# Engineered semi-orthogonal initiator tRNAs initiate translation at non-canonical start codons.

*A thesis submitted in partial fulfilment of the degree of Master of Research*

**Andras Hutvagner** ***ID: 43685250***

Supervisor: ***Dr. Paul R. Jaschke***

Department Molecular Science
Macquarie University

**15th October, 2018**

# Statement of Originality

This thesis entitled "**Engineered semi-orthogonal initiator tRNAs initiate translation at non-canonical start codons.**" is representative of the research study conducted between January 2018 and October 2018 for the completion of Master of Research degree in Molecular Science department at Macquarie University, New South Wales, Australia. The work presented in this thesis is certified to be original by the author, unless otherwise referenced in the literature and/or acknowledged of personal advice and suggestions.

This thesis is formatted according to Master of Research guidelines prescribed by the Faculty of Science & Engineering and Department of Molecular Sciences and has not been submitted for qualification or assessment to any other institution.

Andras Hutvagner,

**Student ID: 43685250**

# Acknowledgments

This second year on my Master of Research has been a journey in which I've learned a lot about myself, and I can confidently say that I come out of it a better student, researcher and ultimately a better person. Most of this journey would not have been possible without the constant support and guidance from my supervisor **Dr. Paul R. Jaschke**, to which I owe countless amounts of gratitude. I can safely say that the fulfilment of achieving the goals Dr. Jaschke and I set out to achieve at the beginning of this year is second to none, and I look forward to working alongside him in the years to come. I would also like to acknowledge the efforts from the members of the Master of Research committee including Dr. Louise Brown, A/Prof Bridget Mabbutt, Prof Paul A. Haynes, Dr Yuling Wang and Dr Alfonzo Garcia-Bennett. Without the support and structure provided by this team, my journey through my Master of Research would not have been possible.

At the beginning of this year, I joined a young lab with many ambitious young researchers. These guys have listened to my troubles, helped me through hard times and have provided constant support and guidance throughout the second year of my Masters of Research. More importantly, they welcomed me into their research team, and for this I would like to thank **Dominic Logel, Bradley Wright and Russel Vincent**.

I would also like to thank members of the Yeast 2.0 team for sharing a laboratory space and frequently providing me with solutions to problems I encountered throughout my research project. I would also like to thank my friends and peers **Thomas Collier, Geraldine Sullivan, Dominic Scopelliti, Sean Svanetti, Federico Van Kooten, James Werner** and **Bronte Simpson** for their help and constant support throughout the past two years, without you this year would not have been possible.

Finally, I would like to thank my mother, **Andrea Toth**, my father **Gyorgy Hutvagner**, my sister **Anna Hutvagner** and last but not least my beautiful girlfriend **Sally Rodzen** and her amazing family for the unconditional love and support throughout this year, even in the hardest of times. Without you this would not have been even remotely possible.

# Table of Contents

# Abstract

Bioengineering is an emerging discipline within molecular biology. It involves the study, the design and the construction of existing and new biological systems through biomolecular techniques. Being a relatively nascent field, bioengineering faces numerous obstacles in its practice such as reduced synthetic circuit efficiency and control, reduced cellular viability due to the introduction of exogenous circuits and possibilities of biocontainment issues of synthetic designs. Incorporating orthogonality into the design of synthetic circuits has previously been suggested as a promising tool for overcoming these prominent obstacles within the field of bioengineering. Within this thesis, I have developed and built an orthogonal set of initiator tRNAs with the aim of achieving translation initiation from non-conical start codons. I designed and constructed 9 new mutant initiator tRNAs, through altering the anticodon region of the *metY* gene in *Escherichia coli,* and tested their efficacy through an anticodon-codon reporter system. I was able to observe successful translation initiation from most of these mutant tRNAs from non-canonical start codons. I believe the results obtained within this dissertation will act as a pilot study for future research, which will explore a more complete set of initiator tRNA variants.

# 1 Introduction

## 1.1 Synthetic Biology: Scope and Current Challenges

Synthetic biology is an emerging field within molecular sciences combining multiple disciplines: including genetic engineering, molecular biology and computational science. Its scope is to design and construct new biological systems, and modify and refine currently existing ones. As a result of this, synthetic biology has multiple applications in medicine, industry and molecular science research (Weber and Fussenegger, 2012, Schmidt, 2012, Khalil and Collins, 2010, Endy, 2005).

Synthetic biology has adapted many principles from the computational sciences and engineering. These include abstracting the complexities of the genetic code to allow for the development of interchangeable genetic circuits or 'parts', the ability to fully control genetic pathways (synthetic or natural) independently from one another and the ability to introduce and remove genetic circuitry without adversely affecting the host organism (Endy, 2005, Ellis et al., 2011). Currently, the field of synthetic biology is focused on overcoming many obstacles which hinder its progress and limit its potential. One of the main obstacles synthetic biology faces is the ability to reduce cross talk between exogenous genetic devices and host cell DNA. Reducing cross talk between host and engineered devices has the potential to enhance the control of these introduced devices, increase their efficiency, and reduce their burden on host cells (Andrianantoandro et al., 2006).

Incorporating orthogonality into the design of synthetic circuits may be a useful tool in overcoming some of the current pitfalls of synthetic biology, such as reduced synthetic circuit efficiency and control, reduced cellular viability due to the introduction of exogenous circuits and possibilities of biocontainment issues. Within the literature, there currently exists a multitude of research into the implementation of orthogonality into the three main components of the central dogma: DNA replication, DNA transcription and RNA translation.

## 1.2 The Central Dogma of Molecular Biology

The central dogma of molecular biology refers to the transfer of information from DNA to RNA and from RNA into protein (**Fig 1**). A universal standard genetic code conserved across most forms of life exists as a language utilised by the central dogma

(Watson and Crick, 1953, Crick, 1970). The standard genetic code possesses 64 combinations of triplet codons which code for a total of 20 amino acids. These amino acids can be assembled in different combinations and may be diversely modified to form proteins which serve as functional molecules in living systems. Out of 64 codon combinations only 20 amino acids are formed, which implies that the genetic code contains redundancy and thus more than one codon must code for the same amino acid or no amino acids at all. This phenomenon was later called codon degeneracy and plays important roles in enhancing the tolerance of point mutations in DNA, as often changing one base pair in the DNA will result the same amino acid being incorporated (Lagerkvist, 1978).

DNA replication is one of the three fundamental processes within the central dogma (**Fig. 1**). It involves the replication of a native strand of DNA into two identical strands. This process is the basis for cell propagation and genetic inheritance, and thus is a highly conserved and tightly regulated process. DNA replication involves the unwinding of the DNA double helix by DNA helicases into two single strands which are stabilised by single stranded binding proteins. DNA polymerases then 'read' the single stranded templates in a 5' to 3' orientation and synthesises complementary strands for each parental strand of DNA.

Transcription is another fundamental step in the central dogma. Transcription involves reading the genetic code, which is present in the host cell as DNA (**Fig. 1**). RNA polymerases (RNAPs) read this template and through the process of transcription produce an RNA copy of this template. From here, the RNA copy is often further translated into a protein of a specific function or further processed into a functional RNA molecule such as a transfer RNA (tRNA), ribosomal RNA (rRNA), and messenger RNA (mRNA).

Translation is the final step of the central dogma. It involves the decoding of an mRNA template from RNA into a polypeptide chain consisting of amino acids (**Fig. 1**). This process involves a large number of different enzymes and molecules, but the key components include transfer RNAs (tRNAs – transfer molecules responsible for bringing the appropriate amino acid to the ribosome), the ribosome (the translation powerhouse) and aminoacyl-tRNA synthetases (aaRS – enzymes responsible for charging the appropriate amino acid onto its cognate tRNA). Much like the previous two steps of the central dogma, this step varies between eukaryotes and prokaryotes, however the gross outcome is the same: an mRNA template is translated into a polypeptide chain, which will later become a functional protein.

## 1.3 Orthogonality in Synthetic Biology

One approach to overcoming the issue of cross talk and other obstacles in synthetic biology is through the implementation of orthogonality into synthetic design. Orthogonality is a concept derived from mathematics and later adapted from computer science. It refers to the existence of two or more entities in a system, which operate completely independently from one another while existing in the same space. This is much like how a virtual machine operates inside the hardware of a computer but is independent from the main operating system (**Fig 1**). The two entities (main operating system and the virtual machine) exist in the same space but share no computational commonalities besides the hardware of the computer itself (Liu et al., 2018).

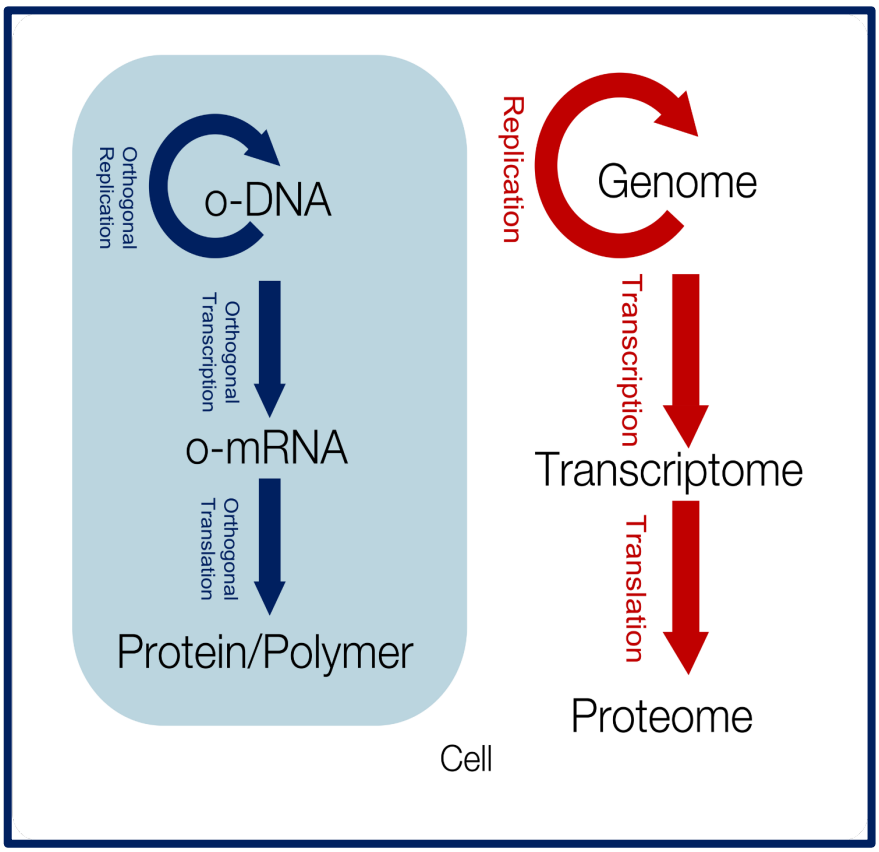



**Figure 1. A conceptual depiction of an Orthogonal Central Dogma running alongside a cell's housekeeping Central Dogma.** This system would exist inside of the cell and operate alongside, but independently of the host central dogma. Adapted from (Liu et al., 2018)

Biological circuits can be viewed in a similar manner, where the host cell is the physical hardware of the computer, and the host operating system is the host cell's machinery, responsible for DNA replication, transcription and translation of native host cell DNA (**Fig. 1**). Synthetic circuits are the virtual machines in this context and exist within the host cell but operate independently of the host cell's mechanisms (Liu et al., 2018). Currently, this is not the case, as synthetic circuitry shares the host cell's machinery for replication and expression. The dependency on host machinery has adverse repercussions on native cellular functions and the efficiency of the introduced synthetic circuit (**Fig.2**) (Borkowski et al., 2016).

This concept of orthogonality can thus be implemented into synthetic designs to potentially resolve some of the obstacles currently facing genetic engineers and aid in achieving the goals of reduced genetic cross talk between host DNA and exogenous circuits. Orthogonality may also be used to increase the efficiency of synthetic circuits, enhance genetic abstraction, furthering the impact synthetic biology currently has in both medicine and industry.

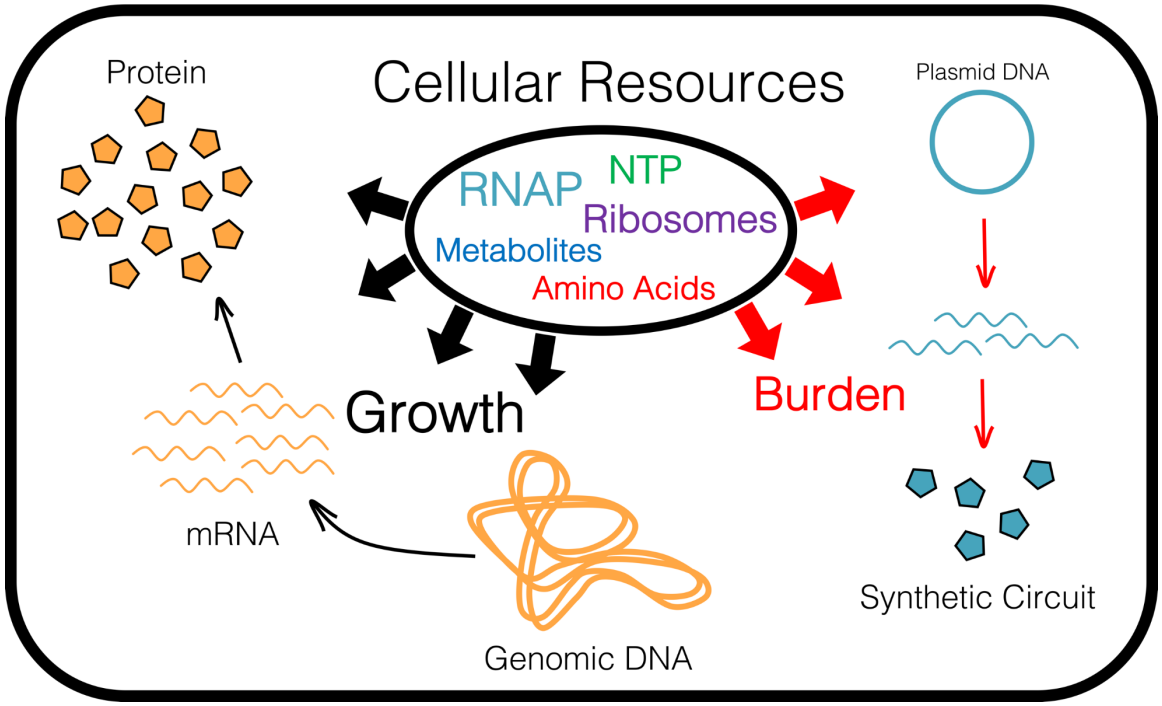



**Figure 2. The competition for shared resources between host and exogenous synthetic circuit.** The introduction of an exogenous circuit into host will cause cellular burden due to the competition for resources, as both the host and the exogenous circuit require the cellular resources for growth and protein production (Borkowski et al., 2016).

Some aspects of orthogonality have already been incorporated into synthetic designs, and researchers are developing orthogonal systems that encompass a large portion of the central dogma. These systems include orthogonal DNA replication, orthogonal transcription and orthogonal translation. There is also work in developing an alternative genetic code, allowing for synthetic biologists to design and build with more tools and less fear of their designs invading natural habitats. With further development of these systems, orthogonal synthetic designs will enable synthetic biology to break free from the constraints of the host cell machinery and mechanisms and bring the field closer to its true potential.

## 1.4 An Orthogonal Central Dogma

### *1.4.1 Development of Orthogonal Designed Genetic Codes*

It was found on several occasions that there are exceptions to the standard genetic code in nature. First discoveries of this were in human mitochondrial DNA (mtDNA), where researchers found that mtDNA encoded methionine from an AUA codon instead of the usual AUG in the standard genetic code (Barrell et al., 1979). Other variations in mtDNA include the assignment of the codon UAG to code for leucine, glutamine and alanine instead of the canonical stop codon found in the standard code (Knight et al., 2001). It is believed that these variances from the genetic code are a result of either codon reassignment due to mutational pressure of certain codons or due to mutations in the functional centre of a tRNA gene leading to changes in codon usage (Schultz and Yarus, 1996, Osawa and Jukes, 1989).

