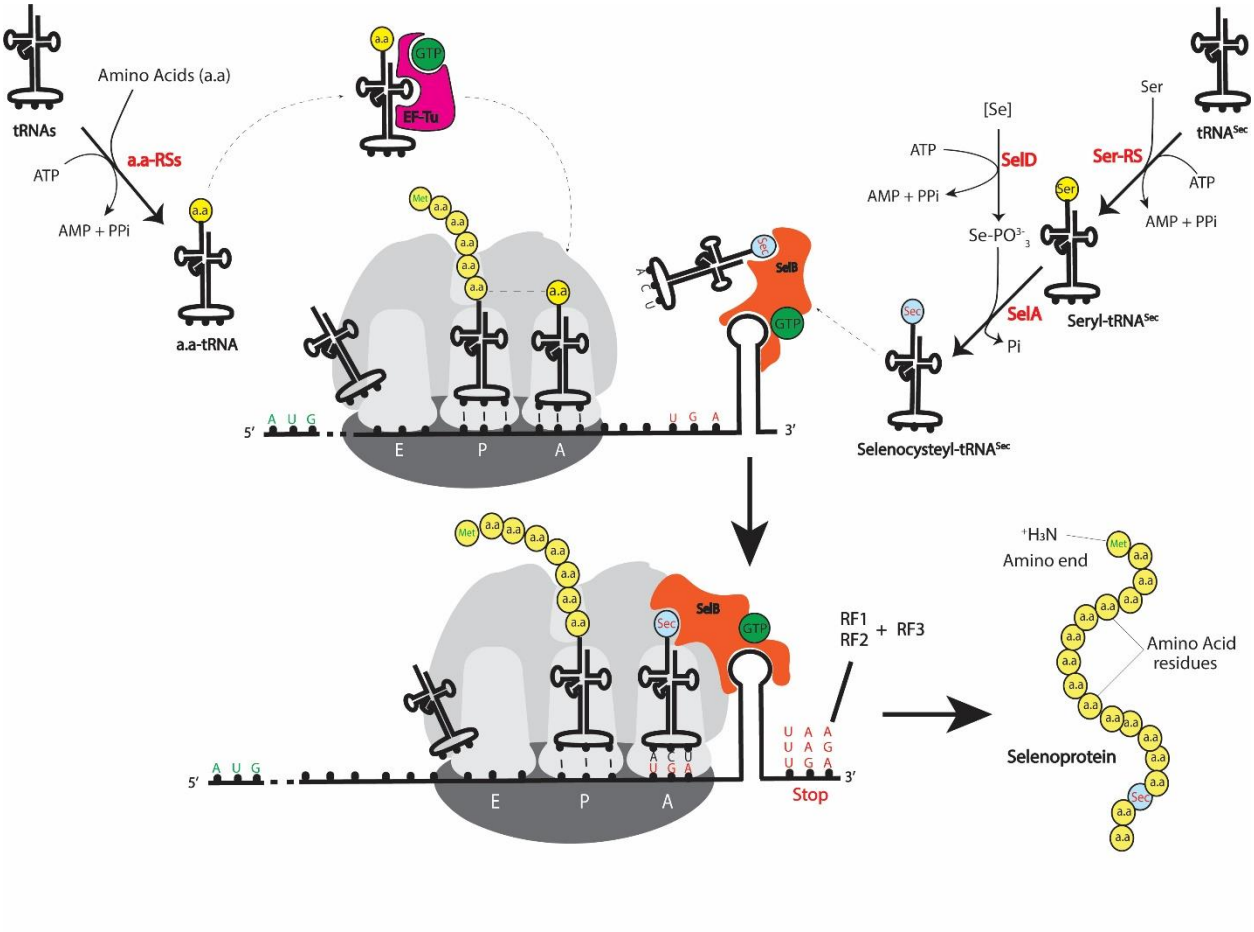


Figure 1. 1 - *E. coli* selenocysteine biosynthesis and insertion machinery. Canonical amino acids are loaded onto their corresponding tRNAs (aa-tRNAs) by their cognate amino acid tRNA synthetases (aa-RSs) and brought to the ribosomal A-site by elongation factor EF-Tu during translation. Selenocysteine (Sec) is differentially synthesized onto tRNA^{Sec} and incorporated into proteins in a multi-step complex process: tRNA^{Sec} is serylated by Seryl-tRNA synthetase (Ser-RS), converted to Sec by SelA using selenium monophosphate as the selenium donor, and brought to the site of translation by Sec-specific elongation factor, SelB. SelB binds to the Sec Insertion Sequence (SECIS) located after the UGA stop codon within the mRNA and properly orients Sec-tRNA^{Sec} into the ribosomal A-site once the ribosome is decoding the stop codon, subsequently leading to the incorporation of Sec into the nascent peptide chain.

During translation, the acylated tRNAs carrying the 20 standard canonical amino acids are brought to the ribosomal A-site by the highly abundant elongation factor EF-Tu whereas Sec-

$tRNA^{Sec}$ utilises a Sec-specific elongation factor, SelB [10]. In bacteria, the SelB•GTP•Sec-$tRNA^{Sec}$ ternary complex binds to the SElenoCysteine Insertion Sequence (SECIS), a coding mRNA stem-loop structure located just downstream of the designated UGA codon. Once the ribosome reaches the UGA codon, Sec-$tRNA^{Sec}_{UCA}$ is properly oriented to the A-site and Sec is subsequently incorporated into the peptide chain.

Although Sec-containing proteins are present in all three evolutionary lines of descents, they are not ubiquitous across all domains [11, 12]. In bacteria, the number of selenoprotein families ranges from 1 to 25, with formate dehydrogenases (*fdh*) the only family present in all Sec-containing organisms [11]. Arguably, formate dehydogenases are crucial metabolic enzymes in bacteria, specially under anerobic conditions where formate is commonly metabolized to supplement growth [13]. These enzymes are highly diverse and employ a wide range of cofactors such as molybdenum, FeS clusters, flavins, cytochromes and tungsten [14].

In *E. coli*, there are only three Sec-containing proteins. They belong to the family of formate dehydrogenases (FDH_N, FDH_H, FDH_O), which are mainly involved in redox reactions between formate and carbon dioxide. The redox efficiency of selenoproteins has been accredited to the catalytic residue, Sec. Despite Cysteine (Cys) and Selenocysteine being structurally equivalent, the substitution of the Sulfur atom for Selenium in Sec makes it more efficient at nucleophilic attack due to notable lower pKa values (thiolate: pKa ~ 8.5 vs. selenolate: pKa ~ 5.2). Moreover, Sec has a lower redox potential than Cys, making it more competent to act as a leaving group [15]. Thus, selenoproteins are generally favoured over cysteine-containing proteins in redox reactions. Nevertheless, the biochemical differences between cysteine and Sec stated above are only relevant in free form and may drastically vary depending on protein's microenvironment, making it an insufficient argument to assume selenoproteins are always preferable in nature [16].

1.2.1. Specialized aminocylation mechanism: Sec biosynthesis

The genetic code is decoded by the ribosome checking the codon-anticodon interaction in the decoding center while the aminoacyl tRNA synthetases (aaRS) interpret the code by making sure each amino acid is loaded onto their cognate tRNAs. However, Sec is one of the rare cases where the amino acid is not directly attached onto the tRNA but rather synthesized on its cognate $tRNA^{Sec}$, involving a more specialized enzymatic mechanism (Figure 1.2). First, $tRNA^{Sec}$ is misacylated with Ser by the non-cognate SerRS to form Ser-$tRNA^{Sec}$ in two consecutive reactions. The first results in the formation of L-seryl adenylate intermediate releasing inorganic pyrophosphate (PPi) in the presence of ATP as the energy source, and the second in the incorporation of L-seryl to the 3`OH end of the acceptor arm of $tRNA^{Sec}$ [17, 18].

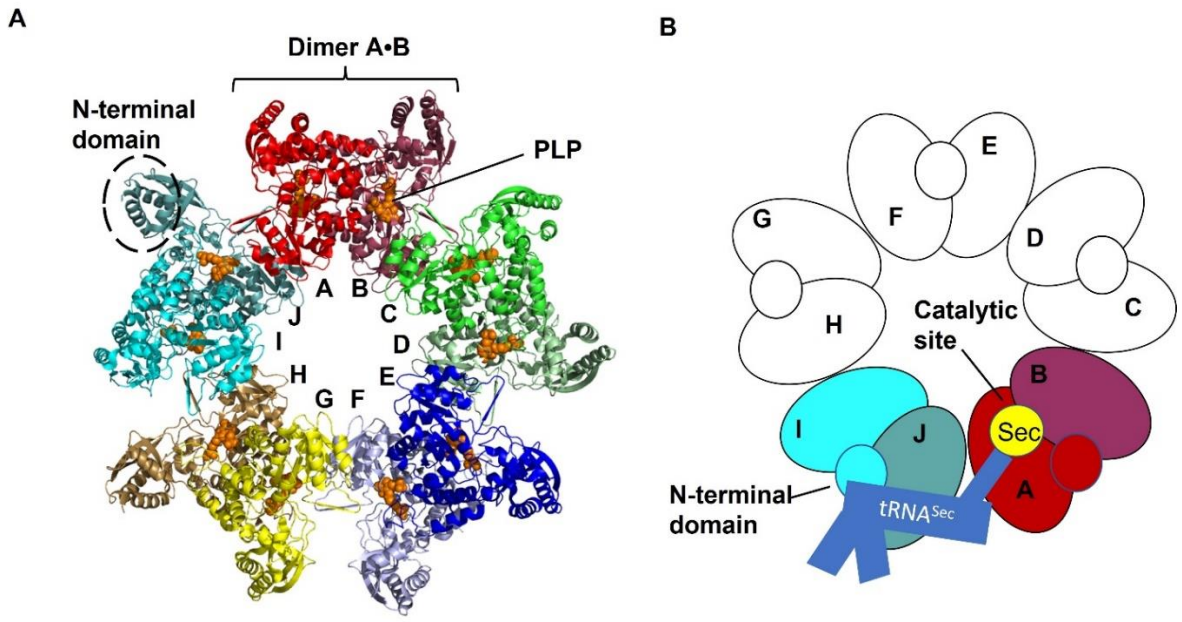


Figure 1. 2 - Structure of SelA. A) SelA color-coded homodecameric ring structure (PDB: 3W1H). PLP is represented as orange sphere models. The N-terminal domain is highlighted. B) Schematic representation of the interaction of $tRNA^{Sec}$ with dimers I•J and A•B. The N-terminal domain of monomer I interacts with $tRNA^{Sec}$ and monomer J helps to orient the tRNA into the catalytic site of monomer A.

Subsequently, Ser-tRNA^{Sec} is converted to Sec-tRNA^{Sec}. In bacteria, the conversion of Ser to Sec is catalyzed by a pyridoxal-5´ phosphate (PLP)- dependent enzyme, selenocysteine synthase (SelA). Structural studies of *A. aeolicus* SelA revealed that SelA forms a star-shape ring structure of approximately 500 kDa, composed of 10 identical subunits (Figure 1.2) [19]. The homodecameric quaternary structure is comprised of 5 pairs of dimers that collectively interact upon the binding of one tRNA^{Sec} to the protruded N-terminal domain of a given monomer. This, triggers a cascade of conformational changes that leads to the binding up to10 tRNA^{Sec} substrates and the conversion of Ser to Sec in a one step reaction, using the product of the selenophosphate synthetase SelD enzymatic reaction, selenomonophosphate, as the selenium donor.

Given the complexity of the SelA activity, it can be best explained as illustrated in Figure 1.2B, where each monomer is alphabetically labelled from A to J, resulting in the A•B, C•D, E•F, G•H and I•J dimers. The N-terminal domain of monomer A first binds to the D-arm of the L-shaped tRNA^{Sec} while monomer B stabilizes the acceptor stem charged with Sec into the catalytic site of monomer C, being accessible due to its interaction with monomer D. Thus, SelA activity is mainly accredited to dimer-dimer interactions between 4 subunits, where one dimer mainly holds a tRNA^{Sec} in place while the neighbor dimer provides the catalytic site, leading up to 10 tRNA^{Sec} bound. However, although two pairs of SelA dimers are involved in catalysing Ser to Sec conversion, a complete synergy between all monomers to form the decameric structure is required as it provides the relative positioning between dimers, ensuring the right space for tRNA^{Sec} to approach the catalytic site.

1.2.2. Stop codon recoding by Sec in *E. coli*

Selenocysteine is co-translationally incorporated into the growing peptide, resulting in the synthesis of Sec-containing proteins. During this process, Sec-tRNA^{Sec} is delivered to the

ribosome by the specialized elongation factor SelB, a GTP-binding protein which is a member of the family of translational GTPases (e.g. EF-Tu & EF-G). The SelB•GTP complex binds to the acceptor-and T-arm of Sec-tRNA^{Sec} through the of N-terminal domains 2 and 3 of SelB [20]. This interaction of SelB•GTP with Sec-tRNA^{Sec} is much tighter with a K_D in the picomolar range compared to binding of deacylated tRNA^{Sec} or even Ser-tRNA^{Sec} [21].

The SelB•GTP•Sec-tRNA^{Sec} ternary complex then binds to the Selenocysteine insertion sequence (SECIS) in the mRNA prior to the ribosome reaching the site, tightly assisted by the special N-terminal domain 4 of SelB and the hair-pin 3D structure of SECIS. Once the ribosome reaches the UGA codon, Sec-tRNA^{Sec} is properly oriented in the ribosomal A site, and the anticodon-codon base pairing triggers GTP hydrolysis by SelB [20]. This causes a conformational change in SelB that leads to the release of Sec-tRNA^{Sec} and the subsequent incorporation of Sec at the UGA codon.

In *E. coli*, variation in the efficiency of UGA recoding into Sec has been reported *in vivo*, ranging from ~5 to 60% depending on medium condition, DNA vectors and cell growth rate [22]. Early analysis suggested that the low efficiency was mainly caused by competing translation termination by RF2. However, this hypothesis was rejected as both, *in vivo* and *in vitro* studies demonstrated that RF2 does not behave as an inhibitor of Sec incorporation, potentially due to the presence of the SelB•GTP•Sec-tRNA^{Sec} ternary complex on the SECIS element blocking RF2 access to the stop codon [22]. However, Sec incorporation, in general, remains an inefficient process when compared to the translation of sense codons with the standard amino acids. Although the exact cause remains unclear, it is likely due to the inherently inefficient activity of the Sec incorporation machinery.

1.3. $tRNA^{Sec}$ and post-transcriptional modifications

Transfer RNAs are essential molecules involved in protein synthesis. After being synthesized, they must undergo several steps during the maturation process to gain their functional state. This includes, among other steps, the incorporation of the CCA sequence at the 3´-end as well as several rounds of chemical modification of canonical nucleotides. The 3´-CCA sequence provides the site for amino acid attachment and for interactions with the ribosome [23]. On the other hand, chemical modifications ensure stable folded structures in addition to providing critical interactions with the ribosome and mRNAs during translation [24].

The chemical modifications of tRNA occur post-transcriptionally. Among all classes of RNAs, tRNAs are the most highly modified, where more than 100 chemically distinct modifications have been reported [25]. Generally, these modifications tend to be clustered at the anticodon loop and at the tRNA elbow, which is formed by the interaction between the D and T loops. In *E. coli* $tRNA^{Sec}$, four post-transcriptional modification have been observed: 5,6-dihydrouracil (D_{20}), N^6-isopentenyladenosine (i^6A_{37}), 5-methyluridine (T or m^5U_{54}) and pseudouridine (Ψ_{55}) (Figure 1.3) [25]. D_{20}, m^5U_{54} and Ψ_{55} are highly conserved modifications, which along with the corresponding modifying enzyme activity collectively contribute to the flexibility, stability as well as the proper folding and maturation of many tRNA [26, 27, 28]. Notably, i^6A_{37} is only present in $tRNA^{Sec}_{UCA}$ and has been suggested to be important during translation of selenoproteins.

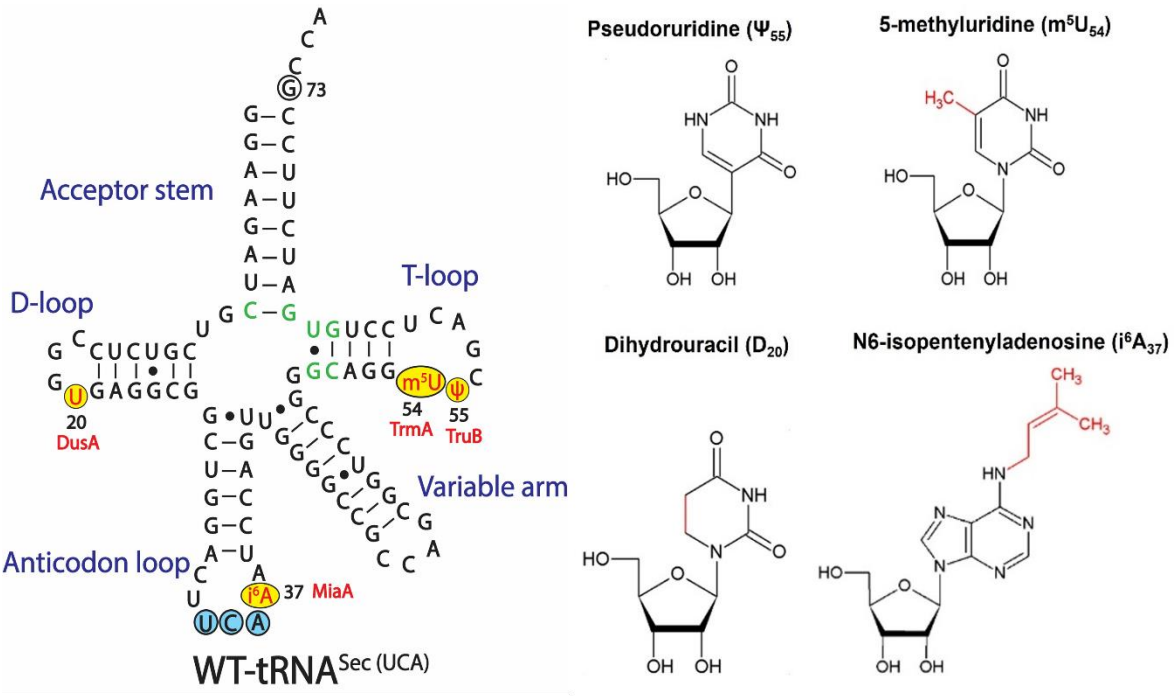


Figure 1. 3 - *E. coli* tRNA^{Sec} structure and post-transcriptional modifications. Secondary cloverleaf structure of tRNA^{Sec}, highlighting general structural features as well as the identity, and location of post-transcriptional modifications with the corresponding modifying enzymes stated in red. Chemical structure of tRNA^{Sec} modifications: Pseudouridine (Ψ_{55}), 5-methyluridine (m^5U_{54}), Dihydrouracil (D_{20}), and N6-isopentenyladenosine (i^6A_{37}).

1.3.1. Dihydrouracil (D20)

Dihydrouracyl is one of the most abundant modified bases in tRNA across all domains of life [25]. Although the content and location of D nucleoside varies within organisms and tRNA species, it is generally observed in the D-loop, contributing to the nomenclature of the tRNA structure. It is catalyzed by flavin-dependent dihydrouridine synthases (Dus), which reduce the C5-C6 double bond in uridine. In prokaryotes, there are three Dus enzymes with conserved catalytic sites, namely DusA, DusB, and DusC, which collectively introduce the D nucleoside to positions 16, 17, 20, 20a and 20b in tRNA D-loop [29, 30].