The design of an alternative genetic code, or the selective modification of the inherent rules in the current central dogma of molecular biology has sparked the interest of numerous research groups in recent years. Completely redesigning or partially modifying the standard code may act as a way of increasing the orthogonality of synthetic designs, and thus relieve synthetic biologists from the constraints of the standard code. Introducing a system which does not compete with the host machinery may also greatly reduce the risk of killing host cells through the introduction of larger, more complex synthetic circuits (Borkowski et al., 2016).

The standard genetic code is degenerate, and many different codons encode for the same amino acid. Ostrov et al. (2016) used the degeneracy of the genetic code to construct an *Escherichia coli* mutant with a recoded genome, that only has the ability to read 57

codons, while maintaining the ability to incorporate all 20 amino acids during protein synthesis. This recoded *E. coli* now has seven codons "free" to use for other purposes such incorporation of unnatural amino acids from engineered aminoacylsynthetase-tRNA pairs with various desired characteristics (Wang et al., 2014, Wang et al., 2007, Liu et al., 2018, Chin, 2014).

The use of organisms with an alternative genetic code can also be used as a form of biocontainment. For example, a recoded *E. coli* strain lacking all UAG stop codons was developed to inhibit horizontal gene transfer between cells (Ma and Isaacs, 2016). Engineering cells to be unable to undergo horizontal gene transfer, may be useful in introducing biocontainment into non-natural engineered species. For instance, the lack of horizontal gene transfer may stabilise introduced genetic circuits into these host cells. Furthermore, cells lacking horizontal gene transfer abilities, may result in the genetical isolation of the modified organisms from wild type populations. This would ensure control of the engineered species if they are released from controlled environments such as a laboratory (Liu et al., 2018).

### *1.4.2 Orthogonal DNA Replication*

There has been research into making aspects of the three main components of the central dogma orthogonal. For instance, an orthogonal DNA replication system in *Saccharomyces cerevisiae* has previously been developed. This system uses a modified DNA polymerase that targets a specific plasmid system and exclusively only replicates this plasmid (Ravikumar et al., 2014). Orthogonal DNA replication systems can also be engineered to replicate sequences containing unnatural nucleotides and even non-natural base pairs, both expanding the genetic code and leaving the host cell unaffected (Malyshev et al., 2014, Liu et al., 2018, Leontis et al., 2002).

### *1.4.3 Orthogonal Transcription*

Orthogonal transcription systems have also been previously developed to reduce host burden due to competition for transcription resources and machinery. An example of such system uses bacteriophage T7 RNA polymerase. The T7 RNA polymerase does not recognise host promoter regions, and likewise, host *E. coli* RNA polymerases do not recognise T7 promoter regions. (Tabor and Richardson, 1985, Studier et al., 1990, Studier and Moffatt,

1986). More recently, RNA-guided transcription initiation has been used to control the transcription of specific exogenous transcripts, reducing any interference with the host's transcriptome (Perez-Pinera et al., 2013).

### *1.4.4 Orthogonal Translation*

Research has previously focused on increasing translation orthogonality in order to achieve enhanced control, reduced cross-talk, and increased efficiency of introduced genetic circuits (**Fig. 3**). Orthogonal ribosome-mRNA pairs show promise in expressing synthetic circuitry with more sensitive control, reduced cell burden, and unaltered host translation efficiency. This concept was originally explored by Rackham and Chin (2005) where they constructed orthogonal ribosome–mRNA pairs which were able to process information alongside, but independently of, their wild type pairs. This orthogonal pair was created by directed evolution of 16S rRNA and the Shine-Dalgarno sequence of the cognate mRNAs.

Recently, a system of quadruplet start codons were developed, whereby mutant tRNA and ribosomes were able to incorporate unnatural amino acids into proteins (Wang et al., 2014).

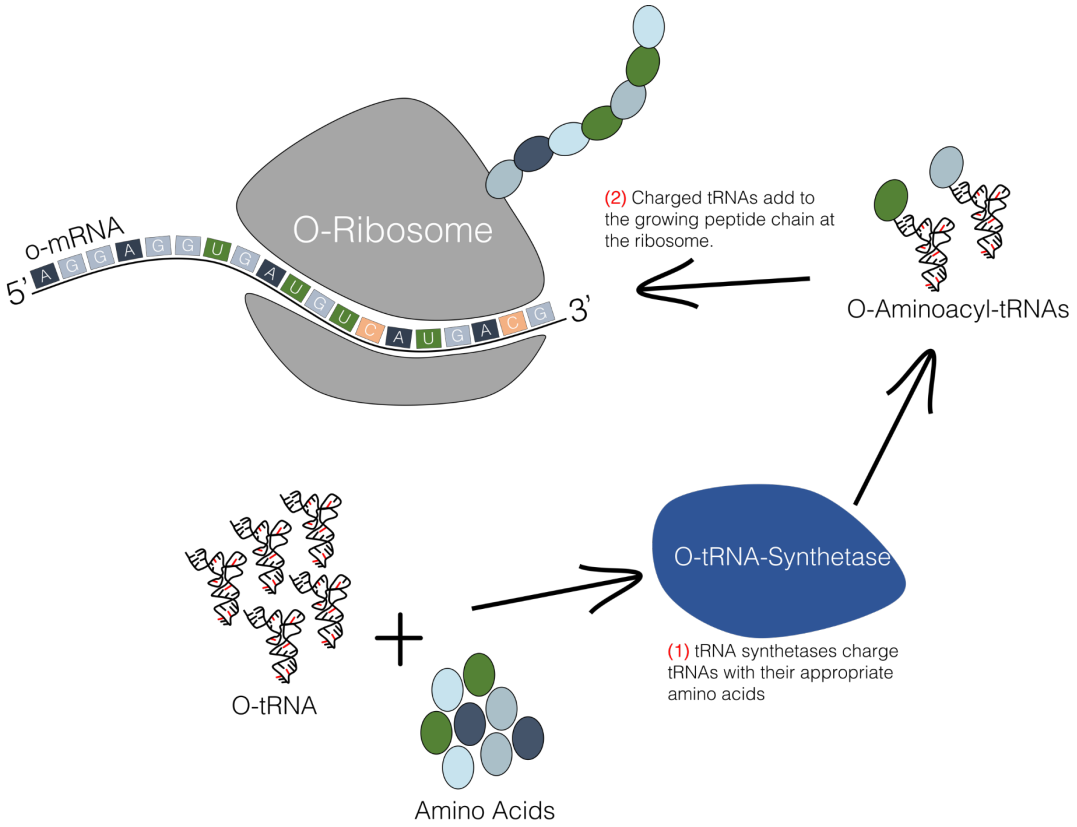



**Figure 3. A conceptual illustration of an orthogonal translation system in prokaryotes.** Different parts of the translation system may be engineered to be orthogonal, and act independently to the native host system.

### *1.4.5 Orthogonal tRNAs*

Orthogonal tRNAs have been developed for a multitude of different aims. There have been sets of orthogonal tRNAs developed which help incorporate unnatural amino acids into proteins (Wang et al., 2014, Wang et al., 2007, Liu et al., 2018, Chin, 2014, Chatterjee et al., 2012). These tRNAs, together with appropriate mutant ribosomes and aminoacyl-RNA synthetases (aaRSs) can act to expand the genetic code.

Developing mutually orthogonal tRNA-aaRS pairs has also been of great interest within the field of bioengineering. For instance, through directed evolution, new tRNA-aaRS pairs which work completely independently to wild type pairs have been developed (Neumann et al., 2010). aaRSs are enzymes which bind with tRNAs and charge them with their appropriate amino acids. Engineering these enzymes to interact with only cognate tRNAs and even to charge these tRNAs with unnatural amino acids has been achieved in

previous studies (Wang et al., 2014, Wang et al., 2007). This application has a large potential in utilising an expanded genetic code to further the scope of synthetic design.

Furthermore, it was shown that translation initiation can be achieved through non-native mechanisms. The native initiator tRNA was modified to have stop codon compliment in its anticodon region and initiate translation through an amber stop codon (UAG). This new non-canonical method of translation initiation showed to be relatively efficient (Varshney and RajBhandary, 1990). While this work was undertaken to expand current knowledge on translation initiation, and not necessarily with an orthogonal translation system in mind, it may serve as a proof of concept for further research regarding the development of orthogonal translation systems which utilise non-canonical start codons (Vincent, 2017, Mayer et al., 2003).

## 1.5 Translation Initiation

### *1.5.1 The initiator tRNA*

The initiator tRNA (i-tRNA) is responsible for bringing the first amino acid, a formylated methionine, to the P-site of the ribosome during translation initiation. This transfer molecule is unique from its sister elongator tRNAs (e-tRNA) as it has three uniquely conserved GC base pairs in the anticodon stem (Selmer et al., 2006) (**Fig. 5).** These GC pairs are essential for P-site binding of the i-tRNA, as they interact with the 16s rRNA on the 30S complex. Furthermore, this GC region has also been shown to cause the projection of the anticodon loop outwards, stabilising the anticodon-codon interaction, a characteristic which is not found in e-tRNAs (Barraud et al., 2008).

The formylation of the i-tRNA has been postulated to play important roles in differentiation from e-tRNAs. Formylation occurs after the i-tRNA is charged with its cognate amino acid (methionine) by an aaRS. The methionine found on i-tRNA are then formylated by methionyl-tRNA formyl aminoacyl (MTF) (Mayer et al., 2003) It was believed for some time that formylation of i-tRNA by MTF is the sole determining factor for binding affinity of IF2 towards the i-tRNA, recruiting it during translation initiation. This concept was further explored and validated in studies with mutant i-tRNA with altered acceptor stems. The results were of these studies found that these mutant i-tRNAs have reduced formylation, lower rates of IF2 interaction and in turn hindrances of migration towards the 30S complex *in vivo* (Varshney and RajBhandary, 1992, Guillon et al., 1993)

More recent studies have shown that IF2 affinity was not solely dependent on i-tRNA formylation, but also a highly conserved sequence (CAACCA) found on the acceptor stem, exclusively on i-tRNAs (**Fig. 5.**). It was also shown that this same sequence was also responsible for the formylation of the tRNA through interaction with (MTF). These studies demonstrated that the formylation of the i-tRNA and IF2 binding were independent from each other, however the same conserved sequence was responsible for both phenomena (Guenneugues et al., 2000, Simonetti et al., 2008, Mayer and RajBhandary, 2002).

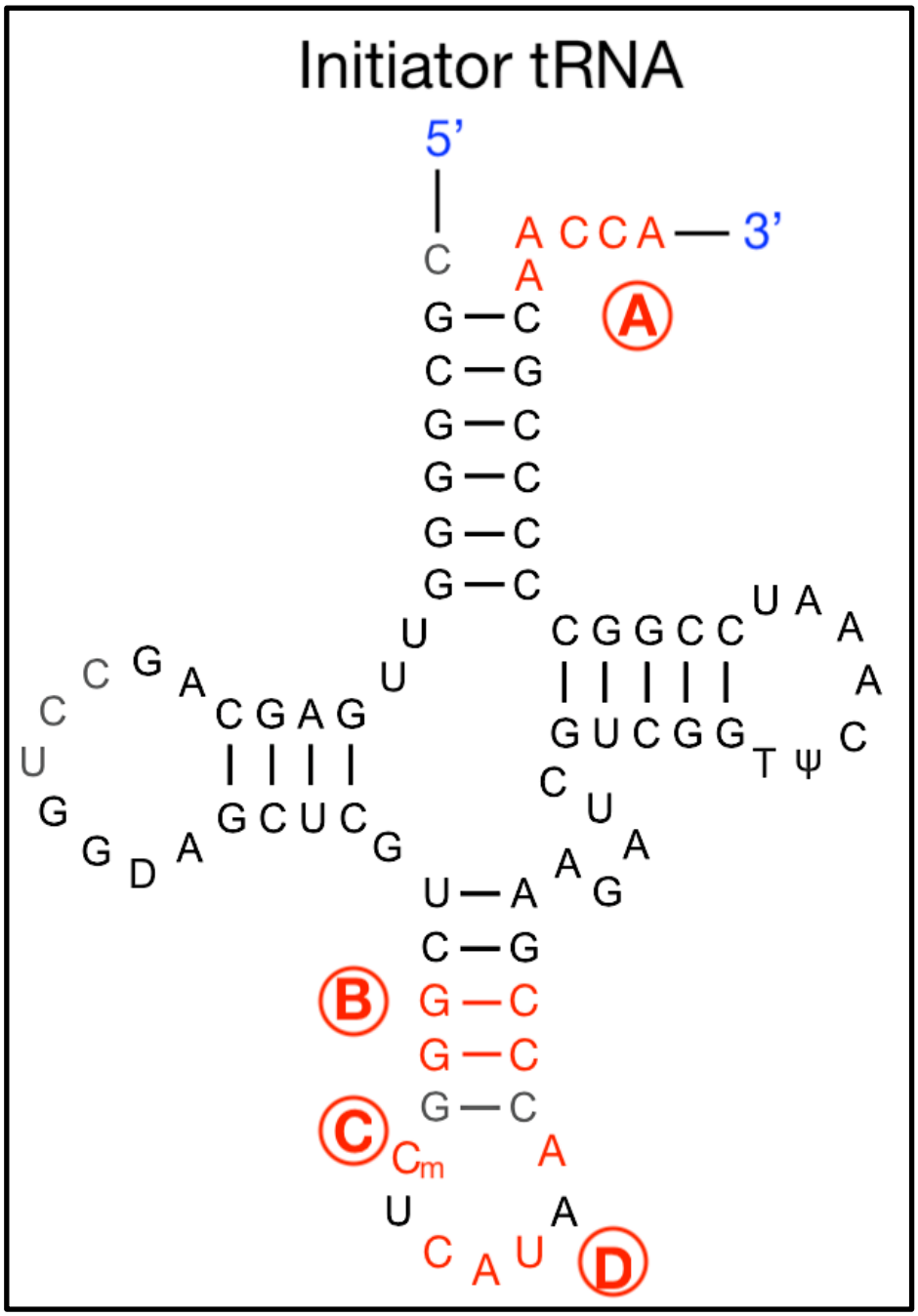



**Figure 4. The initiator tRNA and its important structures.** **(A)** Highly conserved acceptor stem sequence enhances IF2 affinity. **(B)** GC base pairs distinguish initiator tRNA from elongator tRNA. **(C)** $C_m$A pairing causes unusual anticodon loop configuration. **(D)** Anticodon region, complimentary to initiation codon on mRNA transcript.

### *1.5.2 Different modes of translation initiation in prokaryotes*

Translation initiation is the rate limiting step in translation, and in prokaryotes there have been three main modes observed, all of which occur within cells at different rates. Firstly, there is the 30S binding mode, where the 30S subunit of the ribosome recognises the

upstream Shine-Dalgarno sequence on the mRNA template. Following this, initiation factors 1, 2, and 3, the i-tRNA and the 50S ribosomal subunit form an initiation complex and translation initiation commences. Within this phase, it is believed that the 30S subunit dissociates from the mRNA when a stop codon is encountered but is re-assembled at the start of the next gene **(Fig.5A)** (Laursen et al., 2005). The second model of translation initiation is the 70S scanning mode, where the ribosomal subunits read through the whole transcriptome continuously. The various initiation factors and the i-tRNA assemble to the 70S complex as the ribosome reaches the Shine-Dalgarno region upstream of the open reading frame (ORF) of a gene **(Fig.5B)** (Yamamoto et al., 2016). The final, and most rare form of translation initiation is leaderless mRNA initiation. This is where the upstream region of the ORF lacks the Shine-Dalgarno sequence. In this case, it has been observed that a complex between Initiation factor 2 and the i-tRNA aid in translation initiation **(Fig.5C)** (Udagawa et al., 2004, Moll et al., 2002).

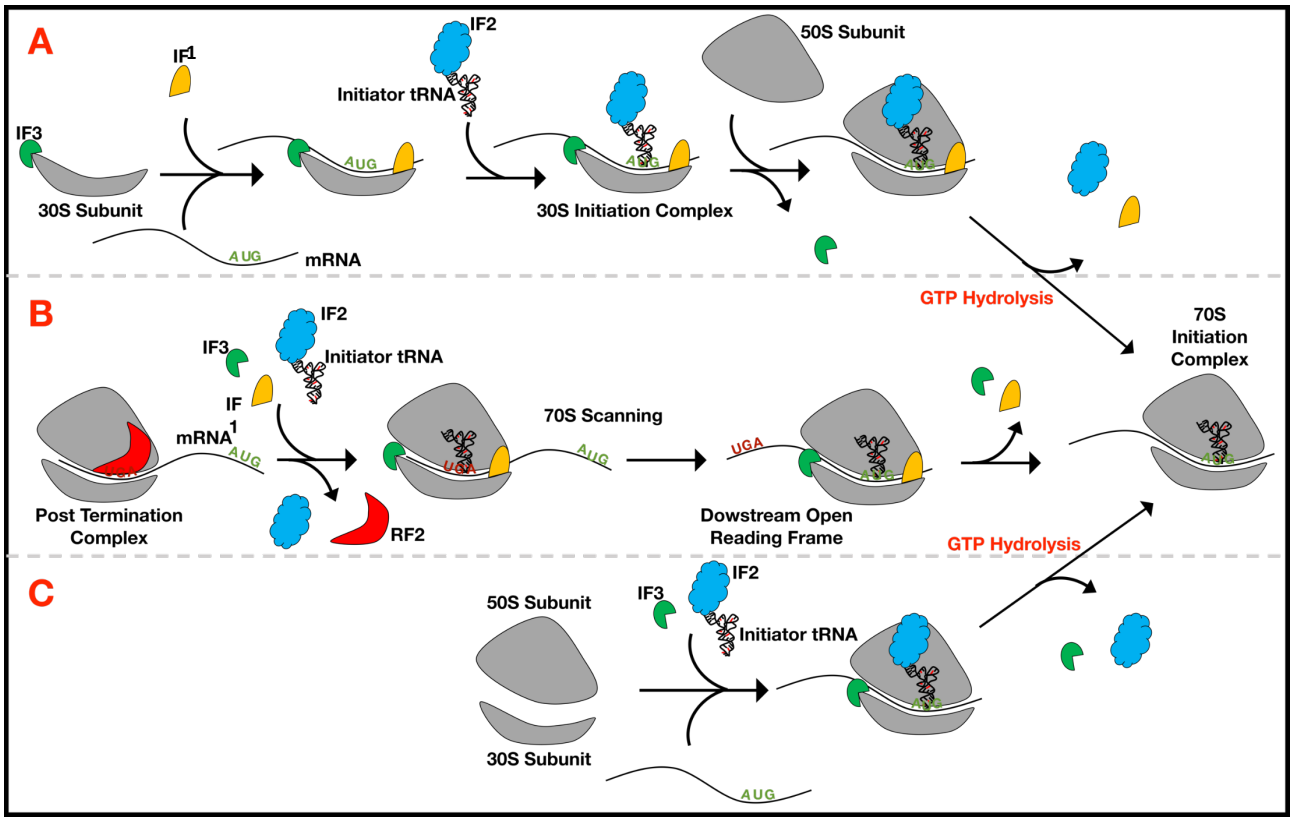



**Figure 5. Three different modes of translation initiation in prokaryotes. (A)** *30S binding mode,* where IF1, IF2 and IF3 work together to recruit and bind the initiator tRNA to the initiation codon on the mRNA. **(B)** *70S scanning mode*, whereby IF1 and the initiator tRNA scan the bicistronic mRNA for the next initiation site. **(C)** *Leaderless mRNA initiation*, involves the translation initiation without the presence of a Shine-Dalgarno sequence (Yamamoto et al., 2016).

### *1.5.3 The role of initiation factors in translation initiation*

Initiation factors are proteins that play pivotal roles in the translation initiation process. They aid in the formatting of the 30S subunit in prokaryotes and ensure stringent and correct codon-anticodon pairing at the start codon to prevent translation initiation at incorrect sites throughout the genome. There are three initiation factors involved within prokaryotic translation initiation: initiation factor 1 (IF1), initiation factor 2 (IF2) and initiation factor 3 (IF3).

IF1 is the smallest initiation factor and the least understood of the three. During 30S formation, it binds to the A-site of the ribosome, ensuring the correct binding of the i-tRNA within the ribosome. It has also been shown to regulate IF2 and IF3 throughout the initiation process, while also being involved in the disassociation of 70S subunits through the release of IF2 and the initiation of 30S subunit formation (Moazed et al., 1995, Carter et al., 2001, Gualerzi and Pon, 1990, Dottavio-Martin et al., 1979, Pon and Gualerzi, 1984, Wintermeyer and Gualerzi, 1983, Benne et al., 1973, Stringer et al., 1977, Brock et al., 1998)

IF2 has a highly conserved C-terminus which selects correctly charged and formylated i-tRNAs. It also has ribosomal dependant GTPase activity which interacts with the ribosome and anchors the i-tRNA (Guenneugues et al., 2000, Caserta et al., 2006, Julián et al., 2011, Simonetti et al., 2008)

IF3 has been shown to ensure the fidelity of codon-anticodon interaction during the binding of the i-tRNA to the start codon of the mRNA (Ayyub et al., 2017). It is also been shown to have roles in the detection of correct start codons on the mRNA transcript, thereby ensuring that the correct transcripts are translated. Furthermore, IF3 plays a vital role in promoting translation initiation though catalysing the dissociation of 70S complexes, freeing 30S complexes to form elsewhere within the cell (Milon et al., 2008).

### *1.5.4 The role of the start codon in translation initiation*

Start codons play an important role in the initiation of translation. They are triplet nucleotides from which translation initiates in both prokaryotes and eukaryotes. The most frequent nucleotide triplet found at the start codon in bacteria is AUG (82%), however translation initiation has been shown to also occur at considerable rates form the codons GUG(14%) and UUG(4%) (Vincent, 2017, Hecht et al., 2017). The canonical start codon, AUG, encodes for methionine, an amino acid which is found internal to proteins, and not

just at the start of the growing polypeptide chain. Thus, there are mechanisms in play which ensure that that correct AUG within an mRNA transcript are being used as initiation codons. One of the main selection mechanisms is the presence of a Shine-Dalgarno (SD) sequence upstream of the start codon (Shine and Dalgarno, 1974). SD sequences are highly conserved within prokaryotes, and function by complimentary base pairing with 16S rRNA on the ribosome, to anchor the mRNA onto the ribosome (Malys, 2012). The start codon also interacts with the anticodon region of the i-tRNA **(Fig 4.)** This interaction is essential to the stringency and fidelity of translation initiation and is mediated through IF3 (Hartz et al., 1990, Julián et al., 2011).

### *1.5.5 The role of codon-anticodon pairing in translation initiation*

The anticodon region of the i-tRNA is complimentary with the start codon on the mRNA transcript, and during the process of translation initiation, strong Watson-Crick base pairing occurs between these two RNA molecules. **(Fig. 6.)** The correct pairing of the anticodon-codon is essential to ensure high-efficiency translation initiation in prokaryotes, and thus the process is highly regulated by IF3 (Hartz et al., 1990).

Through structural analysis, it has been observed that IF3 recognises the anticodon-codon interaction as a unit, and not as two separate entities. Once the two RNAs bind, IF3 causes a conformational change in the 30S subunit, which can only be sustained by a strong anticodon-codon interaction. It is believed that this conformational change caused by IF3 acts as checkpoint in translation initiation to ensure the correct start codon-anticodon binding (Hussain et al., 2016, Julián et al., 2011).

Near cognate start codons (GUG and UUG) have also been shown to be sufficient for the efficient translation of many bacterial genes **(Fig. 6)**(Hecht et al., 2017). This may be due the thermostability of these near cognate sequences, with the nucleotides at the second position accounting for a large portion of the thermal stability, ensuring translation initiation even under IF3 destabilisation (Sussman et al., 1996).

Other non-cognate start codons have also been observed in nature, with translation efficiencies significantly lower than AUG and its near cognates. Most of these alternative start codons exhibit wobble base pairing at the third position **(Fig. 6).** Wobble base pairing involves non-Watson Crick base pairing between nucleotides, and in some instances, this form of pairing provides sufficient thermostability in anticodon-codon interactions to allow

for translation initiation to take place (Sussman et al., 1996, Varani and McClain, 2000, Cochella and Green, 2004, Murphy IV and Ramakrishnan, 2004).

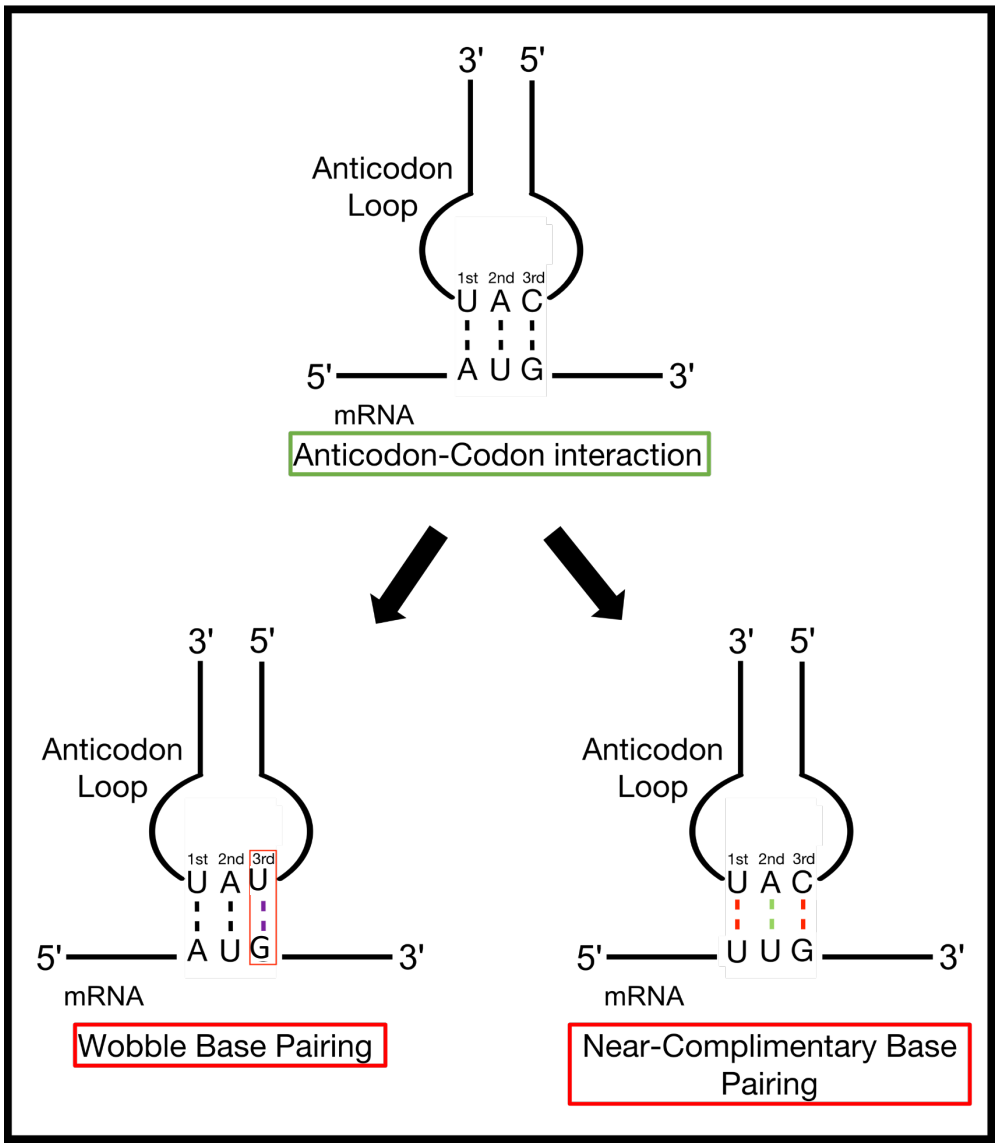



**Figure 6. Anticodon-codon interactions.** Anticodon-codon interaction occur between the start codon of the mRNA and the anticodon of the i-tRNA. Wobble base and near cognate interactions are also found in nature due to their favourable thermostability.

### *1.5.6 Non-canonical translation initiation in prokaryotes*

Translation initiation from non-canonical start codons has been demonstrated using i-tRNA with modified anticodons (Vincent, 2017, Varshney and RajBhandary, 1990, Mayer et al., 2003). Furthermore, studies by Hecht et al. (2017) showed successful translation of a florescent reporter green fluorescent protein (GFP) by 47 near cognate and non-canonical start codons in cells with only the wild-type i-tRNA. The main aim of previous studies on non-canonical translation initiation was to understand the mechanisms of translation rather

than to control the translation process (Mayer et al., 2003, Chattapadhyay et al., 1990, Schulman and Pelka, 1985). Currently, there are few reports outlining the use of i-tRNA to create an orthogonal translation initiation system.

## 1.6 Designing and creating an orthogonal set of i-tRNAs to control translation initiation in *E. coli*.

There has been extensive research regarding translation initiation in prokaryotes, revealing the importance of i-tRNAs and the anticodon-codon interaction for translation. Recent literature has also revealed that slight modifications to specific regions of the anticodon systems will still allow relatively high rates of translation initiation *in vivo*. Together, accumulated evidence suggests that modifying translation initiation through i-tRNA engineering may be a feasible path towards an efficient orthogonal translation system. This system would be expected to enhance the efficiency and control of exogenous synthetic circuits, reduce cross-talk, decrease cellular burden, and potentially enhance biocontainment of synthetic designs.

In this thesis, I created a new set of mutant i-tRNAs to take an important first step towards an orthogonal translation initiation system. For each engineered tRNA I will determine the translation initiation characteristics against all possible start codons and assess its ability to function as an orthogonal entity within *E. coli*.

## 1.7 Project aims and objectives

In this thesis project I aim to build and test a set of 9 new (11 in total) mutant i-tRNAs to expand the set of currently available initiation codons and to determine the anticodon-start codon fidelity characteristics within a host. *E. coli*.

I will be using a translation initiation system which uses mutant i-tRNAs with altered anticodon regions to initiate translation of fluorescent proteins with complimentary start codons. I will analyse the efficacy of mutant i-tRNA-protein pairs in translation initiation, and will also determine how the introduction of mutant tRNAs into a host may affect cell viability.

Specifically, the aims of this thesis are:

1) Design and construct 9 new (11 in total) i-tRNA with altered anticodons.
2) Measure the translation initiation efficiency of each tRNA against all 64 possible start codons.
3) Determine the fitness effects on the host *E. coli* cell from each mutant i-tRNA.
4) Identify mutant i-tRNA anticodon features that modulate translation initiation and orthogonality compared with wild-type i-tRNA.

# 2. Methods

## 2.1 Anticodon mutant plasmid construction

All of the DNA synthesised within this study (including oligonucleotide primers and genes) was obtained from Integrated DNA technologies (IDT). Variants of the *metY* gene were ordered and synthesised with the anticodon region of this gene being modified to altered sequences**.** These synthetic gene fragments were designed to have 20 bp overlaps with a linearised pULTRA backbone (Gibson et al., 2009).

The expression plasmid pULTRA:*metY*:tac was obtained from a previous study (Vincent, 2017) and was linearised with oligonucleotide primers**.** The primers were designed using the Primer3plus webtool, and were designed to have a 20bp overlapping region with the designed *metY* gene fragments.

The linear plasmids were digested with DPN1. DPN1 is an enzyme that cleaves CpG methylation sites in the parental plasmid, ensuring that instances of the native plasmid being re-transformed further in the pipeline of this study will not occur. Following DPN1 digestion, the samples were cleaned up using the GenElute PCR Clean-Up Kit (catalogue number NA1020-1KT) to remove any impurities from the DNA sample such as unwanted primers, nucleotides, enzymes and salts. Finally, the DPN1 digested and PCR-cleaned up plasmid backbones were measured to determine their concentrations using 1 µL of sample on the NanoDrop 500.

## 2.2 Gibson assembly of metY variants and pULTRA backbone

Synthetic *metY* variants were used to assemble mutant i-tRNA plasmids with altered anticodon regions. These synthetic genes contained 20bp overlaps with the linear pULTRA

plasmid, allowing for easy and efficient Gibson assembly. The assembly was undertaken as per the NEBuilder protocol from NEB (catalogue number E2621S).

The Gibson assembled pULTRA:*metY* plasmid variants were transformed into the expression strain NEB Turbo, obtained from NEB (catalogue number C29841), and made competent utilising the Transformation and Storage Solution (TSS) method. 10pg of the Gibson assembly mix was added to 200 µL of TSS competent NEB Turbo cells and were incubated on ice for 30 minutes. Following this, the cells were heat shocked for 30 seconds at 42 °C. The cells were then inoculated in 1 mL SOC medium (2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, and 20 mM glucose) and incubated on a shaker at 37 °C, 250 revolutions/minute (RPM) for 1 hour. Cells were them plated onto LB (miller) agar plates with spectinomycin (50 ug / mL) antibiotic and incubated overnight at 37 °C in a shaker at 200 RPM.