In general, D modifications resist the stacking of nucleotides due the non-planar configuration in the absence of C5=C6 double bond, promoting structural flexibility while facilitating internal interactions within the tRNA [26]. Thus, although D is highly conserved in all tRNAs, this modification could be relevant in the case of $tRNA^{Sec}$ as it must be flexible enough to successfully interact with multiple enzymes throughout the entire Sec biosynthesis and incorporation machinery.

1.3.2. 5-methyluridine (T or m^5U54) and pseudouridine (Ψ55)

Like the D nucleoside, m^5U_{54} and Ψ_{55} are well conserved modifications, being found within the T-arm of every *E. coli* tRNA. They are catalyzed by the methyl transferase TrmA and the pseudouridine synthease TruB, respectively. Independent of their modifying activity, both TrmA and TruB, have been characterized as chaperones, assisting the folding of tRNAs into their overall ternary structure [27, 28]. During methylation, U_{54} flips out of the loop, allowing the S-adenosylmethionine (SAM)-dependent TrmA enzyme to form a covalent bond with C6 in the target base utilising a catalytic cysteine residue [31]. This covalent bond then facilitates the transfer of the methyl group from SAM to the uracil ring, resulting in the subsequent formation of the m^5U_{54} modification.

On the other hand, Ψ_{55} is formed via C5-ribosyl isomerisation of uridine catalyzed by TruB [27]. Ψ_{55} is unique among modified nucleosides in possessing a C-C bond rather than the usual N-C glycosyl bond that links base and the sugar ring, providing greater rotational flexibility and a free N1-H that can act as an additional hydrogen donor. Ψ_{55} makes a long-range contact with the highly conserved G18 in the D loop [55], contributing to the proper folding of the tRNAs into the L-shape structures. Moreover, Ψ_{55} has been suggested to interacts with other modifications

within the T-arm such as the m^5U_{54} to participate in the formation of the complex circuits of hydrogen interactions that ultimately facilitate and maintained proper folding of the elbow region of the L-shape tRNA structure [56]. Lastly, the extra H-bond donor present within the Ψ modification provides greater tRNA stability by enhancing the stacking properties of the tRNA while providing more tolerance to thermal stress [26].

1.3.3. N6-isopentenyladenosine (i^6A37)

Over 25 tRNA modifications have been determined in *E. coli* [25]. The most complex and bulky modifications are located in the anticodon loop, particularly at the base 34 reading the wobble base and the nucleotide at position 37. In general, modifications in these key locations have been shown to play important roles, regulating efficiency and fidelity during translation as they counteract some of the weaker codon-anticodon interactions [32]. One of the early modifications identified at position 37 was i^6A [33]. In *E. coli*, most tRNAs reading codon U (Cys, Leu, Phe, Ser, Trp, Tyr, Leu, Sec) containing the A36-A37-A38 motif are isopentylated by the dimethelylallyltransferase MiaA. However, $tRNA^{Sec}$ is the only tRNA that retains the i^6A modification since the other tRNAs are subsequently methylthiolated by MiaB, forming $ms^2i^6A_{37}$ [34].

To evaluate the effect of tRNA modification machinery on the efficiency of UAG stop codon re-assignment *in vivo*, 42 knockout strains lacking single, or multiple tRNA modification enzymes were prepared [35]. The result of this study shows that deletion of MiaA, diminishes stop codon read through by approximately 2-fold. On the other hand, the dual overexpression of both MiaA and TruB slightly lowers the recombinant protein yield; however, it improves translation fidelity by 82-88 %. Thus, the results highlight the positive role of tRNA modification in efficient stop

codon re-assignment, where the overexpression of TruB is hypothesized to contribute to proper folding of the tRNA. In eukaryotes, a A37G $tRNA^{Sec}$ mutant, which prevents the isopentylation reaction, abolishes translation efficiency of selenoprotein biosynthesis in Chinese hamster ovary (CHO) cell line *in vivo* [36]. However, it is worth mentioning that some eukaryotic $tRNA^{Sec}$ contain additional modifications at the wobble position such as mcm^5Um, which was observed in this study to play slightly more important role in selenoprotein translation than i^6A [36]. Therefore, although slightly different sets of modifications exist in the anticodon loop of $tRNA^{Sec}$ across different species, the finding in these studies strongly suggest that the i^6A_{37} modification is important for efficient selenoprotein synthesis.

1.4. Selenoprotein and Synthetic biology

Sec-containing proteins play crucial roles in all domains of life. They are generally favoured in redox reactions. Thus, they are of general interest due to their involvement in human health and disease regulation and in the field of proteins bioengineering, where they can improve existing technologies or open new doors for novel applications. Biological energy/fuel cells are an emerging and promising technology aiming to contribute to a cleaner environment while producing either electricity or fuel sources. A prominent example are the enzymatic fuel cells, where a significant effort has been put in place into enhancing the efficiency of the enzyme-orchestrated circuits. Recently, an improvement in these systems was made by implementing a formate/O_2 fuel cell, were the *E. coli* Sec-containing formate dehydrogenase (FDH) plays a critical role as a bioanode in the oxidation of formate [37]. Another area that has become a priority, particularly in the field of energy research, is the development of solar-driven bioenergy

cells. One of the early but exiting examples in this area is the bioelectrochemical cell created with FDH as a cathode member coupled with photosystem II in the anode [38]. Thus, this system artificially mimics the coupling of water oxidation with CO_2 reduction that sustains life while having the potential to function as a CO_2 clustering devise and producing valuable chemical.

Nevertheless, additional research and development is required to bring these emerging technologies into the broader market. Currently, significant effort has been made into finding the right enzymes in nature to perform the desired reactions while determining the proper physical-chemical environment to enhance the efficiency of the system. One area that could be explored, however, is enhancing enzyme activity via protein bioengineering, where the Sec amino acid is of high interest due to its redox capabilities. Nevertheless, utilising a natural selenoprotein biosynthesis and insertional machinery to bioengineer selenoproteins poses some limitations due to the inherent complexity of the system.

1.4.1 Limitation of the natural selenoprotein biosynthesis machinery

Despite being intensively studied *in vivo*, the Sec biosynthesis and insertion machinery fails to provide a universal platform for selenoprotein bioengineering with reproducible results. Although selenium and $tRNA^{Sec}$ concentrations play a role, the unreproducible results of selenoprotein biosynthesis is mostly due to the SECIS element [36], which is protein- and organism-specific. The bacterial SECIS element is located in the mRNA coding sequence, limiting the flexibility to add or move Sec residue freely within the polypeptide chain. To be able to do so, a proper SECIS element must be engineered without changing the protein's amino acid sequence, which could be extremely challenging and might not be possible in a desired location. Moreover, one will have to engineer a unique SECIS for every protein, which is time consuming and could affect translation efficiency. In fact, this was proven to be the case when SECIS

elements from various representative selenoproteins, cloned into a luciferase reporter system, resulted in different translation efficiencies *in vivo* [36].

In eukaryotes, unlike in the bacterial system, the SECIS element is located at the 3´-untranslated region. In this case the mRNA, with the help of additional proteins, bends to make the SECIS available to the ribosome during translation [39]. Thus, not only the sequence identity but also the distance between the SECIS and the designated stop codon likely need to be considered in a bioengineering project. As a result, recombinantly synthesizing mammalian selenoproteins in prokaryotes such as *E. coli*, which could be desirable when producing selenoprotein for characterization *in vitro* has been challenging.

1.4.2 Engineered selenoprotein biosynthesis machinery: a viable solution for the flexible incorporation of selenocysteine

To overcome the limitation of the SECIS element, strong effort has been made to engineer the selenoprotein biosynthesis machinery into a SECIS-independent mechanism. A thorough understanding of the interactions within the Sec biosynthesis and insertional machinery at the molecular level, has led the way to engineering synthetic tRNAs that are recognized and brought to the ribosome by EF-Tu [40, 41]. This approach eliminates the requirement of the SelB-SECIS interaction to assist with Sec incorporation, paving a strong foundation for selenoprotein bioengineering across all domains of life using the *E. coli* machinery.

Among some of the tRNA bioengineering projects in the field, one of the most prominent results were achieved using the synthetic $tRNA^{UtuX}$. It was derived from key features between $tRNA^{Sec}$ and $tRNA^{Ser}$ [40] taking into consideration interactions with three protein factors: SerRS, SelA and EF-Tu. In comparison with $tRNA^{Sec}$, $tRNA^{UtuX}$ shares similar structural features such as the

length of the D-arm and sequence identity of the acceptor stem (Figure 1.4). Moreover, $tRNA^{Sec}$ and $tRNA^{UtuX}$ share identical D- and T-loop sequences along with the A36-A37-A38 motif; therefore, both tRNAs are likely to share similar modification profile. However, unlike $tRNA^{Sec}$, $tRNA^{UtuX}$ has a shorter variable arm (6 bp vs. 8 bp), a CUA anticodon and a different antidiscrimination box to be recognized by EF-Tu rather than SelB (green nucleotides in $tRNA^{Sec}$ and blue nucleotides in $tRNA^{UtuX}$, Figure 1.4) [10], allowing EF-Tu-mediated Sec incorporation in response to UAG stop codon (Figure 1.4A-B). In vivo and in vitro studies demonstrated that $tRNA^{UtuX}$ facilitated Sec incorporation with high fidelity, producing active selenoprotein with similar enzymatic activity to those synthesized with $tRNA^{Sec}$ [40]. Thus, the authors showed $tRNA^{UtuX}$ not only solved the 30% Serine misincorporation event observed with other previously engineered tRNAs but also that the EF-Tu mediated mechanism produced selenoprotein with similar activity to those mediated by SelB in vivo. Moreover, the coupled transcription-translation cell-free system showed that the active selenoprotein yield was approximately 2-fold higher when using the repressor $tRNA^{UtuX}$ rather than $tRNA^{Sec}$ [40]. Nevertheless, it is important to highlight that high concentration of Sec-$tRNA^{UtuX}$ was used for optimal conditions. Thus, it remains unclear weather $tRNA^{UtuX}$ poses some improvement in selenoprotein yield relative to $tRNA^{Sec}$ at comparable concentrations *in vitro*.

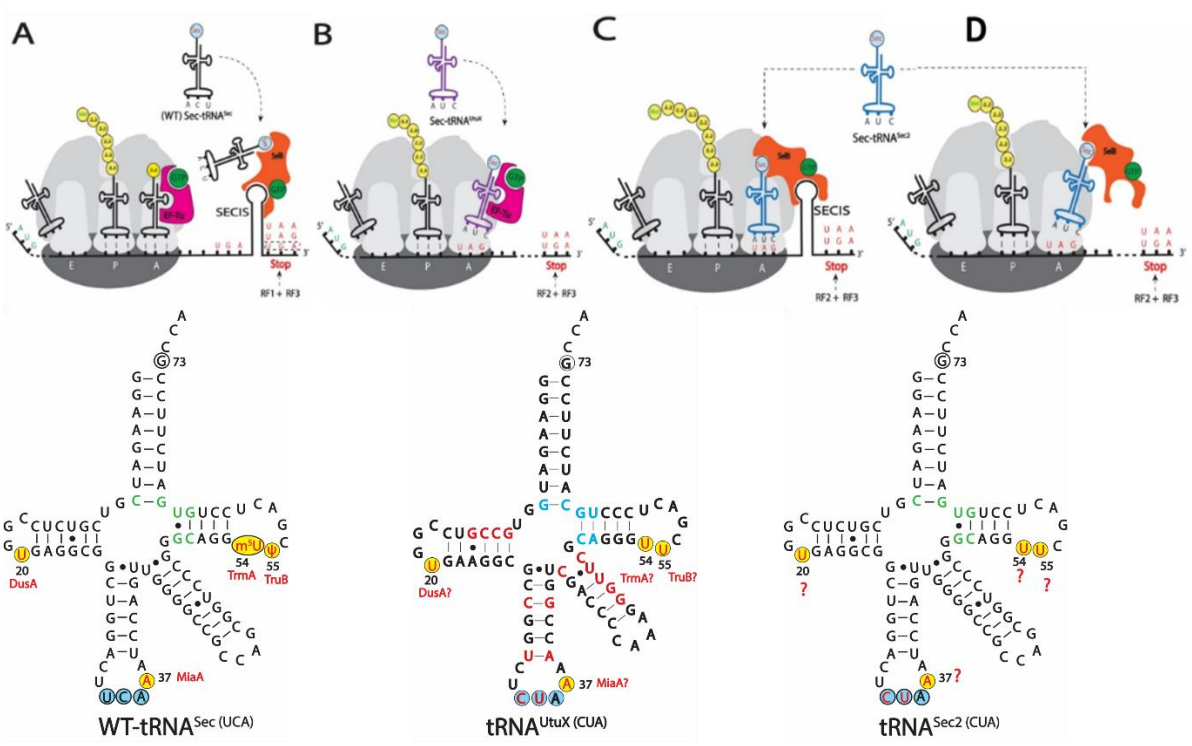


Figure 1. 4 - Comparison of *E. coli* natural and engineered Sec biosynthesis and insertion machinery. A) Natural SECIS-dependent SelB-mediated mechanism. Incorporation of Sec by $tRNA^{Sec}$ in response to UGA stop codon. B) SECIS-independent EF-Tu-mediated mechanism with $tRNA^{UtuX}$ and UAG stop codon. C) SECIS-dependent SelB-mediated with $tRNA^{Sec2}$ reading UAG stop codon. D) SECIS-independent SelB-mediated mechanism with $tRNA^{Sec2}$. Structures of $tRNA^{Sec}$, $tRNA^{UtuX}$ and $tRNA^{Sec2}$. Anti-determinant box highlighted in green for $tRNA^{Sec}$ and $tRNA^{Sec2}$, and in blue for $tRNA^{UtuX}$. Red nucleotides in $tRNA^{UtuX}$ represent changes relative to $tRNA^{Sec}$.

Another relevant discovery in the field is sense codon recoding *in vitro*. Although it has been shown that Sec can be incorporated using multiple sense codons *in vivo* [42], only *in vitro* translation experiments were able to produce selenoenzymes with detectable activity [40]. Moreover, all sense codons tested (CGU, GCC, CCU, UCG) *in vitro* with $tRNA^{Sec}$ variants resulted in active selenoproteins with comparable activity to wildtype UGA stop codon [40, 41], suggesting a potential benefits of cell-free system as a tool for selenoprotein bioengineering using sense codons.

Interestingly, a breakthrough in the field of selenoprotein biosynthesis was arguably achieved when mammalian selenoproteins were synthesized using an *E. coli* SelB-mediated SECIS-independent mechanism *in vivo* (Figure 1.4D) [43]. The results of this study showed that $tRNA^{Sec2}$, a variant of $tRNA^{Sec}$ with its anticodon changed for reading the UAG codon (Figure 1.4), was able to incorporate Sec at a predefined UAG codon. Although protein yields and translation fidelity of the system pale in comparison to those assisted with the SECIS element, it has been suggested that optimising SelB activity independently of the SECIS element could provide a unique tool for selenoprotein bioengineering due its strong affinity and specificity towards Sec-$tRNA^{Sec}$ [10].

Currently, several orthogonal systems have been developed for the incorporation of non-canonical amino acids utilizing genetic code expansion technologies [44, 45]. Typically, these systems use engineered tRNA/aminoacyl-tRNAs (aaRS) synthase pairs that incorporate the non-canonical amino acids at pre-designed UGA stop codon. While the molecular mechanism remains unclear, *in vitro* experiments indicates that UAG stop codon generates selenoprotein with higher enzymatic activity than the counterpart UGA codon [40, 41], suggesting that the UAG is not only a superior codon for Sec incorporation but potentially also for stop codon re-assignments with other non-canonical amino acids.

Recently, a complete plasmid-encoded selenoprotein system was developed using orthogonal SelA, SelD and synthetic allo-$tRNA^{Utu2D}$ for EF-Tu-mediated SECIS-independent Sec incorporation in response to UAG stop codon [46]. The effectiveness of this newly developed orthogonal system relies on the improvement on EF-Tu affinity towards the engineered $tRNA^{Utu2D}$ substrate over the previously developed $tRNA^{Sec2}$ and $tRNA^{UtuX}$. Both $tRNA^{Sec2}$ and $tRNA^{UtuX}$ contains a 13-bp acceptor branches (comprising of the acceptor stem plus the T-stem,

Fig 1.3 and 1.4), making them to act as relatively poor substrates for EF-Tu which generally prefers tRNAs with 12-bp stem branches. To overcome the limitation of *E. coli* SelA inability to successfully recognize the 12-bp stem branch in $tRNA^{Utu2D}$ and ensuring high selenium concentrations, *Aeromonas salmonicida* SelA and SelD are implemented in the system. However, while $tRNA^{Utu2D}$ only interact with the *A. salmonicida* SelA, it is still recognized and aminoacylated by SerRS, ensuring partial utilization of *E. coli* Sec biosynthesis machinery and the subsequent and more efficient EF-Tu mediated Sec incorporation. More importantly, although with relative low protein yield, this orthogonal system was successfully implemented to substitute 4 catalytic cysteine residues for selenocysteine of a mammalian protein using *E. coli* as the host organism [47]. Thus, this partial orthogonal system not only was able to by-pass specie barrier by converting a non-selenium containing mammalian protein into a selenoprotein, it also expanded the number of Sec residues that can be incorporated into a protein sequence.

1.5. Genetic code: Pyrrolysine, another unusual amino acid

The genetic code is composed of four nucleotide bases of RNA which are are congregated into 64 3-base codons, were 61 of these codons in general are used to code for 20 amino acids. The remaining three codons known as stop codons (UGA, UAG, UAA) are used as a termination signal. However, there are special cases where those stop codons are re-assigned to code for amino acids in nature. One example of those special amino acids is Selenocysteine, which uses UGA for sense codon translation assisted by the SECIS element as previously described. Thus, selenocysteine have become known as the 21st natural encoded and special amino acid. However, another amino acid that uses stop codon re-assignment, hence considered as the 22nd special

amino acid, is Pyrrolysine (Pyl). Pyrrolysine is found in a small number of methylamine metabolizing archaea. Unlike Selenocysteine, Pyrrolysine is not synthesized on its cognate $tRNA^{Pyl}$ but rather exists as free amino acid, which is loaded onto $tRNA^{Pyl}$ by pyrrolysyl-tRNA synthase (PylRS) and is incorporated into proteins in response to in-frame UAG codons during translation [57]. Moreover, although a homologue of the SECIS element for Pyl incorporation, known as pyrrolysine insertion sequence (PYLIS), was initially proposed, it was later shown that PYLIS was not essential for Pyl incorporation [57]. Therefore, although translation of selenocysteine and pyrrolysine both involves stop codon suppression, their amino acid biosynthesis and incorporation mechanisms have little in common in nature. While the Sec mechanism re-assigns the UGA stop codon to Sec with the help of the SECIS element, a small group of methanogenic archaea appears to constantly use UAG as both, stop and sense codon without the requirement of an mRNA structure. However, with the engineered repressor $tRNA^{UtuX}$ and the SECIS-independent EF-Tu mediated mechanism, the differences between Sec and Pyl incorporation is reduced since both systems use in-frame UAG codons to incorporate their corresponding amino acid without the assistance of an insertion sequence (SECIS or PYLIS).

Nevertheless, understanding Pyl and its incorporation mechanism led the way to developing valuable tools for genetic code expansion. The archaeal PyRL/tRNAPyl pair have not only been shown to be orthogonal in bacteria but also quite flexible to incorporate non-canonical amino acids, which is likely due to the PylRS high substrate side chain promiscuity and high tolerance for mutations in the substrate binding pocket to accommodate a wide range of unnatural amino acids [57-58].