Colony PCR was undertaken on colonies to determine that successful transformants possessed the pULTRA:*metY* plasmids. It is not uncommon for the plasmid template to re-ligate in on itself, and give false positives during this step, thus designing primers which amplify different sized fragments for correctly assembled plasmids and re-ligated plasmids was essential**.** Using the KAPA2G Robust Hot Start colony screening protocol, transformant colonies were screened using a KAPA2G Master mix (obtained from Sigma Aldrich catalogue number KK5512) and oligonucleotide screening primers**.**

The products of the PCR reaction were run on a 2% agarose gel, with successful transformants expected to show 432 bp sized bands, whereas the re-ligated plasmids an expected size of 355 bp. The successful transformants were then inoculated in 7 mL of LB (miller) with 7 µL spectinomycin (50 ug / mL) overnight at 37 °C in a shaker at 200 RPM, in preparation for plasmid mini-prep and glycerol stocking for long term cryogenic storage.

Overnight cultures of successful transformants were prepared for both sequence confirmation and long term cryogenic storage. Plasmid extraction was undertaken using the Qiagen Plasmid Miniprep kit (catalogue number 27104). Once extracted, the concentration of the plasmids was then measured, to ensure that they were appropriate for Sanger sequencing reactions (above 100 ng / µL). Alongside the plasmids themselves, forward and reverse sequencing primers were also sent off to Macrogen for sequence confirmation

The remaining 1ml of the overnight cultures were prepared for long term cryogenic storage. 900 µL of the remaining 1ml was mixed with sterile 50% glycerol, and stored in a

cryogenic glycerol stock tube, at -80 °C until further use. After sequence confirmation, only the stocks with the correct sequences were kept, while the incorrect stocks were discarded.

## 2.3 Creation of pULTRA:*metY* variant strains with sfGFP reporter plasmids

Reporter plasmids containing a superfolder Green Fluorescent Protein gene (*sfgfp*) and a carbenicillin resistance gene were obtained from a previous study (Hecht et al., 2017). This set of reporter plasmids consisted of 64 pET20b plasmids with sfGFP reporter proteins that each had a modified start codon. Within the set of 64 reporters, each reporter had one of the 64 possible codon combinations in place of the native start codon.

NEB turbo was selected to harbour the *sfgfp* reporters, due to its favourable DNA expression and growth capabilities (obtained from NEB, catalogue number C29841).

Competent cells of NEB turbo were prepared using the *Mix and Go* Transformation and Competent Cells kit from Zymo Research, as the efficiency of this kit was significantly higher than conventional TSS competent cell preparation and transformation kit. Once transformed, the cells were spread onto on pre-warmed LB agar plates, containing carbenicillin (100 ug / mL), and grown overnight at 37 °C in a shaker at 200 RPM.

Successful transformant colonies were picked and inoculated in 6mL of LB miller with carbenicillin (100 ug / mL), and grown overnight at 37 °C in a shaker at 200 RPM. The following day, the pET20b plasmids were extracted using the Qiagen Plasmid Miniprep kit in preparation for sequence confirmation. A portion of the cells were cryogenically frozen in 50% glycerol and stored at- 80C until further use.

## 2.4 Preparation of plasmids for transformation of anticodon variant pULTRA:*metY* variants and pET20b:*sfgfp* variants into BL21(DE3)pLysS

Once both anticodon pULTRA variants and codon pET20b variants were sequence confirmed, the plasmids were extracted from their harbouring strains in preparation for a double transformation into BL21(DE3)pLysS. The extraction of each plasmid was performed using the Qiagen miniprep. Once extracted, the plasmids were diluted to 1 ng / µL with elution buffer 250 mM Imidazole, 0.3 M sodium chloride, 50 mM sodium phosphate), and kept at 4 °C until further use.

BL21(DE3)pLysS bacterial strains were used for expression of the anticodon plasmids and their reporters as the pLysS plasmids within this strain aids in reducing background

expression of plasmids with the T7 promoter. In the case of this experimental series, the T7 promoter was used on the reporter plasmids, thus using the pLysS plasmid to reduce background expression would enhance the quality of the results obtained for these reporters. Since the Zymo transformation protocol is not efficient at double transformations, a sequential double transformation approach was undertaken.

Firstly, the anticodon pULTRA variants were transformed into competent BL21(DE3)pLysS strains using the Zymo Mix and Go protocol. Transformants were then plated onto large LB miller agar plates with spectinomycin (50 ug / mL) and chloramphenicol (25 ug / mL) antibiotics and grown overnight at 37 °C in a shaker at 200 RPM. Successful transformant colonies were picked, re-inoculated overnight at 37 °C in a shaker at 200 RPM. in 6 µl of LB miller with appropriate antibiotics, and glycerol stocked in the same manner as previously. The majority of the overnight cultures however, were used to make competent cells with the anticodons in them as per the Zymo protocol used previously.

Following this, the competent BL21(DE3)pLysS anticodon strains were transformed with the pET20b start codons. The transformants were plated onto large agar plates with spectinomycin (50 ug / mL), chloramphenicol (25 ug / mL) and carbenicillin (100 ug / mL), and incubated overnight at 37 °C. In total, there was 12 anticodon variants and 64 start codon reporter variants, making for a total of 768 new anticodon-codon strains, all of which were cryogenically frozen in 50% glycerol, and stored at -80 °C until further use.

## 2.5 Bulk florescence studies of codon/anticodon strains

To determine the translation initiation patterns from the mutant i-tRNAs and the 64 start different start codons, Bulk fluorescence was performed as per in Hecht et al. (2017). Mutant i-tRNA strains with different start codon mutant sfGFP reporter plasmids were grown overnight in biological triplicates in spectinomycin (50 ug / mL), chloramphenicol (25 ug / mL) and carbenicillin (100 ug / mL) at 37 °C. These overnight cultures were then diluted 1:100 and grown at 37 °C in a shaker at 200 RPM for another 1 hour. After an hour, each culture was induced with 100mM IPTG, and grown for a further 6 hours at 37 °C in a shaker at 200 RPM. After the incubation, the cells from each culture were spun down and resuspended in 250 µL of PBS (NaCl 0.138 M, KCl - 0.0027 M) in a clear bottom 96 well black well plate. These cultures were left at 4 °C overnight, to let the sfGFP mature.

The bulk fluorescence was then recorded using the PheroStar microplate reader with gain setting 111. The $OD_{600}$ of each sample was also recorded, and was used to normalise the data (Arb. U/$OD_{600}$). The data generated from this was then further normalised so that the maximum and minimum values of each reading were the same between samples using a min/max normalisation method.

## 2.6 Bioinformatic analysis of anticodon-codon interactions

The normalised bulk fluorescence data obtained was further analysed utilising a BLAST based bioinformatic binding assay. First the data was sorted into groups based on the mutant i-tRNA within the cell. Within each of these groups, a numeric value was awarded based on the anticodon-codon interactions predicted within the groups. From the three nucleotide bases, when a complimentary base was found a +1 was awarded, whereas if there was a non-complimentary bases pairing event a -1 was awarded. This left the possibilities of 3 (full complement), 1 (partial complement), (-1 only one base pair similarity) and -3 (no similar base pairing). These predicated binding patterns were then assigned to the top 10 anticodon-codon interactions for each mutant i-tRNA group from the data, and a relative abundance was determined between expected and observed codon binding events.

## 2.7 Fitness analysis of anticodon variants in BL21(DE3)pLysS

To determine whether the introduced anticodon variant pULTRA:*metY*:tac were toxic to the BL21(DE3)pLysS strain, a fitness assay was undertaken using growth analysis. Firstly, transformant BL21(DE3)pLysS strains with the anticodon variant pULTRA:*metY*:tac colonies (lacking the pET20B:*sgfp* reporter plasmids) were streaked, picked in triplicates and grown overnight in a shaker at 200 rpm and at 37 °C. in 2ml LB miller with spectinomycin (50 ug / mL) and chloramphenicol (25 ug / mL). The following day, the liquid cultures were diluted 1:100 into 100 µL of fresh LB miller, spectinomycin (50 ug / mL) and chloramphenicol (25 ug / mL) in a 96 well growth plate. The plate was then inserted the NanoStar spectrophotometer, which was set to measure $OD_{600}$ every 5 minutes, at 37 °C. over a period of 10 hours. This was repeated 2 more times in different conditions: firstly with the anticodon plasmids being induced through the addition of 1mM IPTG and secondly with the anticodon plasmids being repressed with 2% glucose.

The data was exported from the spectrophotometer and analysed using a method developed by Lajoie et al. (2013), where the fitness of the new strains was determined through normalised measurements of max OD and the growth rate for each triplicate. Due to the inherent differences in growth data between technical replicates, the measurements were all normalised against internal an control: BL21(DE3)pLysS cells without the anticodon variants.

# 3 Results

## 3.1 Designing and testing mutant i-tRNAs with altered anticodon regions

### *3.1.1 Designing a compatible anticodon-codon reporter system*

i-tRNAs in *E. coli* are transcribed from the gene *metY*. This gene encodes a methionine bearing i-tRNA with all the vital features of an i-tRNA, including a highly conserved acceptor stem region responsible for IF2 affinity, GC base pairing and non-Watson-Crick pairing in the anticodon loop, responsible for P site binding, and an anticodon region complimentary to the start codon on the mRNA transcript. To build mutant i-tRNAs, the anticodon regions of the gene were altered, preserving the other important aspects of the *metY* gene ensuring that the tRNA can still function as an i-tRNA when transcribed. The gene was assembled in the pULTRA plasmid, containing a CloDF13 origin of replication and a spectinomycin resistance gene (**Fig 6.).** A *tac* promoter was also used upstream of the *metY* gene, to allow suppression and induction of the gene throughout the experimental series. The choice of plasmid to harbour the *metY* gene was due to the compatibility of its origin of replication with the designed reporter plasmids from Hecht et al. (2017).

The reporter plasmids used within this study were obtained from a study conducted by Hecht et al. (2017). These plasmids contained a superfolder Green Fluorescent Protein (sfGFP) reporters, which mature more quickly and are more stable than their parent protein GFP. This reporter was harboured on a pET20b plasmid, with a pBR322 origin of replication and ampicillin resistance (**Fig 7.**). There were 64 variants of this reporter plasmid, each with a different start codon sequence on the sfGFP reporters, all prefixed with a standard T7 promoter sequence.

The two plasmids pULTRA:tac:*metY*, and pET20b:T7:*sfgfp* were transferred into one host strain for fluorescence measurements of this reporter system. The BL21(DE3)pLysS strain, to ensure efficient expression of the T7 controlled reporter gene and reduce background expression of this reporter through the pLysS. The pLysS plasmid has a p15A origin and chloramphenicol resistance, making it compatible with the other two plasmids within this study **(Fig. 7).**

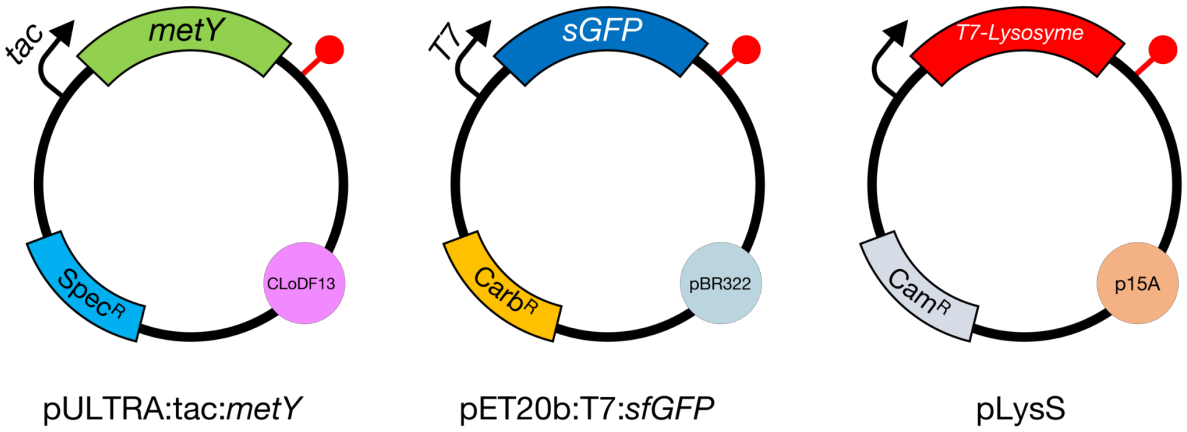



**Figure 7. The three plasmids used within the anticodon-codon reporter system.** pULtra:*metY*:tac (left) harbours the *metY* gene which encodes for the mutant i-tRNAs with altered anticodon regions, spectinomycin resistance gene and a CloDF13 origin of replication. pET20b:sfGFP (middle) harbours the fluorescent reporter gene *sfgfp* with altered start codons, a carbenicillin resistance gene, and a pBR322 origin of replication. pLysS plasmid (right) harbouring a T7 lysozyme gene, a chloramphenicol resistance gene and a p15A origin of replication.

## 3.2 Measuring anticodon mutant initiator tRNA translation initiation efficiency across all possible start codons

The efficacy of the anticodon-codon reporter system with the three different plasmids **(Fig 7.)** was determined through the use of bulk fluorescence measurements. This method involved expressing the transcription of mutant i-tRNAs from the *metY* gene via IPTG induction, and measuring the fluorescence of the sfGFP proteins produced in response to the IPTG induction. This method was repeated for each *metY* anticodon variant versus each start codon sfGFP reporter variant.

### *3.2.1 Anticodon variants elicited varied translation initiation across sfGFP variants*

Bulk fluorescence measurements were undertaken for strains containing different mutant anticodon *metY* plasmids-variant codon sfGFP reporter plasmids. Measurements were taken in biological triplicates or duplicates (**s. Fig. 1**). Normalised data was then compiled to determine any patterns in differences between sfGFP expression between the strains (**Fig. 8**). It can be seen from these data, that there is some variation between the translation of sfGFP between different mutant i-tRNA strains. As a general trend however, it can be seen that the top translation was achieved were from the codons AUG, GUG and UUG, indicating that natural translation initiation was still dominant within these reporter systems, despite the expression of mutant i-tRNAs.

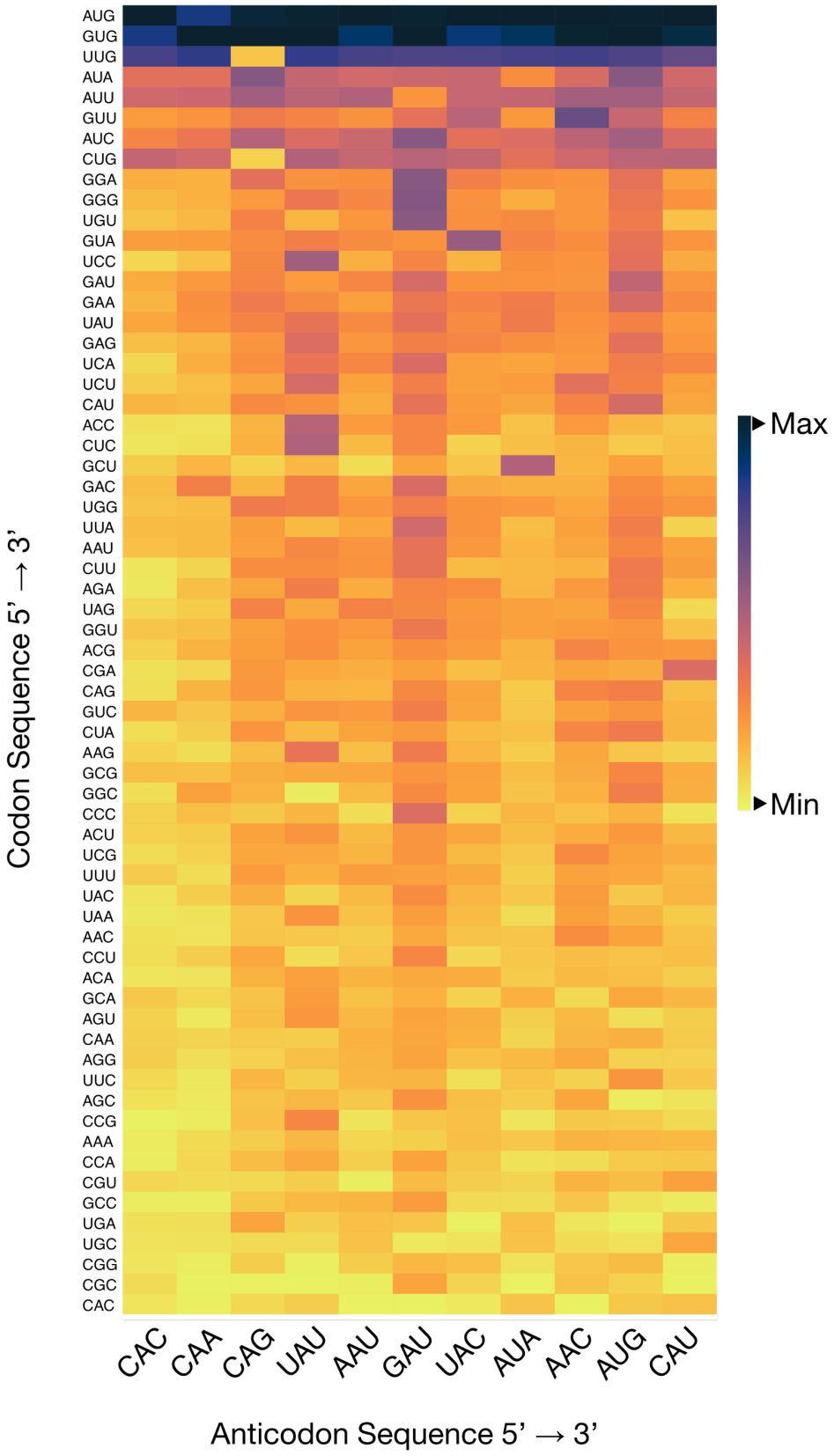



**Figure 8. The expression of non-canonical start codons from mutant i-tRNAs results in highly varied translation initiation.** Normalised bulk fluorescence data (Arb. U/$OD_{600}$) of the interaction between induced mutant i-tRNAs and 64 start codon mutant sfGFP. High variation of translation initiation can be seen across all of the mutant i-tRNA groups. The highest average expressing codons across all mutant i-tRNA groups include AUG, GUG and UUG.

### *3.2.2 Anticodon-codon compliments show varied translation initiation events*

Codon-anticodon interactions between mutant i-tRNAs and their cognate start codons were variable between different anticodon groups (**Fig.9**). When comparing anticodon-codon pairs, the strongest translation initiation of *sfGFP* was found between the anticodons CAU, CAC and CAA and their respective complimentary codon sequences. Conversely, the anticodons AUG, AUA and CAG and their anticodon compliments had the lowest rates of translation initiation of the reporter protein (**Fig. 9.**).

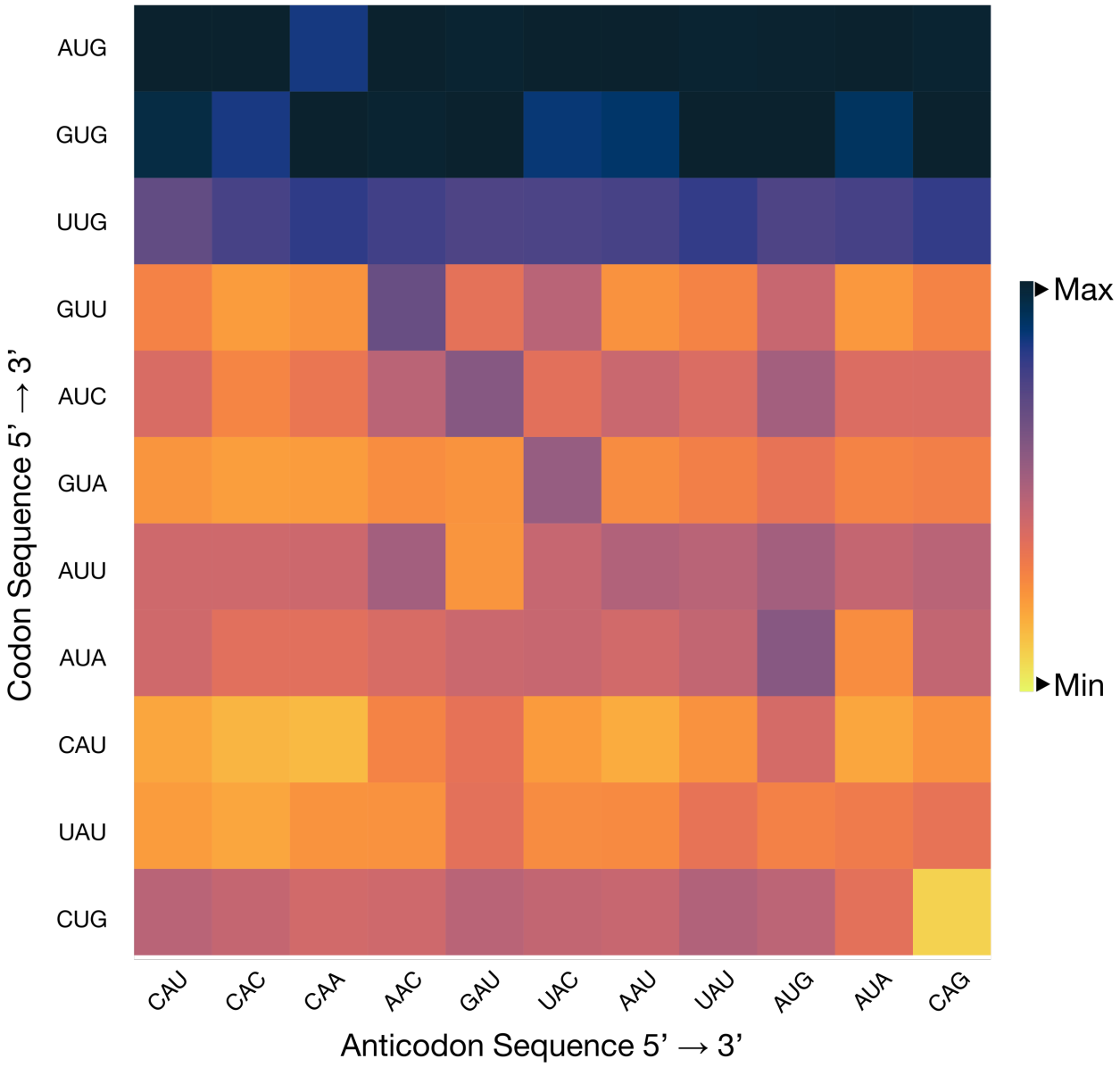



**Figure 9. Most of the mutant i-tRNA groups show translation initiation from their complimentary start codons.** Normalised bulk fluorescence data (Arb. U/$OD_{600}$) of the interaction between induced mutant i-tRNAs and the mutant sfGFP reporters with complimentary start codons. The mutant i-tRNA groups are ranked highest from right to left, with the highest expression from complimentary start codon achieved to the lowest expression achieved from complimentary start codon. The highest expression was observed in the mutant i-tRNAs CAU, CAC and CAA and its complimentary start codons, whereas weak expression can be seen between the mutant i-tRNA CAG and its CUG start codon compliment.

To better examine changes in protein expression due to i-tRNA anticodon and mRNA codon interaction, anticodon-codon pairs were ranked from highest to lowest expressing for each anticodon variant group. The ranks between individual anticodon groups were then compared to the native CAU anticodon group to determine any fluctuation between complimentary anticodon-codon pairing. The aim of this analysis was to determine if the expression of a mutant i-tRNA would alter protein expression of its cognate start codon.

There was considerable variation observed between the different anticodon-groups results, as it was found that some anticodon-codon complimentary pairs did not alter expression when compared to the CAU control group (**Fig. 10.),** but some anticodon groups showed considerable changes in protein expression from their cognate start codons **(Fig. 11, and Fig. 12).**

The anticodons which did not cause any significant changes in the expression from their cognate codons between their anticodon group and the control CAU group were AAU, GAU, CAC and CAA **(Fig. 10).** From this subset, there was no observable difference in cognate codon protein expression between the anticodons CAC and CAA, while anticodons CAU and GAU had a slight increase when compared with the control.

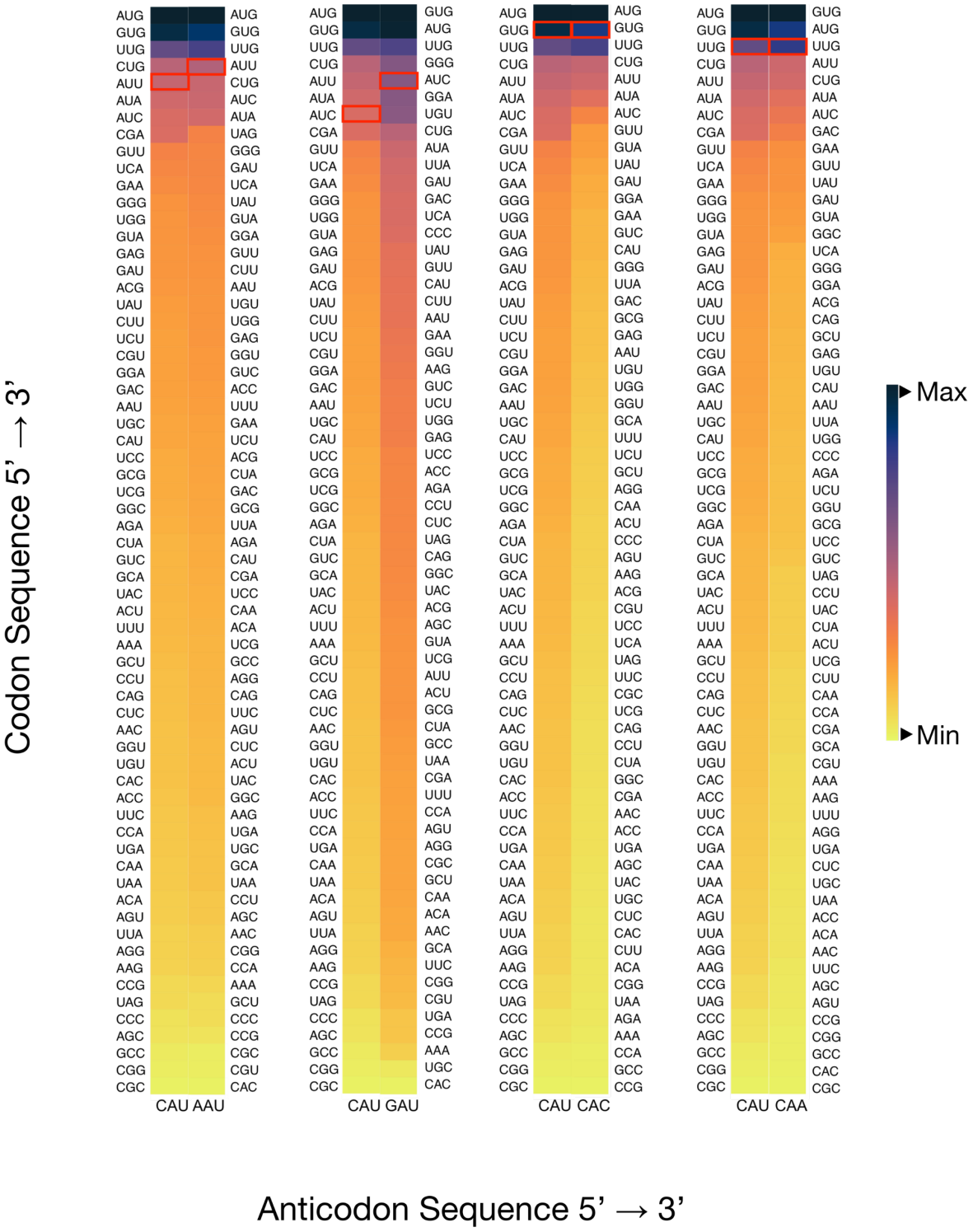



**Figure 10. Mutant i-tRNAs and their complimentary start codons with the little observed increase in translation initiation between anticodon-codon binding.** Normalised bulk fluorescence data (Arb. U/$OD_{600}$) of the interaction between induced mutant i-tRNAs and the mutant sfGFP reporters with complimentary start codons. All of the start codons for each mutant i-tRNA are ranked from highest expressing to lowest expressing. Each mutant i-tRNA group is compared against the native i-tRNA (CAU anticodon). Within this subset, the complimentary start codons to each mutant i-tRNA (boxed in red) show little to no variation between the mutant i-tRNA control group and the mutant i-tRNA group.

Conversely, the anticodons AUA, UAC, AUG and AAC showed an increase in expression from their complimentary start codons when compared to the CAU control, with AUG having the most pronounced expression increase (**Fig. 11**).

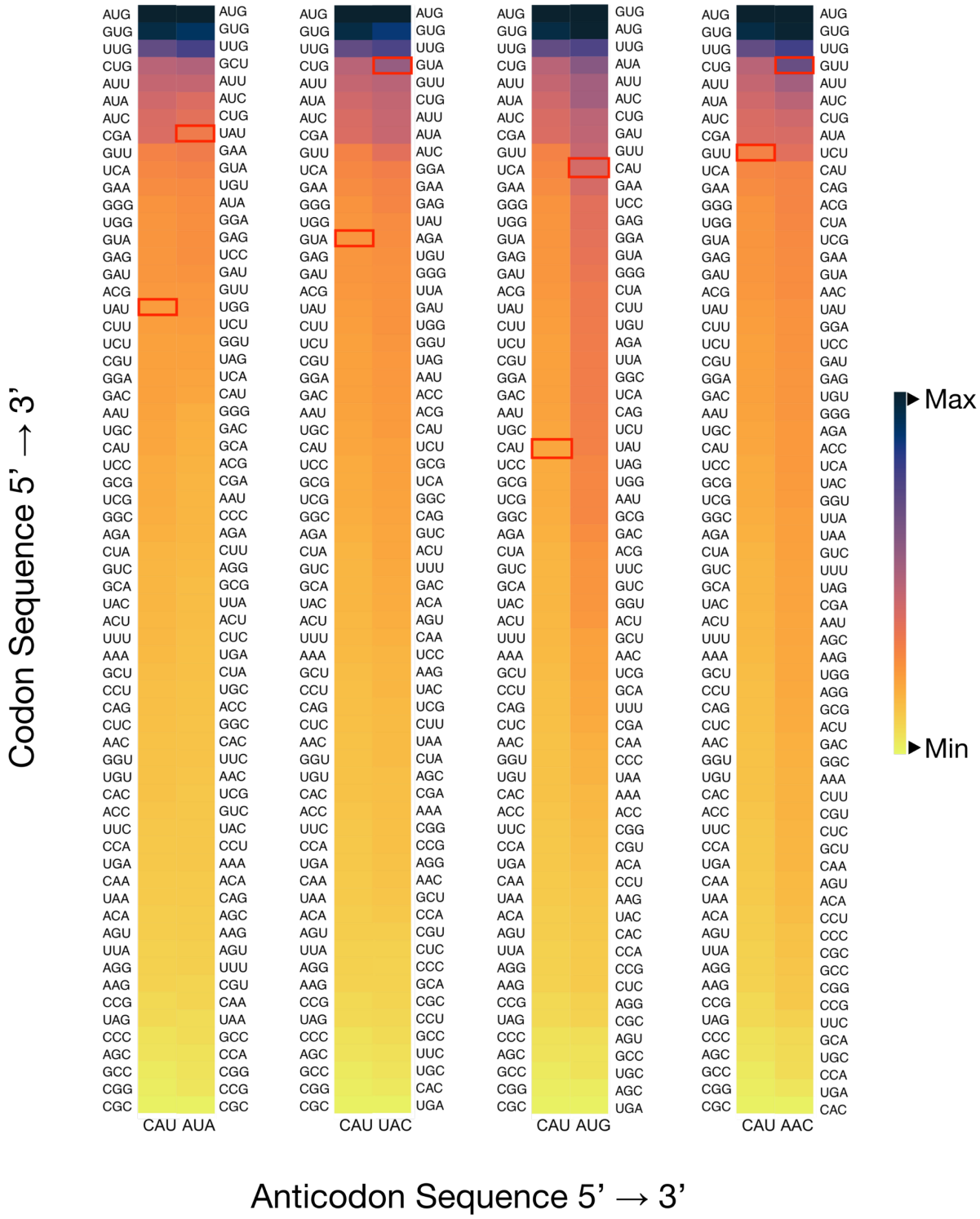



**Figure 11. Mutant i-tRNAs and their complimentary start codons with the most observed increase in translation initiation between anticodon-codon binding.** Normalised bulk fluorescence data (Arb. U/$OD_{600}$) of the interaction between induced mutant i-tRNAs and the mutant sfGFP reporters with complimentary start codons. All of the start codons for each mutant i-tRNA are ranked from highest expressing to lowest expressing. Each mutant i-tRNA group is compared against the native i-tRNA (CAU anticodon). Within this subset, the complimentary start codons to each mutant i-tRNA (boxed in red) show a considerable increase in expression between the mutant i-tRNA control group and the mutant i-tRNA group.