1.6. Objectives

The field of selenoprotein bioengineering has come a long way from the initial SECIS engineering approach to implementing SECIS-independent mechanisms. Great efforts have been made for Sec incorporation using engineered machineries primarily through UAG amber stop codon suppression. However, despite a growing body of evidence that suggest positive roles of tRNA modifications during translation [32, 34, 35, 36], the role of $tRNA^{Sec}$ modifications for wild-type or engineered selenoprotein biosynthesis machineries remain vastly unexplored. Therefore, the objective of this thesis is to investigate the importance of $tRNA^{Sec}$ post-transcriptional modifications for selenoprotein biosynthesis. To investigate the effect of tRNA modifications, RNA modifying enzymes along with individual components of the selenoprotein biosynthesis and insertional machinery will be generated. By conducting binding and activity assays between the modifying enzymes and tRNAs, including wild type $tRNA^{Sec}$ and synthetic variants, one may gain valuable information regarding the interaction between tRNAs and modifying enzymes and the potential modification profile of the synthetic tRNAs. Moreover, by supplementing *in vitro* translation reactions that monitor stop codon readthrough in the presence of tRNA modifying enzymes, one could potentially infer the effect of the tRNA modifications on the overall translation of selenoproteins. Thus, in this research, I aim to narrow the existing gap in the literature by characterising $tRNA^{Sec}$ modifications and their role in translation of selenoprotein biosynthesis, potentially contributing to the optimization of future cell-free selenoprotein bioengineering platforms.

Chapter 2: MATERIALS METHODS

2.1 Materials

All chemicals were purchased from Fisher Scientific. [5-^{3}H] UTP for *in vitro* transcriptions was purchased from Moravek Biochemicals and chromatography materials from GE Healthcare. Oligonucleotides were obtained from Integrated DNA Technology (IDT), and fluorescence reporter constructs were synthesized and purchased from GeneWiz. Plasmid encoding for *E. coli* SelA and SelD were a kind gift from M. Rodnina lab (Max Planck Institute for Biophysical Chemistry).

2.2 Protein Expression and Purification

ME5305 E. *coli* cells from the ASKA collection containing pCA42N plasmid encoded hexahistidine-tagged MiaA and SelB, E. *coli* Lemo21(DE3) containing Pca42N plasmid encoded hexahistidine-tagged SerRS, E. *coli* BL21 (DE3) containing plasmid encoded SelA and hexahistidine-tagged SelD [22], were cultured in LB (Luria Broth) media supplemented with 35 µg/mL chloramphenicol (for MiaA, SerRS, and SelB) and 50 µg/mL Kanamycin (for SelA and SelD) at 37 °C. At OD_{600nm} of ~0.6, protein expression was induced with isopropyl β-D-1-thiogalacropyronoside (IPTG) to a final concentration of 1 mM followed by overexpression overnight (~16 hours) at 18 °C. Overexpression success was validated by sodium dodecyl

sulphate-polycacrylamide gel electrophoresis (SDS-PAGE), and the cells were harvested by centrifugation at 5000 x g for 20 minutes, flash frozen in liquid nitrogen, and stored at -80 °C.

SelB, SelD, MiaA and SerRS were initially purified via Ni-Sepharose affinity purification in batch. To lyse the cells, cell pellets were thawed on ice and resuspended in 5 ml/g (cells) of buffer A (20 mM Tris-HCl pH 8.1, 400 mM KCl, 5 % (v/v) glycerol, 1 mM β-mercaptoethanol (β-Me), 0.5 mM PMSF, 30 mM imidazole), incubated with 1 mg/ml of lysozyme for 30 min at 4 °C followed by 1 hour incubation with 12.5 mg/g (cells) of sodium deoxycholate at 4 °C. To enhance the cell opening event, cell lysates were sonicated (intensity level 6, duty cycle at 60%) while on ice for 1-min intervals for a total of 10 minutes. Subsequently, the cell lysates were centrifuged at 30,000 x g for 30 min at 4 °C, and the protein-containing supernatants were applied to buffer A-equilibrated 5 ml-Ni^{2+} Sepharose resin and incubated for 1 hour at 4 °C. To remove unbound and low-affinity proteins, the resin was collected by centrifugation at 500 x g for 5 min at 4 °C and washed 6 times with buffer A, using 5 min shaking incubation periods at 4 °C followed by centrifugation at 500 x g for 2 min while decanting supernatants in between washes. Subsequently, proteins were eluted 6-8 times with 90% resin volume of buffer B (20 mM Tris-HCl pH 8.1, 400 mM KCl, 5 % (v/v) glycerol, 500 mM imidazole), using the same incubation and centrifugation procedure as in the washing step explained above. The presence of the target proteins within the collected fractions was validated via SDS-PAGE, and MiaA and SelB fractions were pooled and concentrated using Vivaspin concentrator (Vivaspin 30, MWC 30 kDa) while SerRS and SelD where rebuffered in buffer C (50 mM HEPES, pH 7.5, 70 mM NH_4Cl_2, 30 mM KCL, 7 mM $MgCl_2$, 2 mM DTT) and similarly concentrated using the Vivaspoin concentrator. Subsequently, MiaA and SelB where further purified via Superdex 75 size exclusion chromatography (XK 26/100 column, GE Healthcare) in buffer D (20 mM

HEPES, pH 7.5, 600 mM KCL, 20% Gycerol, 1mM EDTA), and selected fractions were analysed by SDS-PAGE, pooled and concentrated as indicated above.

First, SelA cell pellet was resuspended in 7 ml/g (cells) of buffer H (50 mM potassium phosphate pH 7.4, 3 mM DTT, 1 mM EDTA, 10 µM pyridoxal 5-phosphate, 100 µg/ml FMSF) and subsequently cells were lysed, and the cleared cell lysate was prepared by centrifugation as described above. To remove contaminating proteins while maintaining SelA decameric structure under non-denaturing conditions, SelA was subject to ammonium sulfate precipitation as described previously [21, 48]. The supernatant was then diluted to 300 ml in buffer H and incubated for 3 hours with ground ammonium sulfate, which was slowly added to 25% saturation (154 g ammonium/ L of solution) at 4 °C as described [49]. Afterwards, the protein-ammonium sulfate pellet was collected by centrifugation at 20,000 x g for 20 min at 4 °C and resolubilized in buffer I (20 mM potassium phosphate pH 7.5, 2 mM DTT, 0.5 mM EDTA, 10 µM pyridoxal 5-phosphate). To remove the residual precipitate and collect clarified proteins, another centrifugation step was performed at 20,000 x g for 10 min at 4 °C followed by concentration of the supernatant using a Vivaspin concentrator (Vivaspin 30, MWC 30 kDa). Subsequently, SelA was further purified by Sephacryl S-300 size exclusion chromatography in buffer I (20 mM potassium phosphate pH 7.5, 2 mM DTT, 0.5 mM EDTA, 10 µM pyridoxal 5-phosphate), validated by Coomassie-blue stained SDS-PAGE, and finally concentrated with a Vivaspin concentrator.

Final protein concentration was determined using comparative SDS-PAGE and A_{280} using predicted molar extinction coefficients (ε) (SelA, 35785 M^{-1} cm^{-1}; SelB, 81025 M^{-1} cm^{-1}; SelD, 14815 M^{-1} cm^{-1}; MiaA, 39545 M^{-1} cm^{-1}; SerRS, 34630 M^{-1} cm^{-1}) from ExPASy ProtParam tool and protein sequences (UniProt Consortium, 2015).

2.3 *In vitro* transcription and tRNA purification

Double-stranded DNA (dsDNA) templates for *in-vitro* transcription of tRNAs ($tRNA^{Sec}$, $tRNA^{Sec2}$ and $tRNA^{UtuX}$) were generated via primer extension PCR as indicated (Table 2.1). To prevent any PCR reagent from interfering with downstream applications, PCR products were cleaned up using EZ-10 Spin Purification kit (BioBasic Inc.) followed by determining their concentration (BioDrop DUO UV/VIS Spectrophotometer; Biochrom, Ltd.) and verified by 12% DNA-PAGE (60V for 30 minutes). Subsequently, *in vitro* transcription reactions to prepare the target tRNAs were conducted for 4 hours at 37°C using 1 ng/µL of purified PCR template in transcription buffer (40 mM Tris-HCl pH ~7.5, 15 mM $MgCl_2$, 2 mM spermidine, 10 mM NaCl, 10 mM DTT) with 3 mM NTPs (ATP, CTP, GTP, and UTP), 5 mM GMP, 0.01 U/µL inorganic pyrophosphatase (iPPase), 0.3 µM T7 RNA Polymerase, and 0.12 U/µL RNase inhibitor. After transcription, reactions were treated with DNase1 to degrade template DNA for 1 hour at 37°C, and the presence of tRNA transcripts was verified by a 15% Urea-PAGE (40 minutes at 300V) from samples taken throughout the process (0, 15, 30, 60 minutes and after DNase1 digestion). To prepare tritium-labeled tRNAs, [5-^{3}H]UTP was added to the reaction as previously described [51]. To remove proteins from the transcripts, nonradioactive products were purified by phenol-chlorophorm extraction and unincorporated nucleotides added in excess to the reactions were removed via Superdex 75 size exclusion chromatography (XK 26/100 column, GE Healthcare) in $TAKEM_4$ buffer as previously reported [50]. To verify the presence of tRNA within the collected fractions, samples were analyzed by a 15% Urea-PAGE (40 minutes at 300V) and subsequently pooled. To ensure sufficient precipitation, the selected pool of fractions was speed vacuumed to ~2-3 µM and consequently concentrated by ethanol precipitation. Radioactive

tRNAs were purified by Nucleobond Xtra Midi anion-exchange gravity column (Macherey-Nagel) as previously described [51] and similarly concentrated by ethanol precipitation. The final tRNA concentrations were determined by A_{260} using IDT OligoAnalyzer to predict molar extinction coefficients of 852300 M^{-1} cm^{-1} for $tRNA^{Sec}$ and $tRNA^{Sec2}$, and 829400 M^{-1} cm^{-1} for $tRNA^{UtuX}$ and the specific activity of [^{3}H]-tRNAs was determined by scintillation counting.

Table 2. 1- Primer extension PCR - composition and parameters.

Reaction Mixture	**Reagent**		**Final Concentration**	
	MilliQ H_2O		-	
	Q5 DNA Polymerase		40 U/ml	
	Forward Primer		0.9 µM	
	Reverse Primer		0.9 µM	
	dNTPs		0.2 µM	
	Q5 Reaction Buffer (5X)		1X	
Thermocycle Settings	**Cycle Steps**	**Temperature (°C)**	**Time**	**# of Cycles**
	Initial Denaturing	98	2 min	1
	Denaturing	98	10 s	30
	Annealing	70	30 s	
	Extension	72	30 s	
	Final Extension	72	3 min	1
	Hold	4	∞	-
Primer Sequences*	$tRNA^{Sec}$ (Forward): 5´- **taatacgactcactata**GGAAGATCGTCGTCTCCGGTGAGGCGGCTGGACT**TCA**AATCCAGTTGGGGCCGCC-3´ $tRNA^{Sec2}$ (Forward): 5´- **taatacgactcactata**GGAAGATCGTCGTCTCCGGTGAGGCGGCTGGACT**CTA**AATCCAGTTGGGGCCGCC-3´ $tRNA^{Sec/Sec2}$ (Reverse[1]): 5´- TGGCGGAAGATCACAGGAGTCGAACCTGCCCGGGACCGCTGGCGGCCCCAACTGG-3´			
	$TRNA^{UtuX}$ (Forward): 5´ - **taatacgactcactata**GGAAGATGGTGCCGTCCGGTGAAGGCGCCGGTCT**CTA**AAACCGGTCGACCCGAAAGG-3´ $TRNA^{UtuX}$ (Reverse): 5´-TGGCGGAAGATGCAGGGAGTCGAACCCTGCGAACCCTTTCGGGTCGACCG-3´			

* T7 promoter sequence in bold lowercase letters. Overlapping sequences between forward and reverse primer underlined. Anticodon sequence in bold-red.

[1] $tRNA^{Sec}$ and $tRNA^{Sec2}$ utilize the same reverse primer.

2.4 Tritium release assay to detect Ψ55 and m^5U54 formation

To determine if TruB and TrmA modify *in vitro* transcribed $tRNA^{Sec}$, $tRNA^{Sec2}$ and $tRNA^{UtuX}$, tritium release single-turnover reactions were conducted as previously described [50, 51]. [^{3}H]tRNAs were refolded in $TAKM_4$ buffer (50 mM Tris-HCl, 70 mM NH_4Cl, 30 mM KCl, 1mM EDTA and 4mM $MgCl_2$) by heating at 60°C for 5 min followed by slow cooling to room temperature. Subsequently, reactions were incubated in $TAKEM_4$ with 0.6 µM tRNA and 5 µM enzyme final concentrations with the inclusion of 50 µM S-Adenosyl methionine (SAM) or reaction with TrmA. After 2, 5, 10 and 30 minutes, reactions were stopped by the addition of 5% Norit A (w/v) in 0.1 M HCL and after separation from tRNAs by absorption to charcoal as describe [51], the level of tritium released into the supernatant corresponding to the level of Ψ55 and m^5U54 formation was determined via scintillation counting. All reactions were done in triplicates.

2.5 Preparation of Ser-tRNAs and Sec-tRNAs

To prepare Ser-[^{3}H]$tRNA^{Sec}$, an aminoacylation reaction was performed similarly as described [52]. Briefly, 0.5 µM of [^{3}H]$tRNA^{Sec}$ was aminoacylated by SerRS (6 µM), L-Serine (30 µM), ATP (5 mM), and inorganic pyrophosphate (0.01 U/µL) in buffer C (50 mM HEPES, pH 7.5, 70 mM NH_4Cl_2, 30 mM KCL, 7 mM $MgCl_2$, 2 mM DTT) for 1 hour at 37 °C. To prepare Sec-[^{3}H]$tRNA^{Sec}$, "one-pot" reaction was performed as previously described [21]. Briefly, 0.5 µM of [^{3}H]$tRNA^{Sec}$ was incubated with SerRS (6 µM), L-Serine (30 µM), ATP (5 mM), inorganic

pyrophosphate (0.01 U/µL), SelA (3 µM), SelD (10 µM), Na_2SeO_3 (250 µM), SelB (6 µM), GTP (2 mM) in buffer C for 1 hour at 37 °C. Both, Ser-[^{3}H]tRNA^{Sec} and Sec-[^{3}H]tRNA^{Sec}, were subjected to phenol-chloroform extraction and ethanol precipitation as described [21]. The tRNA concentration was determined by scintillation counting.

2.6 Nitrocellulose filtration

To determine the interaction of tRNA modifying enzymes MiaA, TruB and TrmA with tRNA^{Sec}, tRNA^{Sec2} and tRNA^{UtuX}, a series of filter binding assays were conducted. [^{3}H]tRNA was initially refolded by heating at 65°C for 5 min and allow to cool down to room temperature for 15 minutes. The refolded [^{3}H]tRNAs (20 nM) were incubated with increasing concentration of enzymes (0, 0.5, 1, 5, 10, 30 µM for TruB D48N and TrmA C324A, and 0, 0.05, 0.08, 1, 5, 10 µM for MiaA) in TAKEM$_4$ buffer (50 mM Tris-HCl, 70 mM NH_4Cl, 30 mM KCl, 1mM EDTA and 4mM $MgCl_2$) for 10 minutes at room temperature and subsequently subjected to nitrocellulose membrane filtration as described [50, 51]. Similarly, to assess the interactions of the components of the Sec biosynthesis and insertional machinery, [^{3}H]tRNA^{Sec} (20 nM) was incubated with increasing concentration of SerRS (0, 0.5, 1, 2, 3, 5 µM) and Ser-[^{3}H]tRNA^{Sec} (20 nM) was incubated with SelA (0, 0.05, 0.1, 0.5, 1, 3, 5, 10, 20 µM). Moreover, Sec-[^{3}H]tRNA^{Sec} (20 nM) was incubated with SelB (0, 0.05, 0.5, 1, 3, 5 µM) using deacylated [^{3}H]tRNA^{Sec} (20 nM) as control. The fraction of tRNA bound to enzyme and retained on the nitrocellulose membrane was determined by scintillation counting as described [51], and the dissociation constant (K_D) was determined by plotting the amount of bound tRNA (*Bound*) as a function of enzyme concentration ([*enzyme*]) and fitting to a hyperbolic equation:

$$\text{Bound} = \text{Bound}_{max} \times [\text{enzyme}] / (K_D + [\text{enzyme}]).$$

Each reaction was carried out in triplicates, the data were averaged and the K_D was determined by fitting the averaged data. Lastly, K_D values wee evaluated as significantly different when the stand deviations did not overlap with each other.

2.7 Coupled *in vitro* transcription-translation (TXTL) reactions

Reporter plasmids: pSB3C5_UGA+SECIS_sfGFP, pSB3C5_UAG-SECIS_sfGFP, pSB3C5_UUC+SECIS_sfGFP, and pSB3C5_UUC-SECIS_sfGFP, were synthesized and cloned by GeneWiz. The TXTL reactions were performed using the PURExpres® Δ RF123 cell-free system (New England Biolabs, #E6850S) according to the manufacture recommendations with the following modifications (Tables 2.2 & 2.3). Briefly, reactions in the absence of tRNA modifying enzymes were performed with 10 µL of solution A, 7.5 µL of solution B, 0.5 µL of RNase inhibitor (RiboLock, Thermo Fisher), 250 ng of plasmid DNA, 0.5 µL of the corresponding release factor (RF3 and RF1 or RF2) supplemented with 5 µM of specific unmodified tRNAs ($tRNA^{Sec}$, $tRNA^{UtuX}$ or $tRNA^{Sec2}$), 15 µM of SelB, 5 µM of SelA and SelD in addition to 50 µM of L-Serine, Na_2SeO_3 and Pyridoxal-5'-phosphate (PLP) in 30 µL of total reaction volume. Test reactions to determine the effect of tRNA modifying enzymes were supplemented with 2 µM of MiaA, TruB and TrmA each in addition to 50 µM of dimethylallyl pyrophosphate (DMAPP) and 50 µM SAM.

All test reactions and the positive control in the presence of SECIS and modifying enzymes were conducted in duplicates while the negative, ΔSelB, $\Delta tRNA^{Sec}$, $\Delta tRNA^{Sec2}$ and the positive

control in the absence of SECIS were performed as single reactions. Individual reactions were incubated for 5 hours at 37°C and placed on ice to stop the reactions. Subsequently, reactions were diluted to a final volume of 150 µL in $TAKEM_7$ buffer (50 mM Tris-HCl pH 7.6, 70 mM NH_4Cl, 30 mM KCl, 7 mM $MgCl_2$), placed in a quartz cuvette (Starna Cells Inc.) and the sfGFP emission spectra were recorded with QuantaMaster Fluorimeter (Photon Technology International (Canada) Inc.). Super-fold green fluorescence protein (sfGFP) was exited at 488 nm and emission scanned between 503-603 nm (λ_{max} = 508 nm; slid width 8 nm).

Table 2. 2 - *In vitro* TXTL test and control reactions of SelB-mediated SECIS-dependent Sec incorporation mechanism.