Finally, the anticodons UAU and CAG caused a decrease in protein expression from their cognate start codons, when compared to the CAU control group (**Fig 12.**). Between these two anticodons, UAU had the least effect on its cognate codon, as the expression between the UAU anticodon group and the control group were slight. In contrast, the CAG anticodon caused a significant decrease in expression from its cognate codon. The compliment of CAG, CUG is the 4$^{th}$ highest expressing codon in the CAU control group. In the CAG anticodon group, this codon is reduced to the 7th worst expressing codon.

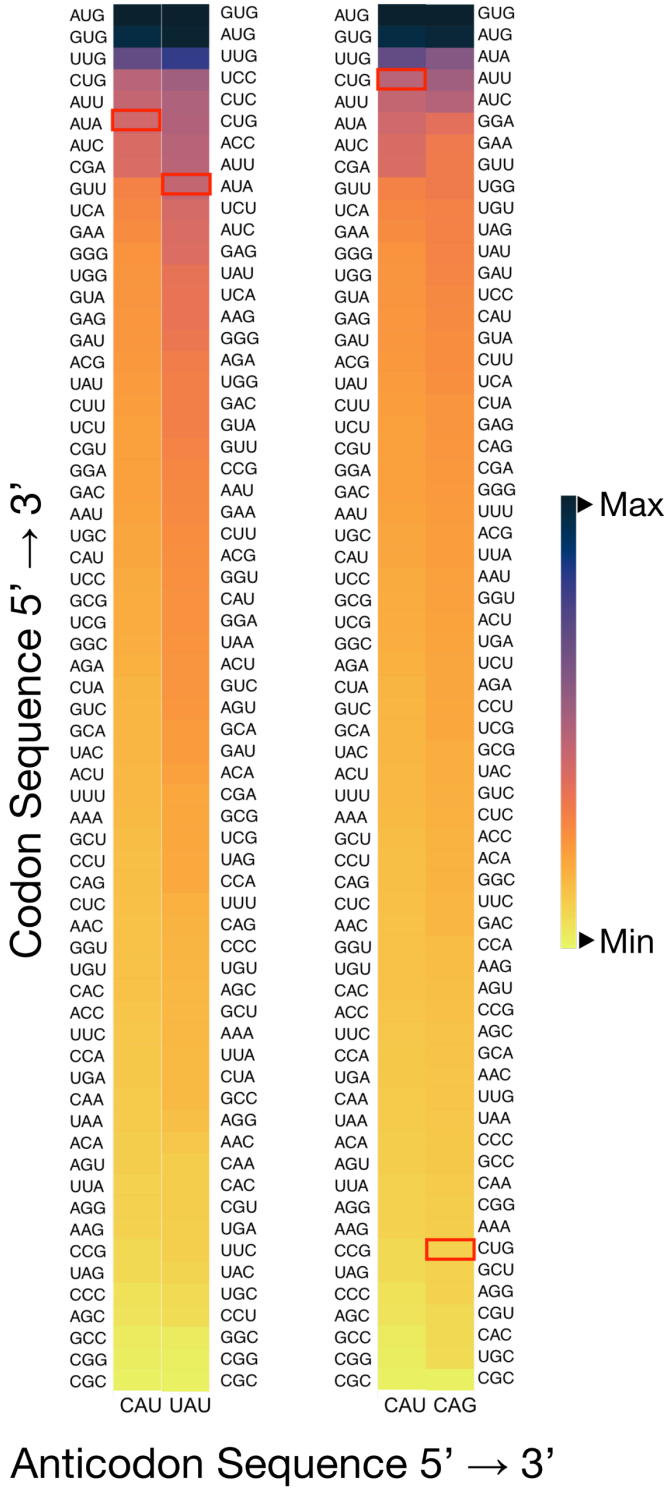



**Figure 12. Mutant i-tRNAs and their complimentary start codons with the highest observed decrease in translation initiation between anticodon-codon binding.** Normalised bulk fluorescence data (Arb.U/$OD_{600}$) of the interaction between induced mutant i-tRNAs and the mutant sfGFP reporters with complimentary start codons. All of the start codons for each mutant i-tRNA are ranked from highest expressing to lowest expressing. Each mutant i-tRNA group is compared against the native i-tRNA (CAU anticodon). Within this subset, the complimentary start codons to each mutant i-tRNA (boxed in red) show a considerable decrease in expression between the mutant i-tRNA control group and the mutant i-tRNA group. This decrease is profound in the CAG mutant i-tRNA, with expression from its complimentary codon CUG decreasing from one of the highest expressing codons in the control i-tRNA CAU group to one of the lowest expressing codons in the CAG mutant i-tRNA group.

### *3.2.3 Determining non-complimentary binding between anticodons and near cognate start codons.*

The results from the bulk fluorescence were compared with bioinformatic data generated through a simple codon binding model to understand the interactions between anticodon-codon pairs. The method involved awarding numbers based on the type of bonds formed between nucleotides in the anticodon-codon pairs. Nucleotides which lined up with their compliment (A-U, U-A, C-G, G-C) were awarded a 1. Conversely, nucleotide pairs which did not have Watson-Crick type binding were awarded a -1. The three nucleotide pair scores for each anticodon-codon interaction were then summed. These scores were then compared with the top 10 highest expressing anticodon-codon pairs for each mutant i-tRNA group (**Table 1.**). Furthermore, the percentage of matches between the top 10 predicted codons vs the actual data was also portrayed for mutant i-tRNA groups (**Table 1.)** From these data, it is evident that for each mutant i-tRNA group, there is an abundance of initiation events occurring from non-complimentary base pairs.

**Table 1. Top expressing codons matched with a BLAST based binding scores show a high abundance of non-Watson-Crick base pairing between mutant i-tRNAs and alternative start codons.**

| CAU | | | CAC | | | CAA | | | AAC | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Start Codon sequence 5' -> 3' | Actual Rank from data | Binding score based on binding test | Start Codon sequence 5' -> 3' | Actual Rank from data | Binding score based on binding test | Start Codon sequence 5' -> 3' | Actual Rank from data | Binding score based on binding test | Start Codon sequence 5' -> 3' | Actual Rank from data | Binding score based on binding test |
| AUG | 1 | 3 | AUG | 1 | 1 | GUG | 1 | 1 | GUG | 1 | 1 |
| GUG | 2 | 1 | GUG | 2 | 3 | AUG | 2 | 1 | AUG | 2 | 1 |
| UUG | 3 | -1 | UUG | 3 | 1 | UUG | 3 | 3 | UUG | 3 | 3 |
| CUG | 4 | 1 | CUG | 4 | 1 | AUU | 4 | -1 | AUU | 4 | -1 |
| AUU | 5 | -1 | AUU | 5 | -1 | CUG | 5 | 1 | CUG | 5 | 1 |
| AUA | 6 | 1 | AUA | 6 | -1 | AUA | 6 | -1 | AUA | 6 | -1 |
| AUC | 7 | 1 | AUC | 7 | -1 | AUC | 7 | -1 | AUC | 7 | -1 |
| CGA | 8 | 1 | GUU | 8 | 1 | GAC | 8 | -3 | GAC | 8 | -3 |
| GUU | 9 | -1 | GUA | 9 | 1 | GAA | 9 | -3 | GAA | 9 | -3 |
| UCA | 10 | -3 | UAU | 10 | -3 | GUU | 10 | -1 | GUU | 10 | -1 |
| **GAU** | | | **UAC** | | | **AAU** | | | **UAU** | | |
| Start Codon sequence 5' -> 3' | Actual Rank from data | Binding score based on binding test | Start Codon sequence 5' -> 3' | Actual Rank from data | Binding score based on binding test | Start Codon sequence 5' -> 3' | Actual Rank from data | Binding score based on binding test | Start Codon sequence 5' -> 3' | Actual Rank from data | Binding score based on binding test |
| AUG | 1 | -1 | GUG | 1 | 1 | AUG | 1 | 1 | AUG | 1 | 1 |
| GUG | 2 | 1 | AUG | 2 | -3 | GUG | 2 | -1 | GUG | 2 | -1 |
| UUG | 3 | -1 | UUG | 3 | -1 | UUG | 3 | -3 | UUG | 3 | 1 |
| GUU | 4 | 3 | GGG | 4 | -3 | GUA | 4 | 1 | AUU | 4 | 1 |
| AUU | 5 | 1 | AUC | 5 | 1 | GUU | 5 | 1 | CUG | 5 | 1 |
| AUC | 6 | -1 | GGA | 6 | 1 | CUG | 6 | -1 | AUC | 6 | 1 |
| CUG | 7 | -1 | UGU | 7 | -1 | AUU | 7 | 1 | AUA | 7 | 1 |
| AUA | 8 | -1 | CUG | 8 | -3 | AUA | 8 | -3 | UAG | 8 | -3 |
| UCU | 9 | -1 | AUA | 9 | -1 | AUC | 9 | -1 | GGG | 9 | -3 |
| CAU | 10 | -1 | UUA | 10 | -1 | GGA | 10 | 1 | GAU | 10 | -1 |
| **AUG** | | | **AUA** | | | **CAG** | | | | | |
| Start Codon sequence 5' -> 3' | Actual Rank from data | Binding score based on binding test | Start Codon sequence 5' -> 3' | Actual Rank from data | Binding score based on binding test | Start Codon sequence 5' -> 3' | Actual Rank from data | Binding score based on binding test | | | |
| GUG | 1 | -1 | GUG | 1 | -3 | GUG | 1 | 1 | | | |
| AUG | 2 | 1 | AUG | 2 | -3 | AUG | 2 | 1 | | | |
| UUG | 3 | -1 | UUG | 3 | -3 | AUA | 3 | -1 | | | |
| UCC | 4 | -3 | AUA | 4 | -3 | AUU | 4 | -1 | | | |
| CUC | 5 | -1 | AUU | 5 | -1 | AUC | 5 | -1 | | | |
| CUG | 6 | -1 | AUC | 6 | -3 | GGA | 6 | -3 | | | |
| ACC | 7 | -1 | CUG | 7 | -1 | GAA | 7 | -3 | | | |
| AUU | 8 | 1 | GAU | 8 | 1 | GUU | 8 | -1 | | | |
| AUA | 9 | 3 | GUU | 9 | -1 | UGG | 9 | 1 | | | |
| UCU | 10 | -3 | CAU | 10 | 3 | UGU | 10 | -3 | | | |

| % of matches between top 10 predicted codons vs actual codons | |
|---|---|
| **Anticodon Sequence 5' -> 3'** | **% of matching predicted codons** |
| **CAU** | **66** |
| **CAC** | **60** |
| **CAA** | **40** |
| **AAC** | **40** |
| **GAU** | **30** |
| **UAC** | **30** |
| **AAU** | **50** |
| **UAU** | **60** |
| **AUG** | **30** |
| **AUA** | **20** |
| **CAG** | **30** |

## 3.3 Fitness Effects of mutant i-tRNAs on host

To determine any fitness effects of the mutant i-tRNAs on the host *E. coli* cells, cultures of BL21(DE3)pLysS cell harbouring mutant i-tRNAs with altered anticodons were grown. To measure growth rate and maximum optical density characteristics, the were

grown under under three separate conditions with different effects on the anticodon mutant i-tRNA expression: (1) LB broth-only, (2) LB broth with 2% glucose repression, (3) and LB broth with 1 mM IPTG induction.

From the baseline growth experiment with LB-only where it was expected that the anticodon mutant i-tRNAs were only created through leaky expression, it was found that most anticodon mutant tRNAs did not affect host cell growth rate and showed only small changes in maximal optical density (**Fig. 13A**). When induced with IPTG, the strains showed a decrease in maximum optical density and the growth rate from a standard control (a BL21(DE3)pLysS strain without any mutant i-tRNAs grown in the presence of chloramphenicol). The native CAU anticodon i-tRNA-containing strain was the only one to show a large growth rate increase under inducing conditions (**Fig. 13C**). Under repression conditions the growth rate and maximal optical density effects from the i-tRNAs was similar to the LB-only control showing only minor alterations from (**Fig. 13B**).

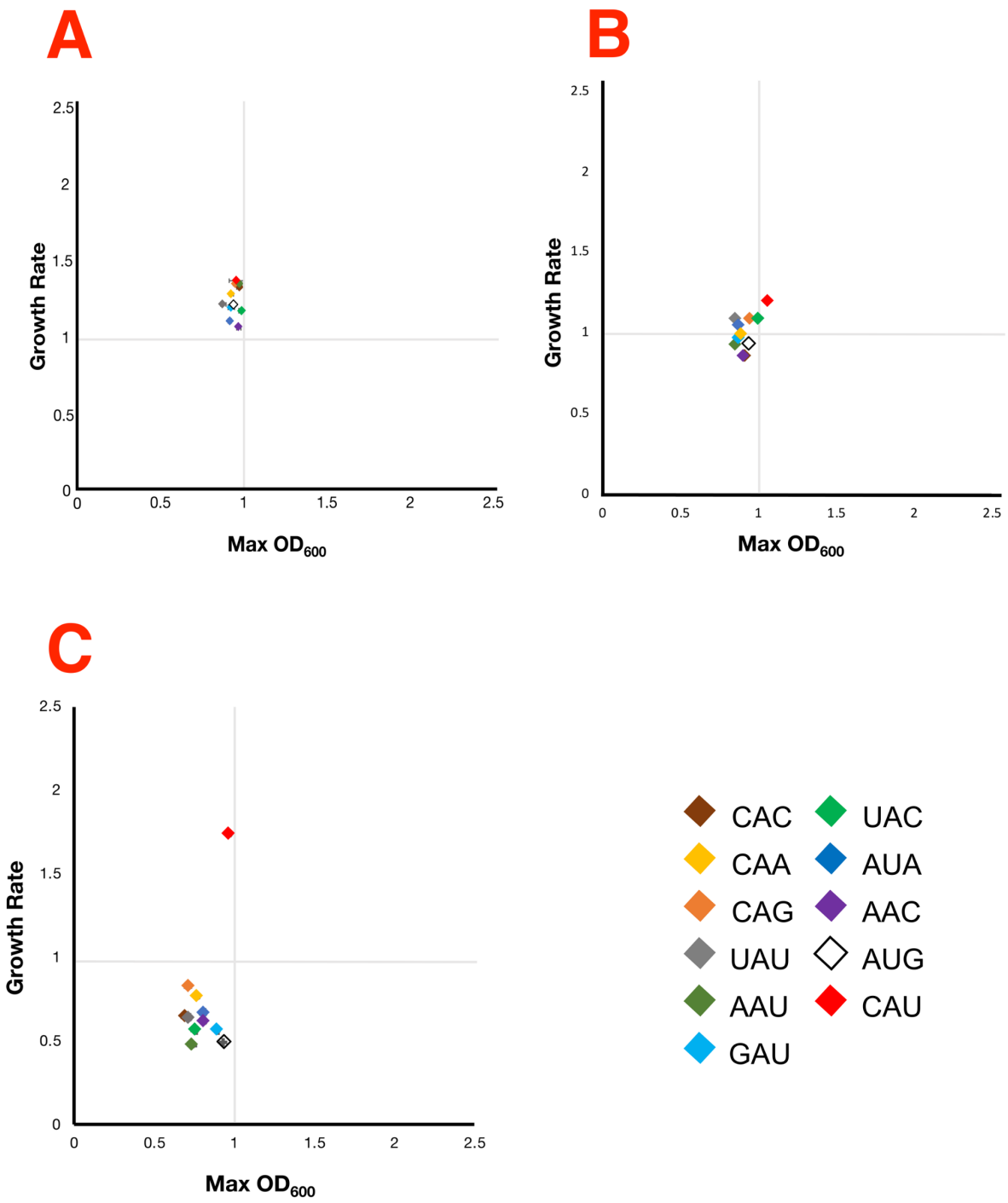



**Figure 13. The effect of induced mutant i-tRNAs on BL21(DE3)pLysS fitness.** Ratios for normalised final maximum $OD_{600}$ (x-axis) and normalised growth rates (y-axis), in three different conditions: **(A)** Baseline LB only growth, **(B)** Repressed growth with 2% glucose and **(C)** induced expression with 1mM IPTG. Growth defects, in mutant i-tRNA strains can be seen in the induced condition, with particular decrease in growth rate as each cell culture is below the grey line on the Y-axis. Induced expression of native i-tRNA resulted in increased growth rate of cell cultures, considerably moving them above the grey line on the Y-axis.

# 4 Discussion

The development of orthogonal translation systems in prokaryotes shows great promise in bringing synthetic biology towards two of its main goals: (1) having high-levels of predictability, and (2) total control of non-natural genetic systems (Liu et al., 2018). The current literature outlines a large and growing body of work testing partially orthogonal systems for DNA replication (Ravikumar et al., 2014), transcription (Tabor and Richardson, 1985, Studier et al., 1990, Studier and Moffatt, 1986, Perez-Pinera et al., 2013) and translation (Wang et al., 2007, Lajoie et al., 2013, Vincent, 2017, Neumann et al., 2010).

In this thesis I designed, constructed, and tested a new set of anticodon mutant i-tRNAs in order to develop an orthogonal translation initiation system in *E. coli*. Using fluorescent reporters representing all 64 possible start codons I measured the mutant anticodon i-tRNA translation initiation specificity. Finally, I measured fitness effects for each anticodon mutant tRNA to determine its feasibility for use in an orthogonal translation initiation system.

## 4.1 High variability in anticodon-codon fidelity in translation initiation between anticodon variants

Within this thesis project 9 new anticodon mutant i-tRNAs were constructed. The translation initiation efficiencies of these i-tRNAs were measured using reporter proteins representing all 64 possible start codons. The bulk fluorescence intensity measurements between anticodon-codon pairs showed high variability in fidelity of mutant i-tRNA to initiation translation from their cognate start codon (**Fig 9**.). I also observed a high frequency of translation initiation from non-cognate start codons throughout the dataset (**Fig 8., Table 1.)**, similar to that observed for the wild-type i-tRNA which strongly initiates at the single-mismatched GUG and UUG start codons (Hecht et al., 2017).

4 out of 11 anticodon groups (AAU, GAU, CAC and CAA) showed no measurable changes in protein expression from their cognate codons (**Fig. 10**). This result was very surprising but may be explained through the sequence interaction between complimentary start codons to each anticodon. The cognate start codons for i-tRNAs with anticodons AAU, GAU, CAC and CAA are the naturally strong start codons AUU, AUC, GUG and UUG, respectively. The sfGFP reporters used in this study were chosen for their comprehensive

coverage of all 64 possible start codons, but their dynamic range is perhaps not sufficient to reliably detect the small difference in expression due to the introduction of a relatively inefficient mutant i-tRNA into the cell compared to the large background translation initiation rate from wild-type i-tRNA.

In contrast, increased expression compared to the CAU anticodon control group was detected, from the cognate start codons of anticodon mutant i-tRNA anticodons AUA, UAC, AUG and AAC **(Fig. 11).** Out of this subset, the more pronounced increase in expression was found between the anticodon AUG and its cognate start codon CAU **(Fig. 10).** This was easily detected with the sfGFP reporters because the background fluorescence of CAU start codon from wild-type i-tRNA is so low. Finally, the anticodon groups with decreased protein expression from their cognate start codons included UAU and CAG, with CAG having marginally worse expression between the anticodon group and the control group. The mutant i-tRNA CAG in particular, shows very interesting results, as the decrease in expression of its cognate start codon between the control i-tRNA CAU and its anticodon groups is pronounced. Expression from the start codon CUG in the control is shown to be the $4^{th}$ highest in the subset, however, in the CAG mutant i-tRNA group, expression from the CUG start codon drops to the one of the lowest expressed in the subset, indicating that the anticodon-codon interactions between CAG-CUG may be causing some undesirable effects to translation initiation.

The observed differences in translation efficiency between the different anticodon-codon pairs may be explained by different factors involving RNA modifications to the tRNA during maturation and the affinity of mutant i-tRNAs to the ribosome assembly machinery such as aaRSs, MTF and IF2.

### *4.1.1 A modification at base A37 in anticodon variants UAU and CAA may stabilise codon-anticodon interactions*

tRNAs undergo a multitude of post transcriptional modifications to different nucleotides, many of which may greatly affect the function of the tRNA. A modification at base A37 of the amber i-tRNA and both Arg and Glu e-tRNAs has recently been shown to play vital roles in A-U bond stabilisation in the first position between anticodon-codon interaction (Schweizer et al., 2017). MiaA/B enzymes have shown to catalyse the modification of A37 to 2-methylthio-N-$^6$($\Delta^2$-isopentenyl)adenosine($ms^2i^6A$). This modification enhances the A-U pairing through the introduction of stacking interactions

between the modified A base and the U base on the codon. The recognition for MiaA/B enzymes has been shown to be the bases A36, A37 and A38, located on the anticodon-stem loop of the tRNA molecule (**Fig 14.**).

In contrast, the wild-type i-tRNA does not possess the A37 modification. Instead, the A37 position of the wild-type i-tRNA interacts with the base pairs G29 and C41 to stabilise the anticodon loop (**Fig 14.**)(Zhou et al., 2015). This $ms^2i^6A$ at base 37 may exist on some of the mutant i-tRNAs within this study, as this bind motif of A36, A37 and A38 is present on some mutant i-tRNAs such as AUA and CAA. If so, this modification may have contributed to the increased translation initiation expression of these mutant i-tRNAs, through the stabilisation of the weak A-U bonds found within the third position of the anticodon-codon pair.

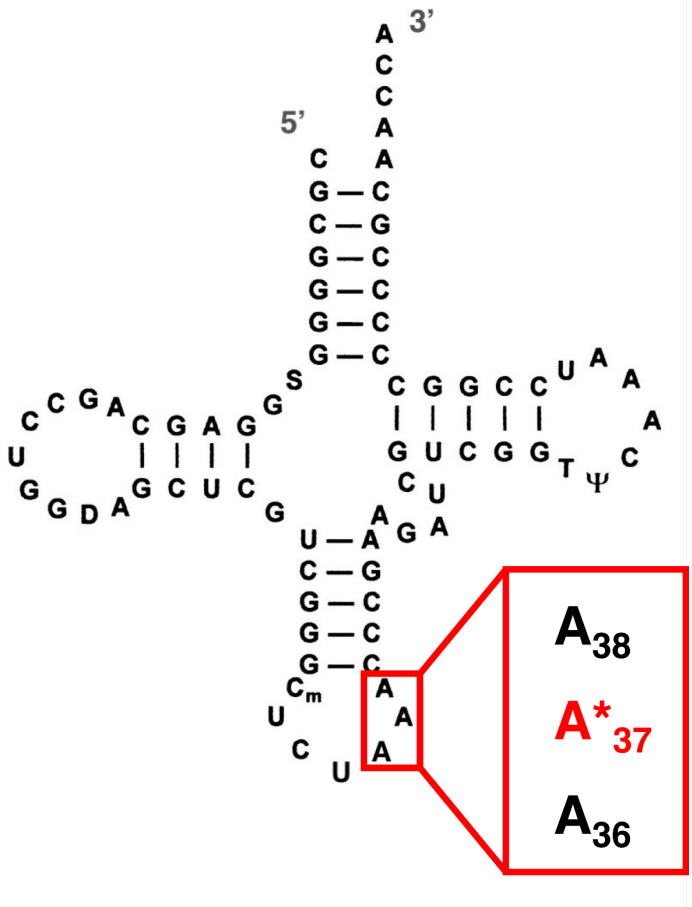



**Figure 14. The A36, A37, A38 Binding motif.** In amber i-tRNAs and Arg and Glu e-tRNAs this binding motif is recognised by Mia/B enzymes, which are responsible for A37 modification. This modification strengthens the weak U-U bond on the third position of the anticodon-codon interaction.

### *4.1.2 Frameshift errors due to modified tRNA bases may reduce apparent translation initiation from sfGFP reporters*

The degeneracy of the genetic code allows for multiple codons to specify the same amino acid (Lagerkvist, 1978). Codons coding for the same amino acid are also found in different abundance. Codon degeneracy allows for the regulation of translational speed, as the ribosome will slow down, stall and often pause in response to the codon on the mRNA transcript. This has evolutionary benefits as it allows for proper protein folding, and in the case of rare amino acids, allows for aminoacyl e-tRNA migration to the ribosome (Agris et al., 2007). The presence and absence of certain tRNA modification has shown to cause ribosome pausing and frameshifts to occur during the translation process. Some of these modifications occur within the anticodon of e-tRNAs. For instance, in the Tyr, His, Asn and Asp e-tRNAs the wobble-base is modified to Queuosine (Q34). If the nucleotide within third position of the anticodon is not modified and transformed into Quenosine34, then it has been observed that +1 frameshifting may occur (Urbonavičius et al., 2001). Similarly, in Glu, Gln, Lys and Leu e-tRNAs, the 34$^{th}$ position is modified to mnm$^5$s$^2$U34. The absence of this modification at the 34$^{th}$ position has shown to also promote +1 frameshifting (Chen et al., 2005).

Even though this modification has not been observed in i-tRNAs, there have been instances of nucleotide modifications occurring in mutant i-tRNAs, which were initially thought to be exclusive to e-tRNAs (Zhou et al., 2015, Schweizer et al., 2017). These modifications were also shown play important roles in ensuring proper mutant i-tRNA function and fidelity (Vincent, 2017, Chen et al., 2005). With this in mind, it may be reasonable to suggest that a similar modification is necessary in modified i-tRNAs with anticodons resembling the e-tRNAs, Tyr, His, Asn and Asp, and thus poor function observed in the mutant i-tRNA CAG may be explained by improper modification **(Fig 11.).**

### *4.1.3 The overproduction of aaRS, MTF and IF2 may be useful in enhancing translation initiation efficiency*

Aminoacylation and formylation of the i-tRNAs has been described as a vital step for IF2 mediated migration of the i-tRNAs to the start site during initiation translation in prokaryotes (Wu and RajBhandary, 1997, Wu et al., 1996). Within this current study, the aminoacylation of the anticodon mutant i-tRNAs was not determined. However, it has

previously been observed that unlike the wild-type i-tRNAs, mutant i-tRNAs are not aminoacylated with Met, but rather by their cognate amino acids dictated by their anticodon regions (Varshney and RajBhandary, 1990, Schulman and Pelka, 1985, Wu and RajBhandary, 1997). For example, it was found that the mutant i-tRNAs with altered anticodon regions to GAC were aminoacylated with Val by ValRS (Wu and RajBhandary, 1997). These changes in aminoacylation from the wild-type i-tRNAs due to changes in anticodon sequence may explain some of the observed weak translation initiation events between cognate anticodon-start codon pairs in this study due to aaRS availability and aminoacylation efficiency.

Previous experiments have shown that the aminoacylation of i-tRNAs is a rate limiting step in translation initiation with mutant i-tRNAs. This is thought to be due to the availability of aminoacyl-RNA synthetases (aaRSs) in the cell, due to codon usage (Sharp and Li, 1987, Chiapello et al., 1999). The over-expression of the aaRSs involved in the aminoacylation of the mutant tRNAs within this study may reveal if aaRS is a limiting factor to translation initiation (Mangroo and RajBhandary, 1995).

The formylation of the aminoacyl-tRNA has also been shown to limit the rate of translation initiation from mutant i-tRNAs (Mayer et al., 2003). MTF is the enzyme responsible for formylating wild-type i-tRNAs, and has also shown to formylate mutant i-tRNAs *in vivo* (Mayer et al., 2003). It was shown that the binding pocket of MTF allows for the binding of non-methionine amino acids, but the catalytic nature of the enzyme is greater depend on the size and charge of the amino acid (Mayer et al., 2003). Previous studies have shown that MTF has a formylation preference for non-methionine amino acids in the preference of Gln ~ Phe > Ile ~ Val > Lys (Mayer et al., 2003). Assuming that the i-tRNAs are charged with their cognate amino acids, between the 10 anticodons tested (excluding the native CAU initiator) the amino acids His, Tyr, Val, Ile, and Leu would be charged on the anticodon mutant i-tRNAs. The most efficient translation initiation events between cognate anticodon-start codons occurred from His, Tyr and Val. Furthermore, the unchanged translation initiation and reduced translation initiation anticodons were presumed to be charged a combination of Ile, Val and Leu. Thus, from this study, there is no distinct correlation between predicted anticodon mutant i-tRNA charged amino acid properties and translation initiation efficiency. Further experiments to identify the amino acid charged to each of the anticodon mutants would shed further light on this correlation. Additionally, MTF supplementation into cells expressing the anticodon mutant i-tRNAs may

resolve potential bottlenecks in the formylation of mutant i-tRNAs, as shown previously (Mayer et al., 2003).

IF2 is believed to bind and carry charged tRNAs to the ribosome during translation initiation, however the recognition of i-tRNAs by IF2 still remains highly under characterised. It has previously been described that i-tRNAs require a highly conserved 5'-CAACCA-3' region on their acceptor stems, in combination with a formyl-aminoacylation for IF2 binding (Simonetti et al., 2008, Mayer et al., 2003, Guenneugues et al., 2000). It may be reasonable to suggest that due to being charged by non-methionine amino acids, and thus reduced formylation efficiency, some proportion of the mutant i-tRNAs within this study are incapable of interacting with IF2 and being carried to the ribosome to initiate translation. This would explain the lack of difference in expression levels seen between the anticodon-codon pairs AAU-AUU, GAU-AUC, CAC-GUG and CAA-UUG, within their anticodon groups when compared to translation initiation from the canonical CAU initiation tRNA. It is likely that translation initiation may be occurring due the wild-type i-tRNAs.

It is possible that a combination of aaRS availability, formylation by MTF and IF2 affinity to mutant i-tRNAs may be responsible for the variability of translation initiation through mutant tRNAs seen within this study. Future experimentation should consider testing these concepts through the supplementation and over expression of some of the genes responsible for aaRS, MTF and IF2 production.

## 4.2 Fitness analysis shows growth effects of mutant and native tRNA expression

Within this study, a growth experiment was conducted to analyse the effect of the mutant and native tRNA expression in the *E. coli* strain BL21(DE3)pLysS. The experiment was repeated three different times under three different conditions. Firstly, cells containing the mutant initiation tRNAs and a plasmid with the native i-tRNA were grown in their appropriate antibiotics and analysed in a plate reader over a time period of 10 hours (**Fig. 13A**). Following this, the same experiment was conducted, but rather this time the i-tRNA plasmids were induced through the addition of IPTG (**Fig. 13C**). Finally, a third repetition of the experiment involved the addition of glucose to supress any tRNA expression (**Fig. 13B**). These three conditions were then analysed, whereby of growth rate and OD were determined and normalised against an internal control. It was found that in both natural

and repressed conditions, there were not notable increases in fitness relating to final OD or growth rate between any of the mutant initiator tRNAs or the native i-tRNA CAU strains. In the IPTG induced repetition however, there is an observable increase in the growth rate of cells containing the native CAU i-tRNA. Conversely, the mutant i-tRNA strains show a decrease in growth rate. These results may be explained through processes involving ribosome maturation and potential ribosome stalling events.

### 4.3 Fitness analysis shows difference in acquired fitness from native i-tRNA when compared to mutant i-tRNAs

Growth effects of the anticodon mutant i-tRNAs were measured to determine if they are suitable for use in an orthogonal translation initiation system. It was found that there was a large increase in the growth rate of cells containing a plasmid-borne copy of the wild-type i-tRNA (**Fig. 13c**). This result may be explained by enhanced ribosome maturation due to additional i-tRNA abundance. It has been long postulated that the wild-type i-tRNA CAU forms an intermediate with the ribosome prior to translation initiation (Mayer et al., 2003). More recently, it was shown that the highly conserved CG pairs in the i-tRNA structure **(Fig. 4B),** interact with the 16s rRNA of the ribosome in later stages of ribosome biogenesis (Shetty and Varshney, 2016). It is therefore reasonable to suggest that the presence of more i-tRNAs within the cells would allow for ribosome maturation at a greater rate and allow for the production of proteins involved in growth to be translated at a greater rate. While this is still speculative, future experiments involving whole cell proteomics may be appropriate to observe the protein landscape of the cell in conditions where the production i-tRNAs are up regulated.