Components	**UGA+ SECIS**	**UGA + SECIS + modification**	**UGA + SECIS + modification (ΔSelB)**	**UGA + SECIS + modification (ΔtRNA^{Sec})**	**UGA + SECIS (MiaA only)**	**-Ve control (ΔDNA)**
Sol. A: 10 µL	✓	✓	✓	✓	✓	✓
Sol. B: 7 µL	✓	✓	✓	✓	✓	✓
RF1: 0.5 µL	✓	✓	✓	✓	✓	✓
RF2: 0.5 µL						
RF3: 0.5 µL	✓	✓	✓	✓	✓	✓
tRNA^{Sec}: 5µM	✓	✓	✓		✓	✓
tRNA^{Sec2}: 5 µM						
tRNA^{UtuX}: 5 µM						
SelB: 15 µM	✓	✓		✓	✓	✓
SelA: 5 µM	✓	✓	✓	✓	✓	✓
SelD: 5 µM	✓	✓	✓	✓	✓	✓
L-Ser: 50 µM	✓	✓	✓	✓	✓	✓
Na_2SeO_3:50 µM	✓	✓	✓	✓	✓	✓
PLP: 50 µM	✓	✓	✓	✓	✓	✓
MiaA: 2 µM		✓	✓	✓	✓	✓
TruB: 2 µM		✓	✓	✓		✓
TrmA: 2 µM		✓	✓	✓		✓
DMAPP: 50µM		✓	✓	✓	✓	✓
SAM: 50 µM		✓	✓	✓		✓
Plasmid DNA: 250 ng	✓	✓	✓	✓	✓	
Number of replicates (n)	n = 2	n = 2	n = 1	n = 1	n = 2	n = 1

Table 2. 3 - *In vitro* TXTL test reactions of EF-Tu and SelB-mediated SECIS-independent Sec incorporation mechanisms and positive control reactions.

Components	**UAG - SECIS**	**UAG-SECIS + modification**	**UAG - SECIS**	**UAG-SECIS + modification**	**UAG-SECIS + modification (ΔtRNA^{Sec2})**	**UUC+SECIS + modification (+ve control)**	**UUC-SECIS + modification (+ve control)**	**UUC-SECIS (+ve control)**
Sol. A: 10 µL	✓	✓	✓	✓	✓	✓	✓	✓
Sol. B: 7 µL	✓	✓	✓	✓	✓	✓	✓	✓
RF1: 0.5 µL						✓	✓	✓
RF2: 0.5 µL	✓	✓	✓	✓	✓	✓	✓	✓
RF3: 0.5 µL	✓	✓	✓	✓	✓	✓	✓	✓
tRNA^{Sec}: 5µM								
tRNA^{Sec2}: 5 µM			✓	✓				
tRNA^{UtuX}: 5 µM	✓	✓						
SelB: 15 µM								
SelA: 5 µM	✓	✓	✓	✓	✓			
SelD: 5 µM	✓	✓	✓	✓	✓			
L-Ser: 50 µM	✓	✓	✓	✓	✓			
Na_2SeO_3:50 µM	✓	✓	✓	✓	✓			
PLP: 50 µM	✓	✓	✓	✓	✓			
MiaA: 2 µM		✓		✓	✓	✓	✓	
TruB: 2 µM		✓		✓	✓	✓	✓	
TrmA: 2 µM		✓		✓	✓	✓	✓	
DMAPP: 50 µM		✓		✓	✓	✓	✓	
SAM: 50 µM		✓		✓		✓	✓	
Plasmid DNA: 250 ng	✓	✓	✓	✓	✓	✓	✓	✓
Number of replicates (n)	n = 2	n = 2	n = 2	n = 2	n = 1	n = 2	n = 1	n = 1

Chapter 3: RESULTS

3.1 Purification of Selenocysteine biosynthesis and insertion machinery

To analyze the effect of tRNA post-transcriptional modification on selenoprotein biosynthesis, components of the Sec biosynthesis and insertion machinery were individually purified. To prepare tritium-labelled and unlabelled tRNAs (wild-type and variants), double-stranded DNA (dsDNA) templates were first generated by primer extension PCR. As illustrated by the Ethidium bromide (Et-Br) stained 12% DNA-PAGE in Figure 3.1A, all PCR products aligned with the 100bp band of the DNA ladder, supporting the conclusion that the dsDNA templates were generated successfully (dsDNAs: $tRNA^{Sec}$ and $tRNA^{Sec2}$ =112 bp, $tRNA^{UtuX}$ = 108 bp). Subsequently, the tRNA dsDNA-templates were *in vitro* transcribed by T7 RNA polymerase, and the presence of tRNAs was verified by 15% Urea-PAGE. As shown in Figure 3.1B, there was a band present after 4 hours (T240), which was consistently preserved after Dpn1 digestion. Moreover, both transcripts of $tRNA^{Sec}$ and $tRNA^{Sec2}$ were aligned while the $tRNA^{UtuX}$ transcript was slightly shifted downwards, which is likely due to a 4 nt difference between the tRNAs ($tRNA^{Sec}$ and $tRNA^{Sec2}$ = 95 nt, $tRNA^{UtuX}$ = 91 nt). All together, the Urea-PAGE supports the conclusion that tRNAs were successfully transcribed.

Following *in vitro* transcription, transcripts were subjected to phenol-chloroform extraction to remove protein contaminants (i.e., RNA polymerase and RNAse inhibitor). Non-radioactive tRNAs were further purified by size exclusion chromatography (SEC) to remove unincorporated nucleotide triphosphates (NTPs) and abortive transcripts. Figure 3.1C shows an overlap of

chromatograms obtained during SEC purification of non-radioactive tRNAs. Two major peaks were observed during the chromatography, one sharp peak around 10-13 mL elution volume and a wider peak eluting between 16-20 mL, which is characteristic of free NTPs. However, it is worth noticing that the first sharp peaks were accompanied by smaller shoulder peaks, being more prominent in $tRNA^{UtuX}$. The early peak fractions (8-16 mL) were analysed by 8M 15% Urea-PAGE, and the fractions containing the tRNAs were pooled and precipitated. After purification, the concentration of tRNAs was determined by UV absorbance (Table 3.1). To visualize the new transcripts in comparison to previously generated transcripts with low concentrations, 5 µL of purified tRNAs were analyzed by Urea-PAGE using $tRNA^{Phe}$ (76 nt) as control. Figure 3.1D shows an increase in band intensity of the new preparation while being slightly shifted upward. Thus, the Urea-PAGE indicates that the new tRNAs were successfully transcribed and purified with a higher concentration than previously obtained. However, it is worth mentioning that the purity of the newly transcribed tRNAs is 56%, 61% and 66% for $tRNA^{Sec}$, $tRNA^{Sec2}$ and $tRNA^{UtuX}$ respectively, indicating that some impurities such as aborted transcript are present, which could affect the downstream assays.

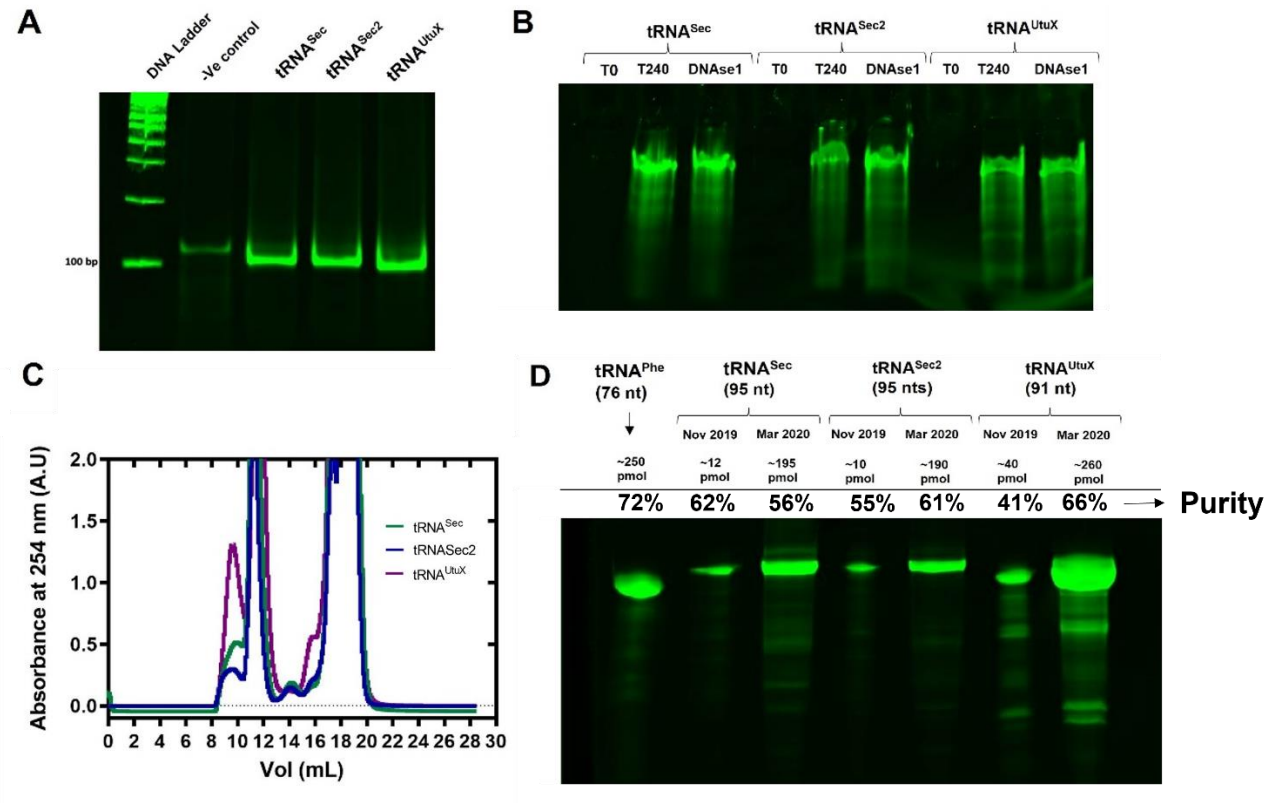


Figure 3. 1 - Preparation of *E. coli* tRNA^{Sec} and variants. A) EtBr-stained 12% DNA-PAGE of primer extension PCR to verify the presence of tRNA dsDNA-templates. Negative (-ve) control: no Q5 DNA polymerase. PCR products approximately align with the 100 bp band, indicating correct products (dsDNA template of tRNA^{Sec} & tRNA^{Sec2}: 112 bp; dsDNA template of tRNA^{UtuX}: 108 bp). B) *In vitro* transcription of tRNAs verified by 15% Urea-PAGE from samples taken at time 0, 240 min, and after DNase treatment. C) Chromatograms measuring the absorbance at 254 nm of fractions collected during the purification of tRNAs by size exclusion chromatography (Superdex 75). D) 8M 15% Urea-PAGE of tRNAs after SEC. To visualize the success of *in vitro* transcription after an initial round of troubleshooting due to low concentrations, 4 µL of each tRNA was loaded on the gel. The size of the tRNAs are as follows: tRNA^{Sec} and tRNA^{Sec2} = 95 nt, tRNA^{UtuX} = 91 nt, tRNA^{Phe} = 76 nt. The purity of each tRNA as determined by the percent of top band of total intensity is stated above the gel: tRNA^{Phe} (72%), 2019 tRNA^{Sec} (62%), 2020 tRNA^{Sec} (56 %), 2019 tRNA^{Sec2} (55%), 2020 tRNA^{Sec2} (61%), 2019 tRNA^{UtuX} (41%) and 2020 tRNA^{UtuX} (66 %)

Table 3. 1 - Concentration of tRNAs

tRNA Preparation	Concentration (μM)	Abs $_{260/280}$
Nov 2019 $tRNA^{Sec}$	2.4 ± 0.3	2.1
Nov 2019 $tRNA^{Sec2}$	2.0 ± 0.2	2.3
Nov 2019 $tRNA^{UtuX}$	8.0 ± 0.5	2.0
Mar 2020 $tRNA^{Sec}$	39.0 ± 1.2	2.4
Mar 2020 $tRNA^{Sec2}$	38 ± 0.4	2.3
Mar 2020 $tRNA^{UtuX}$	52 ± 0.6	2.1
$[^3H]tRNA^{Sec}$	15 ± 1.8	2.2
$[^3H]tRNA^{Sec2}$	43 ± 1.1	2.2
$[^3H]tRNA^{UtuX}$	29 ± 0.2	2.1

Moreover, proteins involved in the selenocysteine translation process and the tRNA modifying enzyme MiaA were purified. First, the N-terminal histidine-tagged SelB, SelD, SerRS and MiaA were overexpressed in *E. coli* cells and subjected to Nickel-sepharose affinity purification to remove most contaminating cellular components. Subsequently, SelD and SerRS were rebuffered and concentrated while SelB and MiaA were further purified via Superdex 75 size exclusion chromatography, and subsequently concentrated. To preserve SelA's decameric structure (510 kDa), a different method of purification was adopted. First, SelA was similarly overexpressed in *E. coli* as verified by SDS-PAGE, which denatures and separates quaternary proteins into monomers that can be visualized on the gel based on their molecular weights. Figure 3.2A shows an increase in band intensity after induction that corresponds to the molecular weight of the SelA monomer (~50 kDa), indicating successful expression. Subsequently, to remove contaminating cellular components under non-denaturing conditions, SelA was subjected to ammonium sulfate precipitation followed by Sephacryl S-300 size exclusion chromatography. Figure 3.2B shows the chromatogram obtained during SEC purification of SelA, indicating primarily a strong and sharp peak around 54 mL elution volume. Selected elution fractions were analyzed by SDS-PAGE (Figure 3.2C), and fraction volumes between 50-62 mL corresponding to SelA were

pooled and concentrated. As judged by the selected fractions in the gel, in comparison to the overexpression (OE) and the ammonium precipitation (AP, before SEC) samples, more than 90% of contaminants were successfully removed. However, it is important to mention that the SelA decametric structure was not validated as the SDS-PAGE analysis is only an indicative of monomer size. Thus, to further validate SelA structure, I could have used a protein of similar size (~ 500 KDa) as reference to compare the chromatograph after size exclusion chromatography. Finally, the concentration of SelA as well as the other proteins was determined photometrically at 280 nm and verified by ImageJ analysis of SDS-PAGE concentration gels through comparison to previously characterized proteins (Table 3.2).

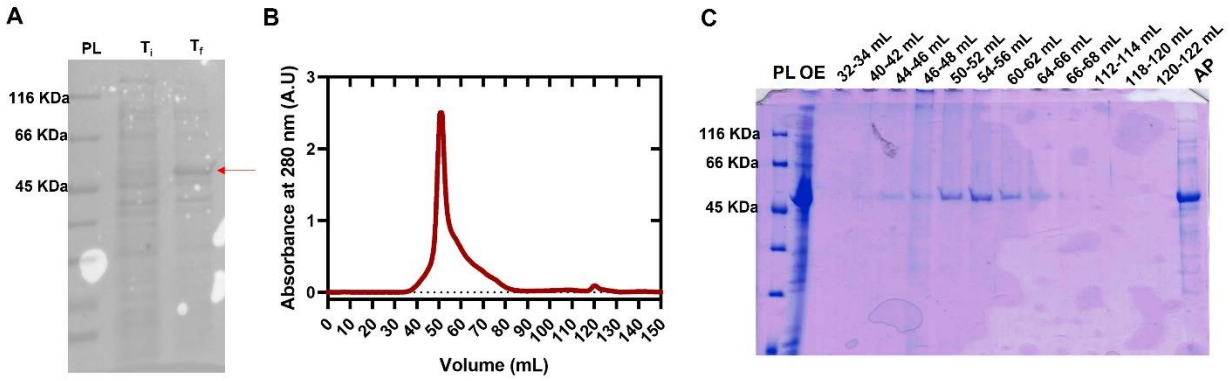


Figure 3. 2 - Representative protein purification: SelA. A) 12% SDS-PAGE of SelA over night expression (~16 hours at 18 °C). PL= protein ladder, T_i = initial time of induced expression; T_f = final time of expression. Red arrow highlights the band corresponding to SelA monomer (~50 KDa). B) Chromatogram generated during SelA Sephacryl-S300 SEC purification. C) 12% SDS-PAGE analysis of selected eluted fraction. OE = over expression sample, AP = sample after ammonium precipitation but before SEC.

Table 3. 2 - Final protein concentrations.

Protein	Concentration (μM)
SelA	112 ± 1
SelB	120 ± 1
SelD	108 ± 2
SerRS	36 ± 1
MiaA	180 ± 3

3.2 Tritium release assays detected Ψ55 and m^5U54 formation on $tRNA^{Sec}$ and variants

Previous work has demonstrated that the *E. coli* Sec biosynthesis and insertion machinery is quite flexible and can be adjusted to incorporate Sec in the absence of the SECIS element mediated either by EF-Tu or SelB in response to a UAG stop codon [40, 41, 43]. Nevertheless, despite the synthetic approaches opening new doors for the incorporation of Sec at any position within the polypeptide chain, further improvements to the mechanism are required to optimize selenoprotein yields. One of the areas in the field that has been widely unexplored is the role of tRNA post-transcriptional modifications on the overall selenoprotein synthesis. Hereby, I wanted to first determine whether the synthetic $tRNA^{UtuX}$ and the $tRNA^{Sec2}$ variant, which play crucial roles on the engineered mechanisms and contain the potential for the same nucleotide modifications (m^5U54 and Ψ55) as wildtype $tRNA^{Sec}$, can be modified by TrmA and TruB similarly to $tRNA^{Sec}$. Along these lines, I performed tritium release assays by incubating 0.6 μM of [^{3}H]-tRNAs with 5 μM of TruB or TrmA. As TruB or TrmA acts on the corresponding tritium-labelled uridines, tritium is released as biproduct of the enzymatic reactions (Figure 3.3). After taking multiple time-points and separating the tRNAs via charcoal absorption, the tritium content of the supernatant is determined via scintillation counting, corresponding to the level of

methylation or speudouridylation on the tRNAs. Figure 3.4 shows that the level of Ψ55 formation in $tRNA^{Sec}$ and variants by TruB is approximately 100% when considering the standard deviation of the datapoints. Similarly, as shown in Figure 3.5, the level of m^5U54 formation in all targeted tRNAs by TrmA is approximately 100% after 5 min of incubation. However, it is worth mentioning that the seemingly decreased level of m^5U54 formation in $tRNA^{UtuX}$ after 30 min is most likely an artifact of the assay, and likely caused by pipetting errors given these methylation events are not reversable. Nevertheless, taken all together, the results support the conclusion that $tRNA^{Sec}$, $tRNA^{Sec2}$ and $tRNA^{UtuX}$ are all similarly modified by TrmA and TruB, and the potential role of tRNA post-transcriptional modification in selenoprotein biosynthesis needs further attention.

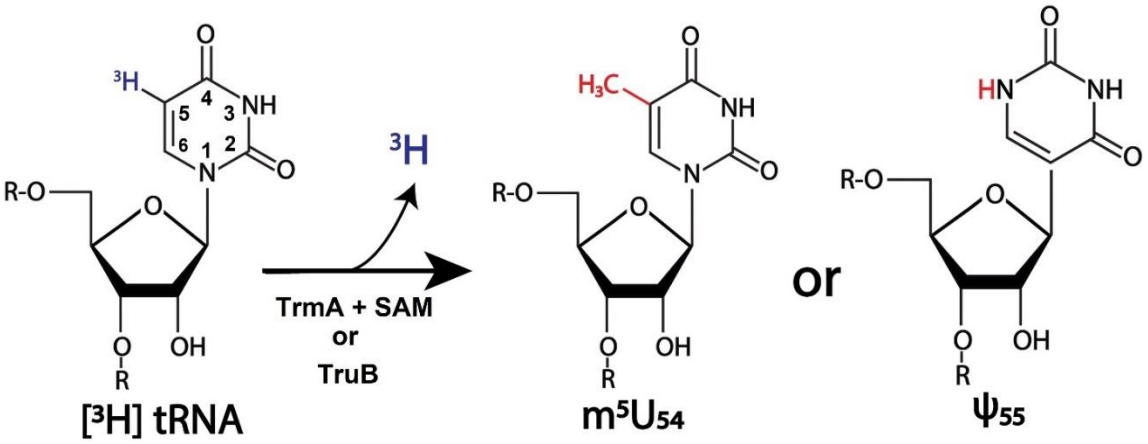


Figure 3. 3 - Principle of tritium release assay. Tritium [^{3}H]-tRNAs are targeted by TrmA or TruB to form the 5-methyl uridine ($m5U_{54}$) and speudouridine (Ψ_{55}) modificaitons at positions 54 and 55 respectively, realsing tritrium as a bi-product of the enzymatic reactions. Samples of the enzymatic reactions are taken at different time-points and the tritum release corresponding to the level of modificaiton is measured via scintilation counting. The added methyl group in the $m5U_{54}$ modificaiton and the new N-H pair in Ψ_{55} are hilighted in read.

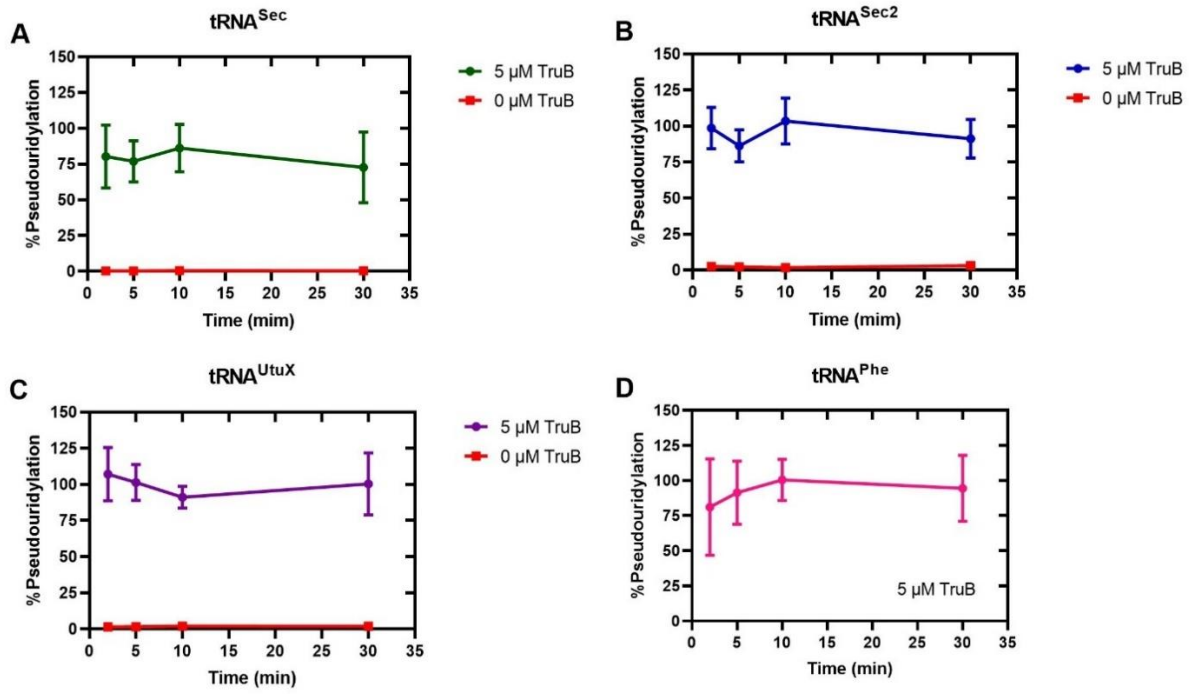


Figure 3. 4 - Pseudouridylation activity of TruB on *in vitro* transcribed tRNA^{Sec} and variants. A) Tritium release assay of tRNA^{Sec}. B) Tritium release assay of tRNA^{Sec2}. C) Tritium release assay of tRNA^{UtuX}. D) Tritium release assay of the positive control tRNA^{Phe}. All test reactions were conducted in triplicates using 5 µM of TruB and 0.6 µM of [^{3}H]-tRNAs. Negative control reactions were performed in the absence of TruB.

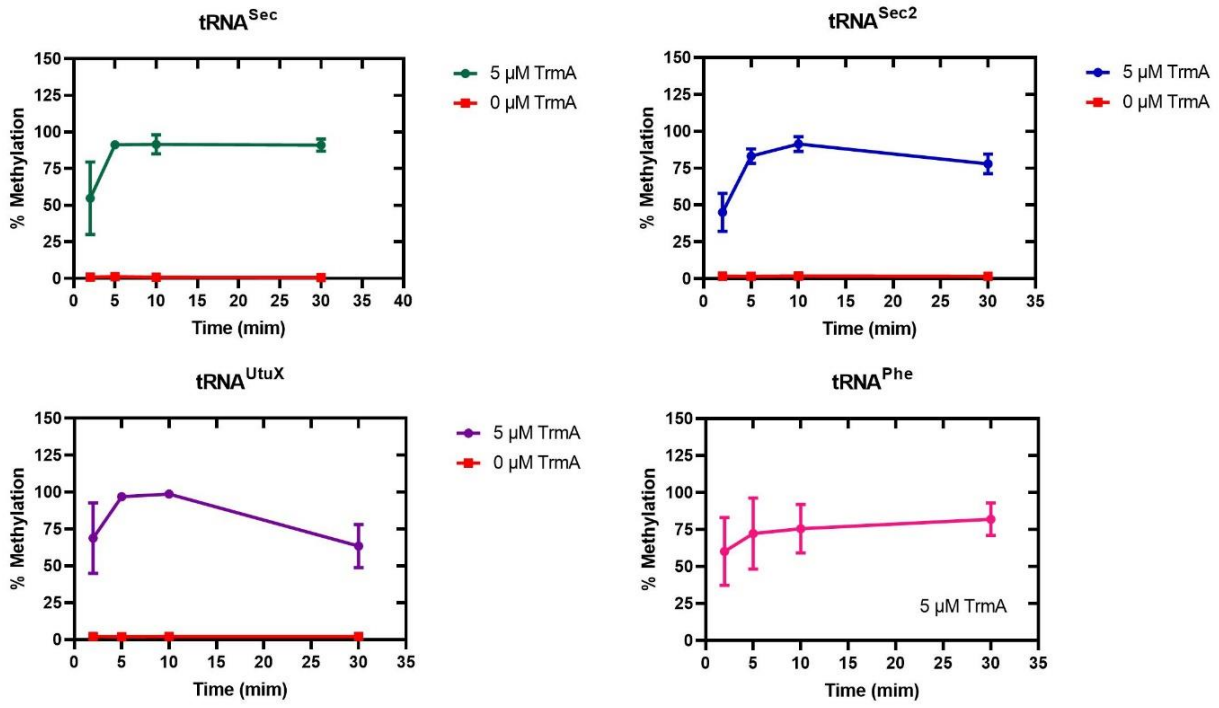


Figure 3. 5 - Methylation activity of TrmA on tRNA^{Sec} and variants. A) Tritium release assay of tRNA^{Sec}. B) Tritium release assay of tRNA^{Sec2}. C) Tritium release assay of tRNA^{UtuX}. D) Tritium release assay of with the positive control tRNA^{Phe}. All test reactions were conducted in triplicates using 5 µM of TrmA and 0.6 µM of [^{3}H]-tRNAs. Negative control reactions were performed in the absence of TrmA.

3.3 Experimental evaluation of the interaction of tRNA^{Sec}, tRNA^{UtuX} and tRNA^{Sec2} with the modifying machinery

To evaluate the interaction of tRNA^{Sec}, tRNA^{Sec2} and tRNA^{UtuX} with the *E. coli* modifying machinery, a series of filter binding assays was carried out using 20 nM of [^{3}H]-tRNAs and increasing concentrations of MiaA, TruB and TrmA (0, 0.08, 0.5, 1, 5 and 10 µM). Each reaction was incubated for 10 minutes and exposed to nitrocellulose membrane filtration, disregarding any unbound tRNA that passes through the filter while those that bind to the proteins are

permeabilized within the filter due to the hydrophobic interactions between the protein side chains and the nitrocellulose membrane (Figure 3.6).

Notably, all the reactions were carried out with either catalytic inactive mutants or in the absence of corresponding cofactor, ensuring the measurement of binding affinity for unmodified tRNA. The percent of bound tRNA was then determined by scintillation counting, and the dissociation constant (K_D) was determined by plotting the fraction of bound tRNA (*Bound*) as a function of enzyme concentration ([*enzyme*]) and fitting to a hyperbolic equation as described [27, 28].

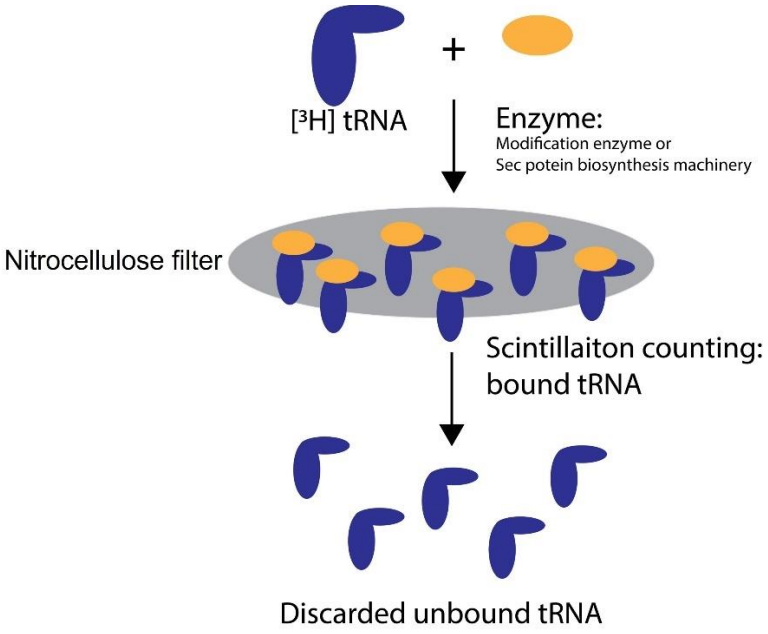


Figure 3. 6 - Nitrocellulose filter binding assay. Tritium-labelled tRNA is incubated with increasing protein concentrations. The reactions are incubated for 10 minutes and consequently exposed to nitrocellulose filtration in vacuum, capturing only protein-bound [^{3}H]-tRNAs while the unbound [^{3}H]-tRNAs go through the filter and are disregarded. The filter is subsequently exposed to scintillation counting and the ratio of unbound versus unbound tRNA is determined, leading to the determination of the dissociation constant (K_D).

The data indicates that TruB has no preferential binding affinity towards the three different tRNA substrates since no significant difference in the K_D values was observed (Figure 3.7 A,

Table 3.3). Similarly, MiaA has no significant substrate specificity as judged by the K_D values (Figure 3.5 C, Table 3.3). However, TrmA seems to have slightly stronger binding affinity towards $tRNA^{Sec}$ compared to either $tRNA^{Sec2}$ or $tRNA^{UtuX}$ (Figure 3.7 B, Table 3.3). When analysing the results with a broader perspective, two major observations could be made. Firstly, the results indicates that there is approximately 50-60% tRNAs bound at high protein concentrations. Secondly, there is an apparent general trend in the K_D values across all three modifying enzymes. The results indicates that TruB binds to $tRNA^{Sec}$, $tRNA^{Sec2}$ and $tRNA^{UtuX}$ with K_D of 3.7 µM, 2.5 µM, and 3.2 µM respectively (Table 3.3). On the other hand, TrmA binds to RNA^{Sec}, $tRNA^{Sec2}$ and $tRNA^{UtuX}$ with K_D of 0.3 µM, 0.8 µM and 1.4 µM respectively, while MiaA showed corresponding K_D values of 0.4 µM, 0.5 µM, and 0.6 µM (Table 3.3). Thus, overall, the data indicates that TruB has the lowest binding affinity towards the substrate tRNAs when compared to TrmA and MiaA based on the K_D values. Moreover, except for the slight difference in K_D for the $tRNA^{UtuX}$ substrate, both TrmA and MiaA bind to the other substrates with similar affinity (Table 3.3).

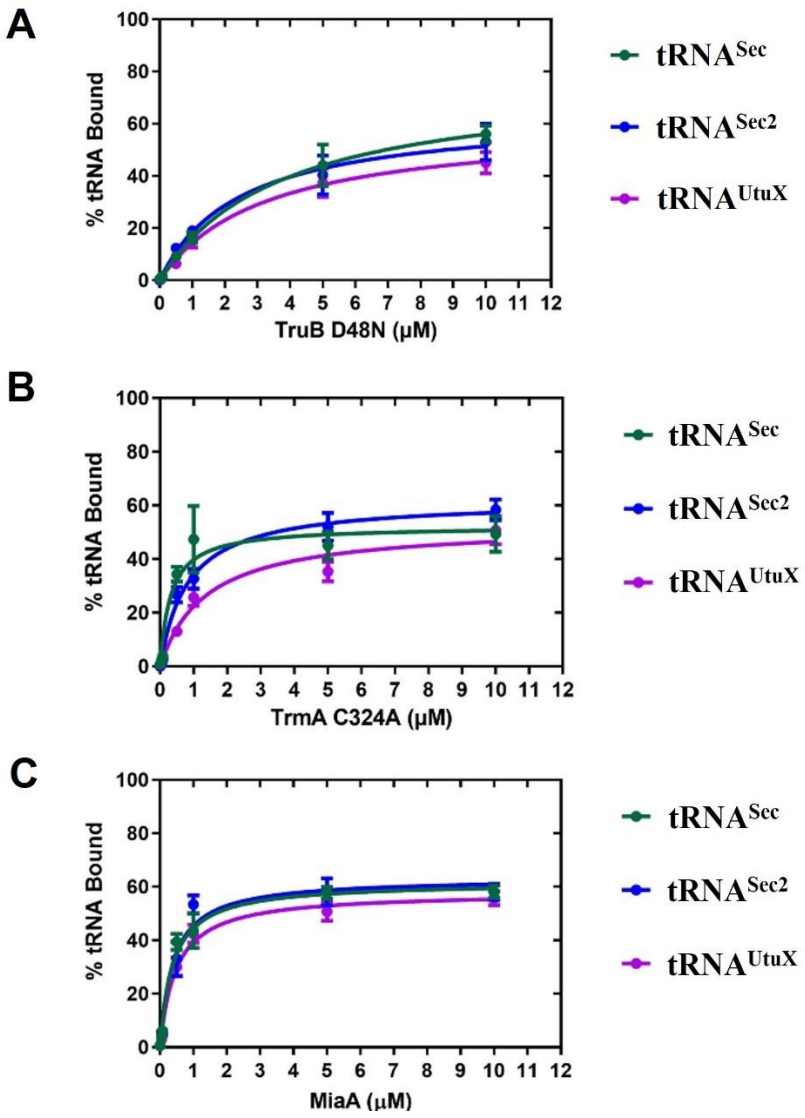


Figure 3. 7 - Nitrocellulose assay of tRNAs binding to the modifying machinery. (A-C) Filter binding of TruB, TrmA and MiaA with $tRNA^{Sec}$, $tRNA^{Sec2}$ and $tRNA^{UtuX}$. All reactions were conducted using 20 nM of tRNA and 0, 0.08, 0.5, 1, 5 and 10 μM of modifying enzyme.

Table 3. 3 - Dissociation constant (K_D) between modifying enzymes and substrate tRNAs.

Protein	Substrate	Dissociation Constant (μM)
TruB D48N	$tRNA^{Sec}$	3.7 ± 1.0
	$tRNA^{Sec2}$	2.5 ± 0.8
	$tRNA^{UtuX}$	3.2 ± 0.9
TrmA C324A	$tRNA^{Sec}$	0.3 ± 0.1
	$tRNA^{Sec2}$	0.8 ± 0.2
	$tRNA^{UtuX}$	1.4 ± 0.6
MiaA	$tRNA^{Sec}$	0.4 ± 0.1
	$tRNA^{Sec2}$	0.5 ± 0.1
	$tRNA^{UtuX}$	0.6 ± 0.1

3.4 Validation of the purified Sec biosynthesis and incorporation machinery

Sec biosynthesis and incorporation mechanism is a complex multi-step process, where the product formation in each step is the prerequisite for the subsequent reactions. One approach to indirectly study a cascade of enzymatic activity like this is to determine the tRNA• protein interaction at each step and compare the binding affinity and percent of bound tRNA to an appropriate control. Along these lines, I incubated $tRNA^{Sec}$ with seryl-tRNA synthetase, ATP and serine to prepare Ser-[^{3}H]$tRNA^{Sec}$ and additionally SelA, SelD, PLP, Na_2SeO_3, SelB, ATP, GTP and serine to generate Sec-[^{3}H]$tRNA^{Sec}$ as previously described [21] and determined the overall tRNA concentration via scintillation counting. Next, I conducted a series of filter binding

experiments by incubating 20 nM tRNA with increasing concentrations of SerRS, SelA or SelB to verify that they bind $tRNA^{Sec}$ as expected. The percent of bound tRNA was determined by scintillation counting, and the K_D was determined by plotting the fraction of bound tRNA (*Bound*) as a function of enzyme concentration ([*enzyme*]) and fitting to a hyperbolic equation (see Materials and Methods).

The first step in the selenoprotein biosynthesis mechanism is the serylation of $tRNA^{Sec}$ by SerRS. To indirectly validate if any of the enzymes involved in the downstream reactions are active, Ser-$tRNA^{Sec}$ and Sec-$tRNA^{Sec}$ needed to be prepared and validated. Given that serylation event could be the bottle neck of the enzymatic cascade, filter binding assays were carried out between unmodified WT $tRNA^{Sec}$ and SerRS to infer potential activity of SerRS. The results showed that SerRS binds to $tRNA^{Sec}$ with a K_D of 1.0 ± 0.4 µM (Figure 3.8 A, Table 3.4). Thus, while this data alone does not indicate activity, it showed that SerRS could potentially be active by interacting with the substrate tRNA.

The second step in the enzymatic cascade is the Ser to Sec conversion by SelA. To evaluate SelA potential activity, filter binding assay were carried out using previously prepared Ser-$tRNA^{Sec}$ or deacylated $tRNA^{Sec}$ as substrates in the absence of the selenomonophosphate cofactor, measuring substrate binding rather than product release. The results indicate that SelA binds to deacylated tRNA with a K_D of 5.7 ± 2.2 µM and to Ser-$tRNA^{Sec}$ with a K_D of 3.6 ± 0.7 µM (Figure 3.8 B, Table 3.4). Therefore, no statistical significance was observed for the binding affinity of SelA towards either deacylated $tRNA^{Sec}$ or Ser-$tRNA^{Sec}$ based on the standard deviation of the K_D values, but my data indicate that SelA is in general active with respect to binding tRNA.

The subsequent step is the recognition of Sec-$tRNA^{Sec}$ by SelB, which preferentially binds to Sec-$tRNA^{Sec}$ over Ser-$tRNA^{Sec}$ or deacylated $tRNA^{Sec}$ [21]. To further evaluate the purified Sec

biosynthesis and insertional machinery, filter binding assay were conducted to determine the experimental binding affinity of SelB toward Sec-tRNA^{Sec} and deacylated tRNA^{Sec}. The data indicate that the purified SelB binds to deacylated tRNA^{Sec} with a K_D of 1.0 ± 0.2 μM and to Sec-tRNA^{Sec} with a K_D of 0.5 ± 0.1 μM (Figure 3.8 C, Table 3.4). Therefore, taken all together, the results suggested that SerRS, SelA and SelD, which was used to prepare Sec-tRNA^{Sec}, were active, and that SelB could differentiate between aminoacylated and deacylated tRNA. Thus, all components of the selenocysteine incorporation machinery seemed to be at least partially active based on this preliminary tRNA binding analysis.

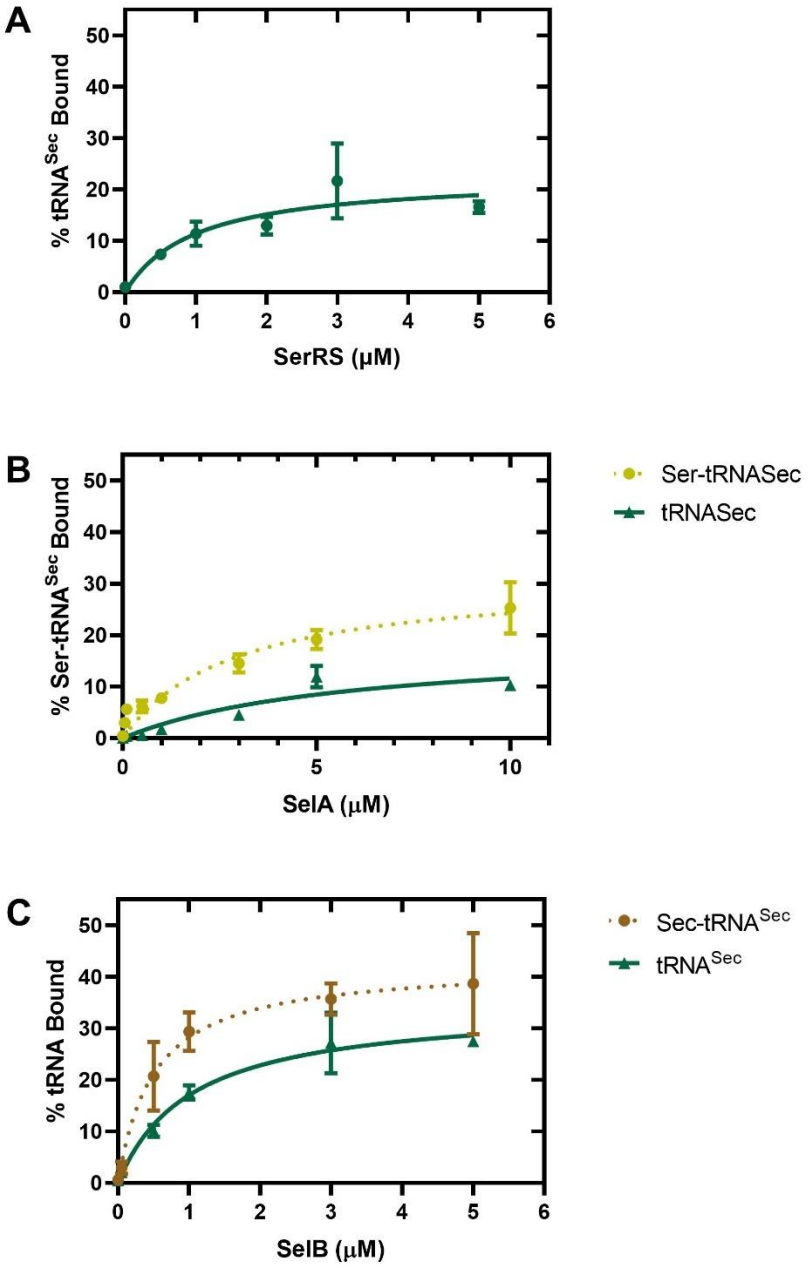


Figure 3. 8 - Nitrocellulose assays of tRNA binding by the selenocysteine biosynthesis and insertion machinery. A) Filter binding assay of $tRNA^{Sec}$ with 0, 0.05, 0.5, 1, 2, 3, 5 μM of SerRS. B) Filter binding assay of $tRNA^{Sec}$ and Ser-$tRNA^{Sec}$ with 0, 0.05, 0.1, 0.5, 3, 5, 10, and 20 μM of SelA. C) Filter binding assay of $tRNA^{Sec}$ and Ser-$tRNA^{Sec}$ with 0, 0.05, 0.5, 1, 2, 3, 5 μM of SelB. All reactions were performed in triplicates with 20 nM of tRNAs. Hyperbolic fitting yielded the K_Ds summarized in the insets.

Table 3. 4 - Dissociation constant (K_D) of the Sec biosynthesis and incorporation machinery.

Protein	Substrate	Dissociation Constant (µM)
SerRS	$tRNA^{Sec}$	1.0 ± 0.4
SelA	$tRNA^{Sec}$	5.7 ± 2.2
	Ser- $tRNA^{Sec}$	3.6 ± 0.7
SelB	$tRNA^{Sec}$	1.0 ± 0.2
	Sec-$tRNA^{Sec}$	0.5 ± 0.1
SelB[a]	$tRNA^{Sec}$	0.5 µM[a]
	Sec-$tRNA^{Sec}$	0.2 pM[a]

[a] *Data from Paleskava et al., 2010*

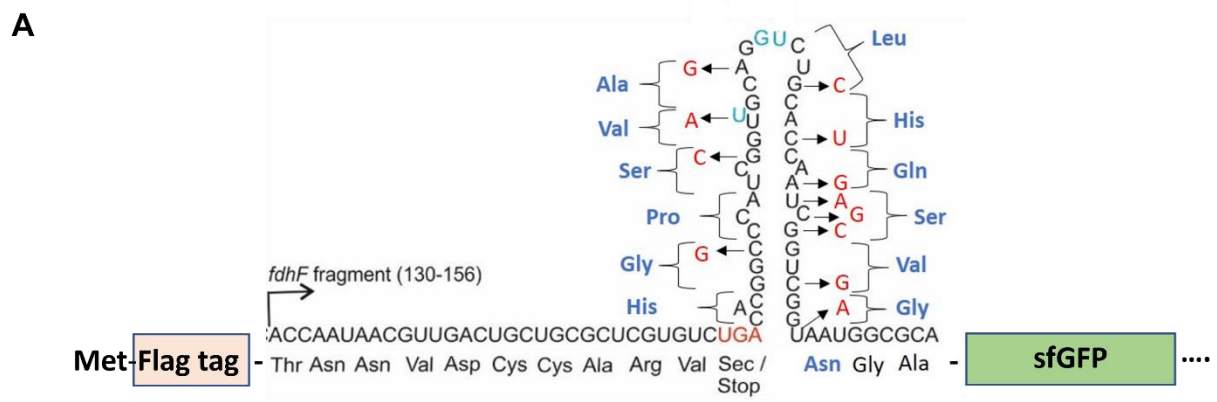
3.5 Role of tRNA modifications on Sec incorporation *in vitro*

To determine the effect of tRNA post-transcriptional modifications on selenoprotein biosynthesis *in vitro*, a series of coupled *in vitro* transcription and translation (TXTL) reactions were carried out in the presence or absence of tRNA modifying enzymes using the PURExpres® Δ RF123 cell-free system and a plasmid-encoded super-fold green fluorescence protein (sfGFP) reporter. The reporter system contains a fragment of *E. coli* selenoprotein *fdhF* with the designated stop codon at position 140 (*fdhF*-UGA_{140} or *fdhF*-UAG_{140}) in the presence or absence of the SECIS element followed downstream by the sfGFP (Figure 3.9). To measure Sec incorporation in the absence the SECIS element without altering the amino acid sequence of the translated product, specific nucleotide changes were introduced to the mRNA sequence of the SECIS element that

likely prevent the formation of the SECIS element (Figure 9). This experimental system allows for the determination of stop codon recoding and Sec incorporation using the natural SelB-mediated SECIS-dependent UGA suppression (UGA + SECIS ($tRNA^{Sec}$)), the engineered EF-Tu-mediated SECIS-independent UAG suppression (UAG – SECIS ($tRNA^{UtuX}$)), and the SelB-mediated SECIS-independent UAG suppression (UAG – SECIS ($tRNA^{Sec2}$)) mechanisms. These reporter systems were designed to evaluate stop codon reassignment as judged by the potential fluorescence signal determined after the PURE reactions (Figure 3.9). If the translation machinery fails to read through the stop codon, a truncated product with no fluorescence signal would be observed. However, if the stop codon is properly reassigned with Sec, but also potentially with the Ser residue, a fluoresce signal emitted by the reporter sfGFP protein would be detectable.

The PURE reactions of the SelB-mediated SECIS-dependent UGA suppression were supplemented with 0.5 µL of RF1 and RF3, 15 µM of SelB, 50 µM of PLP, 50 µM of L-Serine, 50 µM Selenite and 5 µM of SelA, SelD, $tRNA^{Sec}$ and 250 ng of plasmid DNA containing the SECIS mRNA sequence. As a set of controls, reactions were conducted in the absence of SelB, plasmid DNA or $tRNA^{Sec}$, stablishing a background signal or any misincorporation of other amino acids at the UGA stop codon while evaluating the potential presence of $tRNA^{Sec}$ within the commercial kit. PURE reactions for the EF-Tu mediated SECIS-independent UAG suppression approach were supplemented with 15 µM of SelB, 50 µM of PLP, 50 µM of L-Serine, 50 µM of Selenite, 5 µM SelA and SelD, 0.5 µL of RF2 and RF3, 5 µM of $tRNA^{UtuX}$ and 250 ng of plasmid DNA whose mRNA sequence is unlikely to form a tertiary SECEIS element (Figure 3.9). As a control, a reaction lacking $tRNA^{Sec2}$ was conducted, providing better assessment of other amino acid misincorporations, which could be suggested if significant

fluorescence signal would be observed. Similarly, PURE reactions for UAG suppression by $tRNA^{Sec2}$ mediated by SelB were supplemented with 15 µM of SelB, 50 µM of PLP, 50 µM of L-Serine, 50 µM of Selenite, 5 µM SelA and SelD, 0.5 µL of RF2 and RF3, 5 µM of $tRNA^{Sec2}$ and 250 ng of SECIS-deprived plasmid DNA. Lastly, to evaluate the quality of the reporter system, the potential effect of the SECIS element during Sec biosynthesis, and to establish a sense codon translation benchmark, positive control PURE reactions having a Phe_{140} sense codon were carried out in the presence or absence of the SECIS element.



Translation	Sec/Stop	His	Gly	Pro	Ser	Val	Ala	Gly	Leu	His	Gln	Ser	Val	Gly	Asn
UGA + SECIS ($tRNA^{Sec}$ & SelB)	UGA	CAC	GGC	CCA	UCG	GUU	GCA	GGU	CUG	CAC	CAA	UCG	GUC	GGU	AAU
UUC + SECIS ($tRNA^{Phe}$ & EF-Tu)	UUC	CAC	GGG	CCA	UCG	GUU	GCA	GGU	CUG	CAC	CAA	UCG	GUC	GGU	AAU
UUC - SECIS ($tRNA^{Phe}$ & EF-Tu)	UUC	CAC	GGG	CCA	UCC	GUA	GCG	GGU	CUC	CAU	CAG	AGC	GUG	GGA	AAU
UAG - SECIS ($tRNA^{Sec2}$ & SelB; $tRNA^{UtuX}$ & EF-Tu)	UAG	CAC	GGG	CCA	UCC	GUA	GCG	GGU	CUC	CAU	CAG	AGC	GUG	GGA	AAU

Figure 3. 9 - *E. coli* reporter system design. A) Sequence of reporter system, containing a fragment of *E. coli* selenoprotein fdhF, stop codon UGA for selenocysteine incorporation and the natural SECIS element immediately downstream. Three letter code of amino acids coded in the SECIS element is shown in blue and selected nucleotide mutations to disrupt the 3-D structure of the SECIS element without altering the coding sequence in red. B) Summary of nucleotide changes to the SECIS element to implement the SECIS independent mechanisms.

Ultimately, all PURE reactions were also carried out in the presence of modifying enzymes. To evaluate the effect of tRNA modifications during translation, particularly during selenoprotein biosynthesis, PURE reactions were additionally supplemented with 2 µM of modifying enzymes

MiaA, TruB and TrmA along with 50 µM of cofactors, DMAP and SAM. By adding the cofactors and purified enzymes, which had been previously demonstrated to be active or at least partially active, to the TXTL reaction, I can monitor any change in fluorescence signal that could potentially be correlated to the role of tRNA modification during translation.

Fluorescence spectra of the positive control PURE reactions having a Phe_{140} sense codon indicates that the incorporation of modifying enzymes to the reactions had hardly any influence in the fluorescence signal in reactions carried out in the absence of the SECIS element (Figure 3.10 D). Thus, the results suggest that the $tRNA^{Phe}$ within the commercial kit is already modified and that the modifying enzymes do not further enhance translation. Moreover, the fluorescence signal decreases approximately by 11% in the presence of the SECIS element (Figure 3.8D), suggesting only a small negative influence of the SECIS element during translation of sense codons.

However, the fluorescence spectra corresponding to the SelB-mediated SECIS-dependent UGA recoding mechanism show an increase in fluorescence signal in reactions supplemented with modifying enzymes compared to those reaction deprived of the modifications (Figure 3.10 A), representing and increment of approximately 30% in sfGFP peak fluorescence signal (Figure 3.11A) Thus, the results suggest a positive role of tRNA modifications during stop codon reassignment by Sec. In addition, the signal observed in reactions supplemented with MiaA only exhibited similar fluorescence spectra to those supplemented with MiaA, TruB and TrmaA together. Interestingly, the control reaction carried out in the absence of additional $tRNA^{Sec}$ generated a detectable fluorescence signal (Figure 3.10 A), indicating that $tRNA^{Sec}$ is likely present within the commercial PURE system and that the addition of $tRNA^{Sec}$ to the reaction further increases translation of the reporter protein. Moreover, the negative control reaction

carried out in the absence of SelB, resulted in a substantial reduction of the fluorescence signal compared to those supplemented with SelB (Figure 3.10 A), which suggests that SelB is not included in the PURE system and that the supplemented SelB is active. Furthermore, the data also suggest that misincorporation of amino acids other than selenocysteine mediated by EF-Tu is likely insignificant.

Additionally, fluorescence spectra of the SECIS-independent PURE reactions exhibited similar patter in terms of the role of tRNA modification during translation (Figure 3.10 B-C). Both approaches, SelB-mediated and EF-Tu mediated, resulted in higher sfGFP signal in the presence of modifying enzymes. In the case of the SelB-reactions, the increase of peak fluorescence in the presence of tRNA modifying enzymes was approximately 30% higher than reactions carried out in the absence modifications (Figure 3.11B). Moreover, a control reaction in the absence of the repressor $tRNA^{Sec2}$ resulted in a substantial reduction of sfGFP translation as judged by the fluorescence signal, suggesting that amino acids misincorporation is likely insignificant. On the other hand, the additional incorporation of the tRNA modifying enzymes in EF-Tu-mediated reactions resulted in approximately 45% increase in sfGFP peak fluorescence relative to the counterparts lacking modifications (Figure 3.11 C). Thus, the results suggest that tRNA modifying enzymes positively benefit UGA stop codon suppression, possibly by introducing modifications in both $tRNA^{Sec2}$ and $tRNA^{UtuX}$ that appear to provide some advantage during translation. Lastly, when comparing selenocysteine incorporation encoded by a stop codon with translation of sense codon at the corresponding position in the presence of the SECIS element, there was an approximate 30% peak fluorescence decrease for selenocysteine incorporation (Figure 3.11D), indicating that the selenoprotein biosynthesis process could be inherently less efficient than sense codon translation due to the nature of recoding a stop codon.

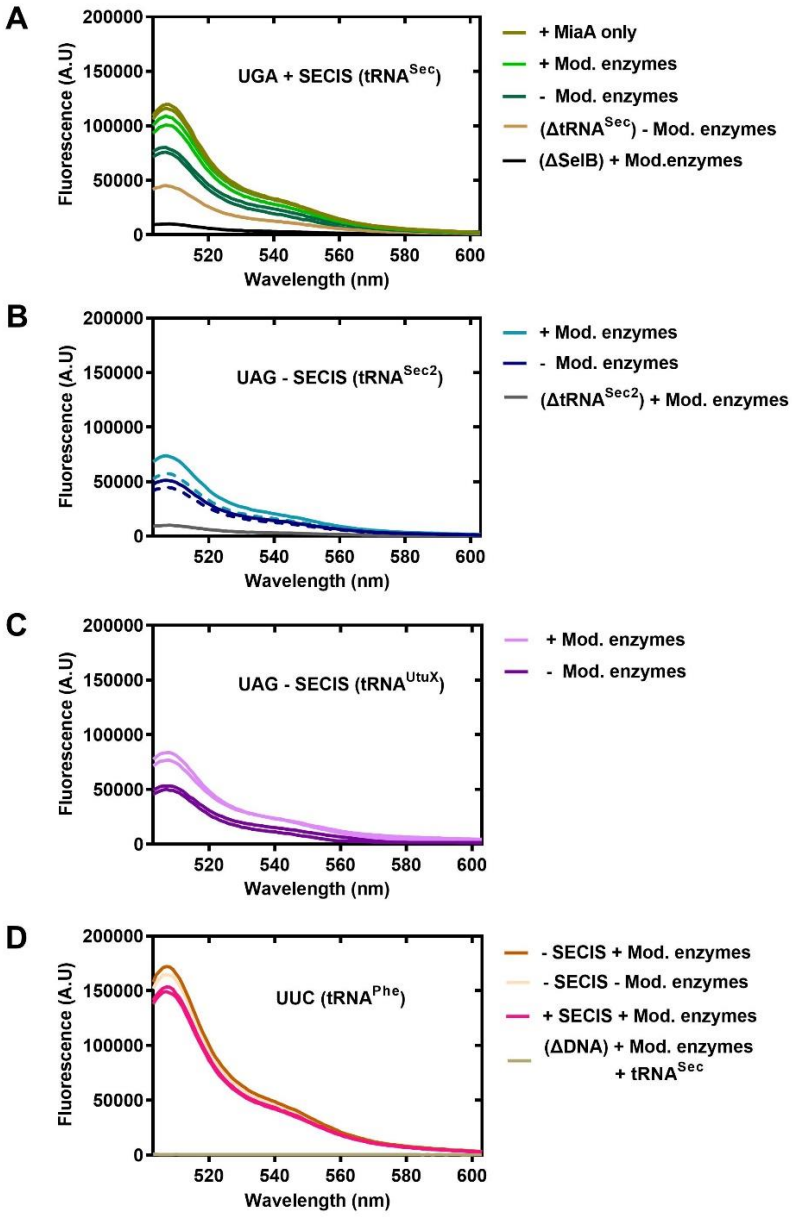


Figure 3. 10 - Supplementing *in vitro* coupled transcription-translation reactions with tRNA modifying enzymes improves Sec incorporation. A) SelB-mediated SECIS-dependent translation of GFP harboring a selenocysteine-encoding stop codon in the N-terminal region. B) SelB-mediated SECIS-independent translation of GFP. Line pattern indicates the date in which the reactions were carried out: dash line reactions were done simultaneously while straight line reactions were done in parallel on a different date. C) EF-Tu-mediated SECIS-independent GFP translation. D) Positive control reactions using DNA templates encoding a UUC sense codon instead of a selenocysteine-encoding stop codon and a negative control lacking DNA template (DDNA). The synthesis of GFP was performed using the commercially available PURE system supplemented with additional factors as summarized in Tables 2.2, 2.3 and the presence of GFP was observed by measuring the fluorescence emission spectra upon excitation at 488 nm and emission scanned between 503-603 nm (λ_{max} = 508 nm; slid width 8 nm).

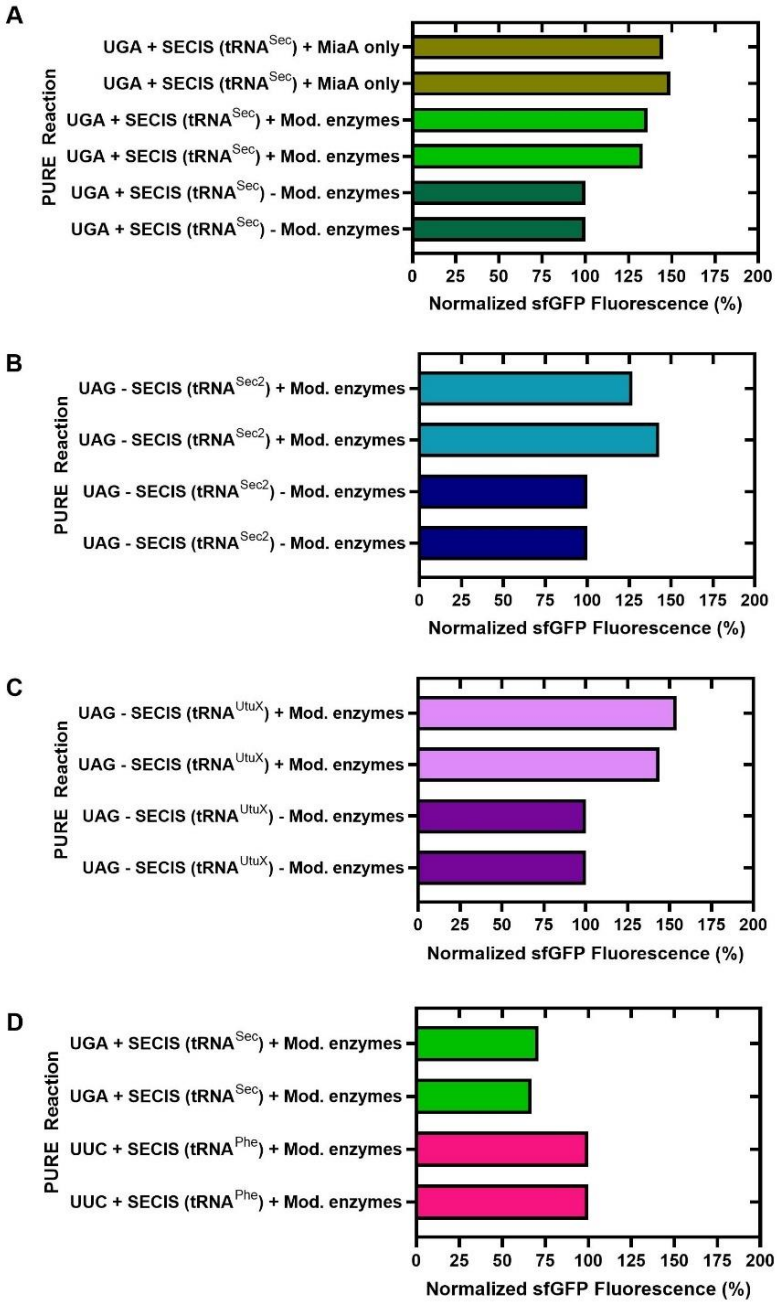


Figure 3. 11 - Normalized sfGFP fluorescence comparing translation efficiency in presence and absence of tRNA modification enzymes. A) SelB-mediated SECIS-dependent sfGFP synthesis. The reaction in absence of modification enzymes is normalized to 100% to compare with reactions in the presence of tRNA modifying enzymes. B) SelB-mediated SECIS-dependent sfGFP translation. C) EF-Tu-mediated SECIS-independent sfGFP synthesis . D) Comparison of stop codon vs. sense codon sfGFP translation where the reactions with the UGA stop codons are normalized against translation in the presence of sense codons. SfGFP was excitation at 488 nm and emission scanned between 503-603 nm (λ_{max} = 508 nm; slid width 8 nm).

Chapter 4: DISCUSSION

4.1 Evaluating $tRNA^{Sec}$ post-transcription modifications

Developing an efficient platform for selenoprotein bioengineering and *in vitro* biosynthesis is of great interest due to the potential benefits selenoproteins provide to the biotechnology industry (see introduction section for details). In recent years, great progress has been made in developing SECIS-independent *E. coli* mechanisms that allow site-specific selenocysteine incorporation into proteins. In general, the flexibility of the SECIS-independent mechanisms depends on $tRNA^{Sec}$ variants that have been optimized to serve as substrates for either EF-Tu or SelB to recode an amber UAG stop codon into selenocysteine. Two of those $tRNA^{Sec}$ variants are $tRNA^{UtuX}$ and $tRNA^{Sec2}$ [40, 43]. However, neither $tRNA^{Sec}$ post-transcriptional modifications nor the potential interaction of $tRNA^{UtuX}$ and $tRNA^{Sec2}$ with modifying enzymes have been evaluated although they may enhance selenoprotein synthesis.

To better understand $tRNA^{Sec}$ modifications, I conducted a series of tritium release assays to monitor the formation of m^5U_{54} and Ψ_{55}, corresponding to the enzymatic activity of TrmA and TruB on $tRNA^{Sec}$, $tRNA^{UtuX}$ and $tRNA^{Sec2}$ (Figure 3.4, 3.5). Both, TrmA and TruB, similarly modify $tRNA^{Sec}$ and $tRNA^{Sec2}$, reaching approximately 100 percent modification within 10 minutes of reaction time under the single-turnover conditions used in our assay. Considering that $tRNA^{Sec}$ and $tRNA^{Sec2}$ share identical sequences except for the anticodon that was altered to read UAG stop codon (Figure 1.4), similar modification levels were expected. However, it was

unknown whether the synthetic $tRNA^{UtuX}$, could be modified by TrmA and TruB. The tritium release assay data showes that $tRNA^{UtuX}$ is also fully modified by TruB and TrmA similarly to $tRNA^{Sec}$ and $tRNA^{Sec2}$. While no modification profile of $tRNA^{UtuX}$ has been previously determined, the results were expected as $tRNA^{UtuX}$ preserves the target nuclotides (U_{54} and U_{55}) and as importantly, the identity of the interacting elements within T-arm present in both parent tRNAs, $tRNA^{Sec}$ and $tRNA^{Ser}$. Thus, my results demonstrate that all three tRNAs can be modified by TrmA and TruB, indicating that the tRNAs were successfully prepared and that they can be utilized to test the effect of post-transcription modification on selenoprotein biosynthesis *in vitro*. Moreover, the fact that $tRNA^{UtuX}$ is modified by TrmA and TruB *in vitro* raises the question of which endogenous modifying enzymes could potentially act on other engineered tRNAs and what the effect of the modification would be during translation, which is generally ignored when implementing synthetic biology experiments.

One approach to study the interaction between modifying enzymes and tRNAs is by determining the dissociation constant (K_D). This approach has previously been implemented to determine the effect of pre-existing tRNA modifications on the tRNA affinity of other tRNA modifying enzymes, supporting the hypothesis of a preferred order of elbow modifications of *E. coli* $tRNA^{Phe}$ [50]. Here, I conducted a series of single turnover filter binding assays to determine the binding affinity of TrmA, TruB and MiaA to the target tRNAs as a first step towards verifying whether these enzymes can modify $tRNA^{Sec}$ *in vitro*. This assay was successfully implemented to determine the K_D of each enzyme-tRNA interaction, resulting in no statistical significance between any of the three tRNA substrates for each modifying enzyme (Figure 3.7, Table 3.3, e.g., TruB vs. $tRNA^{Sec}$, $tRNA^{Sec2}$, $tRNA^{UtuX}$). Thus, in general the results suggest that the modifying enzymes bind to the target tRNAs, which particularly makes sense for TrmA and

TruB since m^5U_{54} and Ψ_{55} are highly conserved modifications found in all elongator tRNAs [24, 25]. Moreover, both TrmA and TruB have been shown to act as tRNA chaperones independently of their modifying activity [27, 28], suggesting that these enzymes likely bind to the substrate tRNAs non-specifically.

In contrast, the i^6A_{37} modification by MiaA is only introduced into a handful of tRNAs (Sec, Cys, Leu, Phe, Ser, Trp, Tyr, Leu) that contain the A36-A37-A38 motif [34]. An early kinetic study showes that MiaA has strong modification activity towards intact $tRNA^{Phe}$ [54]. Moreover, pre-existing modifications and non-substrate tRNAs from bulk tRNA isolates have minor effect on MiaA activity [54], indicating that MiaA has strong affinity towards its tRNA substrates. Therefore, previous work and my data indicate that, unlike TrmA and TruB, MiaA seems to specifically binds to its substrate tRNAs. Notably, MiaA has been reported to have approximately 4-fold stronger binding affinity towards $tRNA^{Phe}$ (K_D = 0.07 μM) [54] than for $tRNA^{Sec}$ and variants reported in this study (Table 3.3).

In general, the result of this study suggests that TruB binds to $tRNA^{Sec}$ and variants with lower affinity than either TrmA or MiaA. Similar binding affinity trend, particularly for TrmA and TruB, was observed for $tRNA^{Phe}$ [50]. TrmA and TruB, due to their chaperone activity and their preference for unmodified or single modified tRNA, are likely to act early during tRNA maturation where TruB modification activity has been proposed to occur prior to that of TrmA [50]. Therefore, considering that MiaA activity is minimally affected by pre-existing modifications [54], the likely preferred order of modification of $tRNA^{Sec}$ and variants *in vivo* would be first Ψ_{55}, then m^5U_{54} and lastly i^6A_{37} by MiaA. To the best of my knowledge, this is the first time that the interaction of $tRNA^{Sec}$ and variants with multiple modifying enzymes is investigated. As mentioned previously, the goal of this section of the project was to evaluate if

the *in vitro* transcribed tRNAs were successfully prepared and that the purified enzymes were active. Thus, the results demonstrated that not only $tRNA^{Sec}$ and $tRNA^{Sec2}$, but also the engineered $tRNA^{UtuX}$ can be post-transcriptionally modified *in vitro*, indicating an interaction with the *E. coli* modifying machinery, which was generally unexplored prior to this study.

However, it is worth noticing that out of the four $tRNA^{Sec}$ post-transcription modifications (D_{20}, i^6A_{37}, m^5U_{54}, Ψ_{55}) reported in the literature [25], the modifying enzyme DusA (D_{20}) was excluded from this study due to the lack of reliable and well-stablished enzyme assay in our lab. Moreover, an attempt to monitor MiaA activity was made, but no success was achieved due to technical difficulties and only filter binding data are reported. While binding data does not necessarily prove whether an enzyme is active, it provides a good indication that MiaA recognizes its substrates and is likely active, but further work on assay optimization is required to get a more direct validation.

4.2 Stablishing an in vitro selenoprotein translation system

Despite the advances in the field of selenoprotein bioengineering, protein yields remain low relative to non-Sec containing proteins. There have been numerous proposed hypotheses that could potentially explain the low yield with the complexity of the Sec biosynthesis and insertional machinery being the common factor [10]. However, one area that remains vastly unexplored is the role of tRNA post-transcriptional modification for selenoprotein biosynthesis even though there is a growing body of evidence that suggests tRNA modifications play a role in tRNA stability and during translation, particularly those at position 34 and 37 [34, 35]. Thus,

getting a better understanding of tRNA post-transcriptional modification could provide some insight into developing more efficient selenoprotein synthesis systems, especially as more synthetic tRNAs are developed.

Having demonstrated that all three tRNAs used in this study, $tRNA^{Sec}$ and its variants, can be post-transcriptionally modified, we thought of investigating the extent to which the presence of those modifications could affect selenoprotein biosynthesis *in vitro*. However, an active Sec biosynthesis and incorporation machinery needed to be individually purified and validated prior to carrying out any protein synthesis experiment. Knowing that SelB preferentially binds to Sec-$tRNA^{Sec}$ over Ser-$tRNA^{Sec}$ or deacylated $tRNA^{Sec}$ [21], I prepared Ser-$tRNA^{Sec}$ and Sec-$tRNA^{Sec}$ and conducted a series of filter binding experiments to indirectly validate the activity of the entire enzymatic cascade including SelB (Figure 3.8). Interestingly, no statistically significant difference was observed for the binding affinity of SelA towards either deacylated $tRNA^{Sec}$ or Ser-$tRNA^{Sec}$ (Table 3.4) as judged by the standard deviation of the K_D values. Thus, the data only suggests that SelA is capable of binding to the substrate RNAs non-specifically. However, there was a 2-fold increase in the affinity of SelB towards Sec-$tRNA^{Sec}$ over the deacylated $tRNA^{Sec}$ (Figure 3.8 C, Table 3.4). Therefore, SelB filter binding data provided valuable insight into the purified Sec biosynthesis and insertional machinery. First and foremost, the data suggested that SelB is active as it can efficiently discriminate between Sec-$tRNA^{Sec}$ and deacylated tRNAs as indicated in the literature [21]. Secondly, the data also suggest that Sec-$tRNA^{Sec}$ was successfully prepared, advocating for at least partial activity of SerRS, SelA and SelD since each of those enzymes are crucial in the enzymatic cascade.

Nevertheless, it is worth mentioning that despite the increase in the binding affinity of SelB towards Sec-$tRNA^{Sec}$ observed in this study, the K_D value measured here is higher (K_D = 0.5 µM)

than the K_D reported in the literature (K_D = 0.2 pM) [21]. Although the discrepancy in the binding affinity could be accredited to different factors such for example enzyme activity and purity, the most likely explanation lies on the quality of the tRNA samples. Indeed, it is highly possible that the tRNA samples I prepared are composed of a mixture of deacylated tRNAs, Ser-tRNAs and Sec-tRNAs because of inefficiencies in the enzymatic reactions during the serylation and or Serine to Selenocysteine conversion by SelA. In this case, the filter binding experiments would not measure the affinity of SelB for Sec-tRNA^{Sec} but would reflect an average affinity for the heterogeneous tRNA mixture. Moreover, it is also possible that some degraded tRNA products or aborted transcripts are present in the assay considering that the radiolabelled tRNAs were not subject to size exclusion chromatography. Thus, any of those scenarios could explain the difference in the binding affinities measured in this study compared to the literature. However, determining approximately 2-fold increase in SelB affinity toward Sec-tRNA^{Sec} while being aware of the potential limitations of the data was compelling evidence to assume that the purified Sec biosynthesis and insertional machinery is in principle active relative to tRNA binding.

4.3 tRNA modifications enhance stop codon reassignment

To evaluate the effect of tRNA modifications on selenoprotein biosynthesis, a fluorescence-based reporter system was designed to test stop codon (UAG or UGA) reassignment (Figure 3.9). The reporter system along with the purified Sec biosynthesis machinery (tRNA^{Sec}, SelA, SelB, and SelD) was then supplemented to commercially available PURExpres® ΔRF123 cell-free system in the presence or absence of modifying enzymes, measuring the effect of tRNA

modification based on the fluorescence signal generated by the sfGFP reporter protein. Along those lines, the wild-type Sec biosynthesis and insertional machinery was first evaluated (Figure 3.10 A), using strategically selected controls to determine the suitability of the PURE system. The results shows that the addition of modifying enzymes and cofactors to the reaction results in an increase of the fluorescence signal compared to the reaction deprived of it (Figure 11), strongly suggesting a positive effect of the modifications on stop codon readthrough. More importantly, the increase in fluorescence can be accredited the most to MiaA since the reactions containing only MiaA (but no other tRNA modifying enzymes) generated comparable fluorescence signal to those reactions supplemented with MiaA, TrmA and TruB together. The results are in accordance with knock out *in vivo* experiments that demonstrated the positive role of MiaA in the efficiency and translational fidelity of stop codon reassignment [35], which makes sense considering that the modification is located within the anticodon loop. Although the exact mechanism of stop codon decoding with the help of the i^6A_{37} modification during translation is not fully understood, it has been proposed that the modification stabilizes the weaker base pair interaction between A_{36} and the first Uridine of the codon by preventing hydrogen bonding between U_{33} and A_{37} that could disrupt the integrity and proper structure of the anticodon loop [34].

Similar increase in fluorescence in the presence of modifying enzymes was observed for reactions supplemented with $tRNA^{Sec2}$ and $tRNA^{UtuX}$ (Figure 3.10 B-C), reading the UGA stop codon in the absence of the SECIS element. Interestingly, a previous study showed that the SECIS-independent approach mediated by EF-Tu and $tRNA^{UtuX}$ yielded similar results to the wild-type SECIS-dependent SelB mediated mechanism [40], which does not appear to be the case in this study as approximately 20% increase in peak fluorescence for the SECIS-dependent

recoding by $tRNA^{Sec}$ was detected compared to either of the SECIS-independent mechanisms (Figure 3.10). However, it is worth highlighting that the *in vitro* translation data reported in the literature were generated by supplementing cell-free systems with much higher concentration of synthetic tRNA (70 µM Sec-$tRNA^{UtuX}$ vs 12 µM $tRNA^{Sec}$). Moreover, the data reported in this study was generated using tRNA modifications to the PURE reactions while the previous study in the literature only used unmodified tRNAs, suggesting that the results are nor fully comparable. Interestingly, the reaction that was carried out in the absence of supplemented $tRNA^{Sec}$ (Figure 3.10 A) shows comparable fluorescence signal to those supplemented with $tRNA^{UtuX}$, indicating that some $tRNA^{Sec}$ is present within the PURE system and that the additional $tRNA^{Sec}$ improves translation.

However, the obvious trend is that in both, SelB-mediated or EF-Tu mediated approaches, the incorporation of modifying enzymes increased peak fluorescence of the reporter protein, ranging from 35 to 45% increase compared to the reactions lacking the modifying enzymes (Figure 11). Thus, my results strongly suggest that tRNA modifications play positive roles in stop codon readthrough *in vitro*, which in turn could correlate to higher yields of selenoprotein. *In vivo*, these modifications would likely have been introduced by the endogenous modification enzymes, suggesting that the presence of modification alone may not be sufficient to counteract some of the low rate of Sec incorporation reported in the literature [22]. Altogether, the data of this thesis provides some insight into the role of tRNA modifications that could be useful for the field of selenoprotein bioengineering.

Firstly, the data show that synthetic $tRNA^{UtuX}$ interacts with the tRNA modifying machinery and that the presence of modifications improve translation of the selenocysteine reporter protein. Accordingly, it is also possible that other synthetic tRNAs fail to get modified or simply get

excessively modified *in vivo*, potentially affecting translation negatively. Thus, from a bioengineering perspective, one should place some consideration to the potential modification profile of a synthetic tRNA by considering the possible interaction with tRNA modifying enzymes while engineering the tRNAs. Secondly, the data show that MiaA alone is probably sufficient to spike an increase in fluorescence signal. Therefore, when trying to optimize selenoprotein biosynthesis *in vitro*, adding MiaA or modifying enzymes that act on the anticodon loop should have a priority. Lastly, although an improvement of fluorescence signal was observed in the presence of modifying enzymes, the peak fluorescence remains approximately 30% lower when compared to translation of sense codon by $tRNA^{Phe}$ even in the presence of SECIS element (Figure 3.11D). Thus, while removing the SECIS element provides great flexibility to add Sec residues freely within any given protein sequence, the SECIS element alone does not appear to be the major limiting factor for protein yields but rather the stop codon re-assignment mechanism or other factors during the Sec biosynthesis process as described [10].

4.4 Future direction

Recently, a new generation of *allo*-tRNAs have been developed for selenoprotein bioengineering, acting as better EF-Tu substrates that lead to higher *in vivo* protein yields than $tRNA^{UtuX}$ [46]. Unlike $tRNA^{UtuX}$ that was rationally engineered from $tRNA^{Ser}$ and $tRNA^{Sec}$, these *allo*-tRNAs are neither related to $tRNA^{Ser}$ nor to $tRNA^{Sec}$ but contain the required elements for interacting with the Sec biosynthesis machinery. Therefore, it would be highly interesting to see whether the peak fluorescence difference between translation of stop and sense codons observed

in this study is reduced using the *allo*-tRNAs, leading to a better understanding of the limiting factors of the Sec biosynthesis and insertional machinery.

It has been suggested that the reason why $tRNA^{Sec}$ has a lower number of modifications compared to $tRNA^{Ser}$ is because $tRNA^{Sec}$ needs to be more flexible to be able to interact with multiple enzymes involved in the Sec biosynthesis and insertion machinery. Thus, getting better understanding of the modification profile of the *allo*-tRNAs would provide some insight that could potentially help to answer the previous hypothesis. Moreover, knowing what modification is present in the anticodon loop of these tRNAs could help to solve the poorly understood mystery of the lack of $ms^2i^6A_{37}$ hypermodification in $tRNA^{Sec}$ compared to other tRNAs containing the A36-A37-A38 motif.

Although the results in this study provide evidence for the positive role of tRNA modification on stop codon reassignment that could be directly correlated to selenoprotein biosynthesis, it is important to highlight some of the limitations of the data with the goal of optimising future experiments. The main limitation of the data is that there is no clear evidence whether the tRNA modifications influence translation fidelity. The reporter system relies on stop codon readthrough to generate a fluorescence signal. However, the fluorescence signal does not reflect whether serine was misincorporated, which could certainty be the case if SelA fails to fully convert all the serine into selenocysteine prior to translation. Thus, to get a better insight into translation fidelity, the translated product would need to be purified and subject to mass spectrometry, which could be limited by the yields of the PURE system. Moreover, the data does not show the readthrough efficiency as we did not measure the fraction of terminated peptides.

Therefore, in future experiments a new reporter system needs to be implemented that can discriminate between the Ser and Sec such as the intein-based reporter recently developed [46].

In addition, the data must be completed in triplicates to accurately calculate statistical values. Moreover, it would be best to conduct western blotting using the FLAG tag to determine the readthrough efficiency while further work is required to stablish a reliable assay to determine the MiaA activity.

Another relevant limitation of the PURE reaction data could be the effect of macromolecule crowding, which is a known factor that can affect *in vitro* studies [60]. Generally, *in vitro* studies are conducted in aqueous buffered conditions, reducing the complexity of the system normally found in nature, which affect among other factors molecular collision, binding interaction, and enzymes kinetics [60]. In general, the data in this study indicates that the presence of modifying enzymes increase expression of the reporter protein as judged by the fluorescence emission profile. However, PURE reactions in the absence of modifying enzymes and cofactor are less crowded; therefore, the components in the system are diluted relative to the reactions supplemented with modifying enzymes and co-factors. Thus, as a future direction of the project, control reactions would be needed were the absence of modifying enzymes and cofactors is compensated for by supplementing the system with crowding agents. While there are various crowding agents that could potentially be implemented [61, 62], serum albumin (BSA) is commonly used to mimic cellular environments in coupled *in vitro* transcription-translation reactions [63, 64] and is therefore the ideal initial candidate to test in control reactions.

Beside the crowding agent controls with BSA, another set of controls that needs to be carried out are reactions in the presence of all release factors. Currently, all the PURE reactions were conducted with different combinations of release factors depending on the stop codon used (RF1 and RF3 for UGA; RF2 and RF3 for UAG). However, it is unclear what the effect of adding the remaining release factor to the corresponding PURE reactions would be. While it is not well

understood if RF1 competes with Sec-tRNA for UAG binding, RF2 has been shown to be a weak competitor, if ever, for the binding of UGA stop codon *in vivo* [22]. Therefore, removing the corresponding release factors could either be influencing the crowding effect in the reactions or ensuring translation is not terminated, particularly in the least understood case of RF1. However, regardless of the effect, a set of controls are required to get a better understanding of the system in presence of all release factors to gain information for future potential optimisations.

Lastly, having observed a positive effect of tRNA modifications on potentially selenocysteine biosynthesis, the optimal goal for the project would be to optimize a coupled *in vitro* transcription-translation system for selenoprotein bioengineering. Currently, there have been numerous improvements to cell-free systems in terms of protein yields that include higher concentrations of existing components such as elongation factors (EF-Tu and EF-G), GTP, and RF123 but also supplementation with other proteins such as the chaperone system GroEL/ES and ribosomal protein S1 [63, 64]. Therefore, it would be crucial to carry out PURE reactions supplemented with high concentration of components mentioned above in addition to high concentration of Sec-tRNAs carrying the post-transcription modifications.

4.5 Conclusion

I successfully completed the objective of investigating the potential role of tRNA post-transcriptional modifications on selenoprotein biosynthesis. The data in this research demonstrates that adding the $tRNA^{Sec}$ modifying enzymes TrmA, TruB and MiaA to coupled *in vitro* transcription-translation reactions have positive effects on the peak fluorescence signal of

the reporter protein, which could potentially lead to increasing selenoprotein yields. Moreover, it was demonstrated that a single modification in the anticodon loop at position A37 by the modifying enzyme MiaA is likely sufficient to stimulate the increase in translation efficiency. Thus, the work of this thesis can provide valuable insight into engineering or optimizing future platforms for selenoprotein bioengineering *in vitro*. Nevertheless, there are limitations to the data obtained in this study such as incomplete data set, lack of more informative reporter system to evaluate translation fidelity during stop codon reassignment and the lack of crowding effect evaluation which should be addressed in the future.

REFERENCES

1. Turner, R. J., Weiner, J. H., & Taylor, D. E. (1998). Selenium metabolism in Escherichia coli. *Biometals*, *11*(3), 223-227.

2. Chen, Y., Mo, H. Z., Hu, L. B., Li, Y. Q., Chen, J., & Yang, L. F. (2014). The endogenous nitric oxide mediates selenium-induced phytotoxicity by promoting ROS generation in Brassica rapa. *PloS one*, *9*(10), e110901.

3. Zeng, H., & Combs Jr, G. F. (2008). Selenium as an anticancer nutrient: roles in cell proliferation and tumor cell invasion. *The Journal of nutritional biochemistry*, *19*(1), 1-7.

4. Zeng, H., & Combs Jr, G. F. (2008). Selenium as an anticancer nutrient: roles in cell proliferation and tumor cell invasion. *The Journal of nutritional biochemistry*, *19*(1), 1-7.

5. Agris, P. F., Eruysal, E. R., Narendran, A., Väre, V. Y., Vangaveti, S., & Ranganathan, S. V. (2018). Celebrating wobble decoding: Half a century and still much is new. *RNA biology*, *15*(4-5), 537-553.

6. Nasim, M. J., Zuraik, M. M., Abdin, A. Y., Ney, Y., & Jacob, C. (2021). Selenomethionine: A Pink Trojan Redox Horse with Implications in Aging and Various Age-Related Diseases. *Antioxidants*, *10*(6), 882.

7. Xia, X., Zhang, X., Liu, M., Duan, M., Zhang, S., Wei, X., & Liu, X. (2021). Toward improved human health: efficacy of dietary selenium on immunity at the cellular level. *Food & Function*, *12*(3), 976-989.

8. Rayman, M. P. (2012). Selenium and human health. *The Lancet*, *379*(9822), 1256-1268.

9. Rayman, M. P. (2020). Selenium intake, status, and health: a complex relationship. *Hormones*, *19*(1), 9-14.

10. Fu, X., Söll, D., & Sevostyanova, A. (2018). Challenges of site-specific selenocysteine incorporation into proteins by Escherichia coli. *RNA biology*, *15*(4-5), 461-470.

11. Zhang, Y., Romero, H., Salinas, G., & Gladyshev, V. N. (2006). Dynamic evolution of selenocysteine utilization in bacteria: a balance between selenoprotein loss and evolution of selenocysteine from redox active cysteine residues. *Genome biology*, *7*(10), R94.

12. Foster, C. B. (2004). Selenoproteins and the metabolic features of the archaeal ancestor of eukaryotes. *Molecular biology and evolution*, *22*(3), 383-386.

13. Ferry, J. G. (1990). Formate dehydrogenase. *FEMS microbiology reviews*, *7*(3-4), 377-382.

14. Hartmann, T., Schwanhold, N., & Leimkühler, S. (2015). Assembly and catalysis of molybdenum or tungsten-containing formate dehydrogenases from bacteria. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, *1854*(9), 1090-1100.

15. Nauser, T., Steinmann, D., & Koppenol, W. H. (2012). Why do proteins use selenocysteine instead of cysteine?. *Amino acids*, *42*(1), 39-44.

16. Arner, E. S. J. Selenoproteins-What unique properties can arise with selenocysteine in place of cysteine? *Exp. Cell Res.* **316**, 1296–1303 (2010).

17. Härtlein, M., & Cusack, S. (1995). Structure, function and evolution of seryl-tRNA synthetases: implications for the evolution of aminoacyl-tRNA synthetases and the genetic code. *Journal of molecular evolution*, *40*(5), 519-530.

18. Leinfelder, W., Forchhammer, K., Zinoni, F., Sawers, G., Mandrand-Berthelot, M. A., & Böck, A. (1988). Escherichia coli genes whose products are involved in selenium metabolism. *Journal of Bacteriology*, *170*(2), 540-546.

19. Itoh, Y., Bröcker, M. J., Sekine, S. I., Hammond, G., Suetsugu, S., Söll, D., & Yokoyama, S. (2013). Decameric SelA• tRNASec ring structure reveals mechanism of bacterial selenocysteine formation. *Science*, *340*(6128), 75-78.

20. Fischer, N., Neumann, P., Bock, L. V., Maracci, C., Wang, Z., Paleskava, A., ... & Stark, H. (2016). The pathway to GTPase activation of elongation factor SelB on the ribosome. *Nature*, *540*(7631), 80-85.

21. Paleskava, A., Konevega, A. L., & Rodnina, M. V. (2010). Thermodynamic and kinetic framework of selenocysteyl-tRNASec recognition by elongation factor SelB. *Journal of Biological Chemistry*, *285*(5), 3014-3020.

22. Kotini, S. B., Peske, F., & Rodnina, M. V. (2015). Partitioning between recoding and termination at a stop codon–selenocysteine insertion sequence. *Nucleic acids research*, *43*(13), 6426-6438.

23. Betat, H., & Mörl, M. (2015). The CCA-adding enzyme: A central scrutinizer in tRNA quality control. *Bioessays*, *37*(9), 975-982.

24. El Yacoubi, B., Bailly, M., & de Crécy-Lagard, V. (2012). Biosynthesis and function of posttranscriptional modifications of transfer RNAs. *Annual review of genetics*, *46*, 69-95.

25. Boccaletto, P., Machnicka, M. A., Purta, E., Piątkowski, P., Bagiński, B., Wirecki, T. K., ... & Helm, M. (2017). MODOMICS: a database of RNA modification pathways. 2017 update. *Nucleic acids research*, *46*(D1), D303-D307.

26. Lorenz, C., Lünse, C. E., & Mörl, M. (2017). tRNA modifications: impact on structure and thermal adaptation. *Biomolecules*, *7*(2), 35.

27. Keffer-Wilkes, L. C., Veerareddygari, G. R., & Kothe, U. (2016). RNA modification enzyme TruB is a tRNA chaperone. *Proceedings of the National Academy of Sciences*, *113*(50), 14306-14311.

28. Keffer-Wilkes, L. C., Soon, E. F., & Kothe, U. (2020). The methyltransferase TrmA facilitates tRNA folding through interaction with its RNA-binding domain. *Nucleic acids research*, *48*(14), 7981-7990.

29. Bishop, A. C., Xu, J., Johnson, R. C., Schimmel, P., & de Crécy-Lagard, V. (2002). Identification of the tRNA-dihydrouridine synthase family. *Journal of Biological Chemistry*, *277*(28), 25090-25095.

30. Yu, F., Tanaka, Y., Yamashita, K., Suzuki, T., Nakamura, A., Hirano, N., ... & Tanaka, I. (2011). Molecular basis of dihydrouridine formation on tRNA. *Proceedings of the National Academy of Sciences*, *108*(49), 19593-19598.

31. Alian, A., Lee, T. T., Griner, S. L., Stroud, R. M., & Finer-Moore, J. (2008). Structure of a TrmA–RNA complex: A consensus RNA fold contributes to substrate selectivity and catalysis in m5U methyltransferases. *Proceedings of the National Academy of Sciences*, *105*(19), 6876-6881.

32. Ranjan, N., & Rodnina, M. V. (2016). tRNA wobble modifications and protein homeostasis. *Translation*, *4*(1), e1143076.

33. Persson, B. C., Esberg, B., Olafsson, O., & Björk, G. R. (1994). Synthesis and function of isopentenyl adenosine derivatives in tRNA. *Biochimie*, *76*(12), 1152-1160.

34. Schweizer, U., Bohleber, S., & Fradejas-Villar, N. (2017). The modified base isopentenyladenosine and its derivatives in tRNA. *RNA biology*, *14*(9), 1197-1208.

35. Crnković, A., Vargas-Rodriguez, O., Merkuryev, A., & Söll, D. (2018). Effects of heterologous tRNA modifications on the production of proteins containing noncanonical amino acids. *Bioengineering*, *5*(1), 11.

36. Jameson, R. R., & Diamond, A. M. (2004). A regulatory role for Sec tRNA [Ser] Sec in selenoprotein synthesis. *RNA*, *10*(7), 1142-1152.

37. Sahin, S., Cai, R., Milton, R. D., Abdellaoui, S., Macazo, F. C., & Minteer, S. D. (2018). Molybdenum-dependent formate dehydrogenase for formate bioelectrocatalysis in a formate/O2 enzymatic fuel cell. *Journal of the Electrochemical Society*, *165*(3), H109.

38. Sokol, K. P., Robinson, W. E., Oliveira, A. R., Warnan, J., Nowaczyk, M. M., Ruff, A., ... & Reisner, E. (2018). Photoreduction of CO2 with a formate dehydrogenase driven by photosystem II using a semi-artificial Z-scheme architecture. *Journal of the American Chemical Society*, *140*(48), 16418-16422.

39. Squires, J. E., & Berry, M. J. (2008). Eukaryotic selenoprotein synthesis: mechanistic insight incorporating new factors and new functions for old factors. *IUBMB life*, *60*(4), 232-235.

40. Miller, C., Bröcker, M. J., Prat, L., Ip, K., Chirathivat, N., Feiock, A., ... & Söll, D. (2015). A synthetic tRNA for EF-Tu mediated selenocysteine incorporation in vivo and in vitro. *FEBS letters*, *589*(17), 2194-2199.

41. Thyer, R., Robotham, S. A., Brodbelt, J. S., & Ellington, A. D. (2015). Evolving tRNASec for efficient canonical incorporation of selenocysteine. *Journal of the American Chemical Society*, *137*(1), 46-49.

42. Bröcker, M. J., Ho, J. M., Church, G. M., Söll, D., & O'Donoghue, P. (2014). Recoding the genetic code with selenocysteine. *Angewandte Chemie International Edition*, *53*(1), 319-323.

43. Cheng, Q., & Arnér, E. S. (2017). Selenocysteine insertion at a predefined UAG codon in a release factor 1 (RF1)-depleted Escherichia coli host strain bypasses species barriers in recombinant selenoprotein translation. *Journal of Biological Chemistry*, *292*(13), 5476-5487.

44. Bartholomae, M., Baumann, T., Nickling, J. H., Peterhoff, D., Wagner, R., Budisa, N., & Kuipers, O. P. (2018). Expanding the genetic code of Lactococcus lactis and Escherichia coli to incorporate non-canonical amino acids for production of modified lantibiotics. *Frontiers in microbiology*, *9*, 657.

45. Ma, Y., Di Salvo, M. L., & Budisa, N. (2018). Self-directed in cell production of methionine analogue azidohomoalanine by synthetic metabolism and its incorporation into model proteins. In *Noncanonical Amino Acids* (pp. 127-135). Humana Press, New York, NY.

46. Chung, C. Z., Miller, C., Söll, D., & Krahn, N. (2021). Introducing selenocysteine into recombinant proteins in Escherichia coli. *Current protocols*, *1*(2), e54.

47. Evans, R. M., Krahn, N., Murphy, B. J., Lee, H., Armstrong, F. A., & Söll, D. (2021). Selective cysteine-to-selenocysteine changes in a [NiFe]-hydrogenase confirm a special

position for catalysis and oxygen tolerance. *Proceedings of the National Academy of Sciences*, *118*(13).

48. Thanbichler, M., & Böck, A. (2002). Selenoprotein biosynthesis: purification and assay of components involved in selenocysteine biosynthesis and insertion in Escherichia coli. *Methods in enzymology*, *347*, 3-16.

49. Griffith, O. M. (2011). Large Volume Precipitation of Proteins with Ammonium Sulfate using Thermo Scientific Fiberlite Carbon Fiber Rotors. *Thermo Scientific: ANCFGAMSULF*, *511*.

50. Schultz, S. K. L., & Kothe, U. (2020). tRNA elbow modifications affect the tRNA pseudouridine synthase TruB and the methyltransferase TrmA. *RNA*, *26*(9), 1131-1142.

51. Wright, J. R., Keffer-Wilkes, L. C., Dobing, S. R., & Kothe, U. (2011). Pre-steady-state kinetic analysis of the three Escherichia coli pseudouridine synthases TruB, TruA, and RluA reveals uniformly slow catalysis. *RNA*, *17*(12), 2074-2084.

52. Fischer, N., Paleskava, A., Gromadski, K. B., Konevega, A. L., Wahl, M. C., Stark, H., & Rodnina, M. V. (2007). Towards understanding selenocysteine incorporation into bacterial proteins.

53. Gerrits, M., Budisa, N., & Merk, H. (2018). Site-specific chemoselective pyrrolysine analogues incorporation using the cell-free protein synthesis system. *ACS Synthetic Biology*, *8*(2), 381-390.

54. Leung, H. C. E., Chen, Y., & Winkler, M. E. (1997). Regulation of Substrate Recognition by the MiaA tRNA Prenyltransferase Modification Enzyme of Escherichia coliK-12. *Journal of Biological Chemistry*, *272*(20), 13073-13083.

55. Kim, S. H., Sussman, J. L., Suddath, F. L., Quigley, G. J., McPherson, A., Wang, A. H. J., ... & Rich, A. (1974). The general structure of transfer RNA molecules. *Proceedings of the National Academy of Sciences*, *71*(12), 4970-4974.

56. Ishida, K., Kunibayashi, T., Tomikawa, C., Ochi, A., Kanai, T., Hirata, A., ... & Hori, H. (2011). Pseudouridine at position 55 in tRNA controls the contents of other modified nucleotides for low-temperature adaptation in the extreme-thermophilic eubacterium Thermus thermophilus. *Nucleic acids research*, *39*(6), 2304-2318.

57. Rother, M., & Krzycki, J. A. (2010). Selenocysteine, pyrrolysine, and the unique energy metabolism of methanogenic archaea. *Archaea*, *2010*.

58. Wan, W., Tharp, J. M., & Liu, W. R. (2014). Pyrrolysyl-tRNA synthetase: an ordinary enzyme but an outstanding genetic code expansion tool. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, *1844*(6), 1059-1070.

59. Crnković, A., Suzuki, T., Söll, D., & Reynolds, N. M. (2016). Pyrrolysyl-tRNA synthetase, an aminoacyl-tRNA synthetase for genetic code expansion. *Croatica chemica acta. Arhiv za kemiju*, *89*(2), 163.

60. Silverstein, T. P., & Slade, K. (2019). Effects of macromolecular crowding on biochemical systems. *Journal of Chemical Education*, *96*(11), 2476-2487.

61. Ge, X., Luo, D., & Xu, J. (2011). Cell-free protein expression under macromolecular crowding conditions. *PLoS One*, *6*(12), e28707.

62. Norred, S. E., Caveney, P. M., Chauhan, G., Collier, L. K., Collier, C. P., Abel, S. M., & Simpson, M. L. (2018). Macromolecular crowding induces spatial correlations that control gene expression bursting patterns. *ACS Synthetic Biology*, *7*(5), 1251-1258.

63. Li, J., Gu, L., Aach, J., & Church, G. M. (2014). Improved cell-free RNA and protein synthesis system. *PLoS One*, *9*(9), e106232.

64. Sheahan, T., & Wieden, H. J. (2022). Ribosomal Protein S1 Improves the Protein Yield of an In Vitro Reconstituted Cell-Free Translation System. *ACS Synthetic Biology*.