In contrast to the wild-type i-tRNAs, a growth defect in cells harbouring all anticodon mutant i-tRNAs was observed. This result is interesting, as the mutant tRNAs also possess the identical internal CG pairs as the wild-type i-tRNAs and thus would be expected to similarly enhance ribosome maturation and increase growth rate. These data indicate that it may not be ribosome maturation alone which is causing the enhanced growth rate in the wild-type i-tRNA strain. Alternatively, it is also likely that the altered anticodons in the mutant i-tRNAs may be hindering translation initiation through means independent of ribosome maturation, such as ribosome stalling.

Ribosome stalling, occurs within the ribosome during translation to ensure proper translation of mRNA into protein, and may be elicited in response to a number of

translation elements. Ribosome stalling is often observed when the ribosome encounters a rare codon, in response to stress factors and incorrect e-tRNA binding. Partial and permanent ribosome stalling has often been observed in cells that are starved of certain required amino acids. As a result of lower amino acid abundance, e-tRNAs bind to the ribosome, un-aminoacylated. This results in a pause of the ribosome until regulatory mechanisms disassemble the ribosome and terminate translation (Buskirk and Green, 2017). While this type of ribosome stalling has not been observed in response to un-aminoacylated i-tRNAs, similar mechanisms may be in place which potentially contribute to the reduced growth rate of cells due to lowered abundance of available ribosomes throughout the cell.

Most of the analysis within this study has speculated that the mutant i-tRNAs have been aminoacylated with the amino acids as dictated by their anticodons, as previous studies have suggested (Wu and RajBhandary, 1997, Varshney and RajBhandary, 1990, Schulman and Pelka, 1985). However, from this study there is no evidence for the aminoacylation of the mutant i-tRNAs, and thus many of the mutant i-tRNAs may not be aminoacylated within the cell, causing ribosomes to stall, hindering the rate of translation throughout the cell.

### 4.4 Future considerations

The results of this study indicate high variability in translation initiation from non-canonical anticodon-codon pairs, with some anticodon-codon interactions having increased levels of protein expression, while others showed no apparent changes in expression levels or even showed considerable decreases in protein expression levels. This variability may be largely due to a multitude of factors, including the improper aminoacylation of the i-tRNAs, the reduced formylation of aminoacyl-tRNAs or the acquiring or lack of post transcriptional modifications to achieve sufficient translation initiation. This study was unable to identify any distinct patterns which distinguished certain mutant i-tRNAs at being better at translation initiation than others. To better understand the nature of these mutant tRNAs, and potentially uncover certain patterns, future experimentation should replicate the protocols from within this study, and extend them to the remaining 53 anticodon variations. The aminoacylation of these tRNAs should also be further investigated. Once this large subset of mutant tRNAs is generated, and understood, different methods may be employed which can be used to enhance this translation initiation system and bringing it closer towards being truly orthogonal from host translation initiation.

# 5 Conclusion

Within this thesis, I have designed, constructed and tested a set of mutant i-tRNAs, in order to try and establish a foundation for an orthogonal translation initiation system in *E. coli*. The results within this thesis showed that translation initiation can occur through non-canonical anticodon-codon interactions, however these initiation events are less pronounced than initiation through the native anticodon-codon interaction. I also showed that these interactions are highly variable, and speculated on potential mechanisms which may be at play preventing efficient translation of certain anticodon-codon pairs. Furthermore, I showed that the expression of mutant initiator tRNAs reduces cell viability by hindering cell growth rate possibly due to ribosome stalling. Conversely, I showed that expression of an exogenous native i-tRNA has growth benefits to host cells, supporting previous work which shows the importance of i-tRNAs in ribosome maturation.

Although a large amount of research is currently being conducted to make areas within the central dogma of molecular biology orthogonal, there is still a long way to go before true orthogonality can be achieved. Current models across the central dogma, including the one within this study, still rely heavily on host cell mechanisms, enzymes and molecules. As a result, these methods are still highly non-orthogonal and may have undesired effects on host cell viability and circuit functioning, as shown within this study. Furthermore, many of the orthogonal systems discussed are still operate on a relatively small scale, and thus might behave differently when placed into the context of larger synthetic circuits, or when industrialised for the mass production of synthetic materials.

This study, along with a multitude of previous studies into orthogonality are still largely in their infancy, and thus further experimentation and research is crucially needed to bring parts of the central dogma of biology, or even the whole of it, towards an orthogonal future. This feat can greatly benefit synthetic biology as it can serve as platform for predictability and scalability in biological designs with applications in medicine and industry. Orthogonal systems may also lead to the development of circuits which are no longer constrained by a natural genetic code, and thus can function with higher efficiencies and with increased biosafety. Finally, developing orthogonal systems have the potential to reduce the overall burden synthetic circuits impose onto host cells, extending the potential of these designs.

# 6 References


AGRIS, P. F., VENDEIX, F. A. & GRAHAM, W. D. 2007. tRNA's wobble decoding of the genome: 40 years of modification. *Journal of molecular biology,* 366**,** 1-13.

ANDRIANANTOANDRO, E., BASU, S., KARIG, D. K. & WEISS, R. 2006. Synthetic biology: new engineering rules for an emerging discipline. *Molecular systems biology,* 2.

AYYUB, S., DOBRIYAL, D. & VARSHNEY, U. 2017. Contributions of the N-terminal and C-terminal domains of initiation factor 3 towards its functions in the fidelity of initiation and anti-association of the ribosomal subunits. *Journal of bacteriology***,** JB. 00051-17.

BARRAUD, P., SCHMITT, E., MECHULAM, Y., DARDEL, F. & TISNE, C. 2008. A unique conformation of the anticodon stem-loop is associated with the capacity of tRNAfMet to initiate protein synthesis. *Nucleic acids research,* 36**,** 4894-4901.

BARRELL, B., BANKIER, A. & DROUIN, J. 1979. A different genetic code in human mitochondria. *Nature,* 282**,** 189.

BENNE, R., NAAKTGEBOREN, N., GUBBENS, J. & VOORMA, H. O. 1973. Recycling of initiation factors IF-1, IF-2 and IF-3. *European journal of biochemistry,* 32**,** 372-380.

BORKOWSKI, O., CERONI, F., STAN, G.-B. & ELLIS, T. 2016. Overloaded and stressed: whole-cell considerations for bacterial synthetic biology. *Current opinion in microbiology,* 33**,** 123-130.

BROCK, S., SZKARADKIEWICZ, K. & SPRINZL, M. 1998. Initiation factors of protein biosynthesis in bacteria and their structural relationship to elongation and termination factors. *Molecular microbiology,* 29**,** 409-417.

BUSKIRK, A. R. & GREEN, R. 2017. Ribosome pausing, arrest and rescue in bacteria and eukaryotes. *Phil. Trans. R. Soc. B,* 372**,** 20160183.

CARTER, A. P., CLEMONS, W. M., BRODERSEN, D. E., MORGAN-WARREN, R. J., HARTSCH, T., WIMBERLY, B. T. & RAMAKRISHNAN, V. 2001. Crystal structure of an initiation factor bound to the 30S ribosomal subunit. *Science,* 291**,** 498-501.

CASERTA, E., TOMŠIC, J., SPURIO, R., LA TEANA, A., PON, C. L. & GUALERZI, C. O. 2006. Translation initiation factor IF2 interacts with the 30 S ribosomal subunit via two separate binding sites. *Journal of molecular biology,* 362**,** 787-799.

CHATTAPADHYAY, R., PELKA, H. & SCHULMAN, L. H. 1990. Initiation of in vivo protein synthesis with non-methionine amino acids. *Biochemistry,* 29**,** 4263-4268.

CHATTERJEE, A., XIAO, H. & SCHULTZ, P. G. 2012. Evolution of multiple, mutually orthogonal prolyl-tRNA synthetase/tRNA pairs for unnatural amino acid mutagenesis in Escherichia coli. *Proceedings of the National Academy of Sciences,* 109**,** 14841-14846.

CHEN, P., CRAIN, P. F., NÄSVALL, S. J., POMERANTZ, S. C. & BJÖRK, G. R. 2005. A'gain of function'mutation in a protein mediates production of novel modified nucleosides. *The EMBO Journal,* 24**,** 1842-1851.

CHIAPELLO, H., OLLIVIER, E., LANDÈS-DEVAUCHELLE, C., NITSCHKÉ, P. & RISLER, J.-L. 1999. Codon usage as a tool to predict the cellular location of eukaryotic ribosomal proteins and aminoacyl-tRNA synthetases. *Nucleic acids research,* 27**,** 2848-2851.

CHIN, J. W. 2014. Expanding and reprogramming the genetic code of cells and animals. *Annual review of biochemistry,* 83**,** 379-408.

COCHELLA, L. & GREEN, R. 2004. Wobble during decoding: more than third-position promiscuity. *Nature Structural and Molecular Biology,* 11**,** 1160.

CRICK, F. 1970. Central dogma of molecular biology. *Nature,* 227**,** 561.

DOTTAVIO-MARTIN, D., SUTTLE, D. P. & RAVEL, J. M. 1979. The effects of initiation factors IF-1 and IF-3 on the dissociation of escherichia coli 70 S ribosomes. *FEBS letters,* 97**,** 105-110.
ELLIS, T., ADIE, T. & BALDWIN, G. S. 2011. DNA assembly for synthetic biology: from parts to pathways and beyond. *Integrative Biology,* 3**,** 109-118.
ENDY, D. 2005. Foundations for engineering biology. *Nature,* 438**,** 449.
GIBSON, D. G., YOUNG, L., CHUANG, R.-Y., VENTER, J. C., HUTCHISON III, C. A. & SMITH, H. O. 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature methods,* 6**,** 343.
GUALERZI, C. O. & PON, C. L. 1990. Initiation of mRNA translation in prokaryotes. *Biochemistry,* 29**,** 5881-5889.
GUENNEUGUES, M., CASERTA, E., BRANDI, L., SPURIO, R., MEUNIER, S., PON, C. L., BOELENS, R. & GUALERZI, C. O. 2000. Mapping the fMet-tRNAfMet binding site of initiation factor IF2. *The EMBO journal,* 19**,** 5233-5240.
GUILLON, J.-M., MECHULAM, Y., BLANQUET, S. & FAYAT, G. 1993. Importance of formylability and anticodon stem sequence to give a tRNA (Met) an initiator identity in Escherichia coli. *Journal of bacteriology,* 175**,** 4507-4514.
HARTZ, D., BINKLEY, J., HOLLINGSWORTH, T. & GOLD, L. 1990. Domains of initiator tRNA and initiation codon crucial for initiator tRNA selection by Escherichia coli IF3. *Genes & development,* 4**,** 1790-1800.
HECHT, A., GLASGOW, J., JASCHKE, P. R., BAWAZER, L. A., MUNSON, M. S., COCHRAN, J. R., ENDY, D. & SALIT, M. 2017. Measurements of translation initiation from all 64 codons in E. coli. *Nucleic acids research,* 45**,** 3615-3626.
HUSSAIN, T., LLÁCER, J. L., WIMBERLY, B. T., KIEFT, J. S. & RAMAKRISHNAN, V. 2016. Large-scale movements of IF3 and tRNA during bacterial translation initiation. *Cell,* 167**,** 133-144. e13.
JULIÁN, P., MILON, P., AGIRREZABALA, X., LASSO, G., GIL, D., RODNINA, M. V. & VALLE, M. 2011. The Cryo-EM structure of a complete 30S translation initiation complex from Escherichia coli. *PLoS biology,* 9**,** e1001095.
KHALIL, A. S. & COLLINS, J. J. 2010. Synthetic biology: applications come of age. *Nature Reviews Genetics,* 11**,** 367.
KNIGHT, R. D., FREELAND, S. J. & LANDWEBER, L. F. 2001. Rewiring the keyboard: evolvability of the genetic code. *Nature Reviews Genetics,* 2**,** 49.
LAGERKVIST, U. 1978. " Two out of three": An alternative method for codon reading. *Proceedings of the National Academy of Sciences,* 75**,** 1759-1762.
LAJOIE, M. J., ROVNER, A. J., GOODMAN, D. B., AERNI, H.-R., HAIMOVICH, A. D., KUZNETSOV, G., MERCER, J. A., WANG, H. H., CARR, P. A. & MOSBERG, J. A. 2013. Genomically recoded organisms expand biological functions. *science,* 342**,** 357-360.
LAURSEN, B. S., SØRENSEN, H. P., MORTENSEN, K. K. & SPERLING-PETERSEN, H. U. 2005. Initiation of protein synthesis in bacteria. *Microbiology and molecular biology reviews,* 69**,** 101-123.
LEONTIS, N. B., STOMBAUGH, J. & WESTHOF, E. 2002. The non-Watson–Crick base pairs and their associated isostericity matrices. *Nucleic acids research,* 30**,** 3497-3531.
LIU, C. C., JEWETT, M. C., CHIN, J. W. & VOIGT, C. A. 2018. Toward an orthogonal central dogma. *Nature chemical biology,* 14**,** 103-106.
MA, N. J. & ISAACS, F. J. 2016. Genomic recoding broadly obstructs the propagation of horizontally transferred genetic elements. *Cell systems,* 3**,** 199-207.
MALYS, N. 2012. Shine-Dalgarno sequence of bacteriophage T4: GAGG prevails in early genes. *Molecular biology reports,* 39**,** 33-39.

MALYSHEV, D. A., DHAMI, K., LAVERGNE, T., CHEN, T., DAI, N., FOSTER, J. M., CORRÊA, I. R. & ROMESBERG, F. E. 2014. A semi-synthetic organism with an expanded genetic alphabet. *Nature,* 509**,** 385.

MANGROO, D. & RAJBHANDARY, U. L. 1995. Mutants of Escherichia coli Initiator tRNA Defective in Initiation EFFECTS OF OVERPRODUCTION OF METHIONYL-tRNA TRANSFORMYLASE AND THE INITIATION FACTORS IF2 AND IF3. *Journal of Biological Chemistry,* 270**,** 12203-12209.

MAYER, C., KÖHRER, C., KENNY, E., PRUSKO, C. & RAJBHANDARY, U. L. 2003. Anticodon sequence mutants of Escherichia coli initiator tRNA: effects of overproduction of aminoacyl-tRNA synthetases, methionyl-tRNA formyltransferase, and initiation factor 2 on activity in initiation. *Biochemistry,* 42**,** 4787-4799.

MAYER, C. & RAJBHANDARY, U. L. 2002. Conformational change of Escherichia coli initiator methionyl-tRNA fMet upon binding to methionyl-tRNA formyl transferase. *Nucleic acids research,* 30**,** 2844-2850.

MILON, P., KONEVEGA, A. L., GUALERZI, C. O. & RODNINA, M. V. 2008. Kinetic checkpoint at a late step in translation initiation. *Molecular cell,* 30**,** 712-720.

MOAZED, D., SAMAHA, R. R., GUALERZI, C. & NOLLER, H. F. 1995. Specific protection of 16 S rRNA by translationalinitiation factors. *Journal of molecular biology,* 248**,** 207-210.

MOLL, I., GRILL, S., GUALERZI, C. O. & BLÄSI, U. 2002. Leaderless mRNAs in bacteria: surprises in ribosomal recruitment and translational control. *Molecular microbiology,* 43**,** 239-246.

MURPHY IV, F. V. & RAMAKRISHNAN, V. 2004. Structure of a purine-purine wobble base pair in the decoding center of the ribosome. *Nature Structural and Molecular Biology,* 11**,** 1251.

NEUMANN, H., SLUSARCZYK, A. L. & CHIN, J. W. 2010. De novo generation of mutually orthogonal aminoacyl-tRNA synthetase/tRNA pairs. *Journal of the American Chemical Society,* 132**,** 2142-2144.

OSAWA, S. & JUKES, T. H. 1989. Codon reassignment (codon capture) in evolution. *Journal of molecular evolution,* 28**,** 271-278.

OSTROV, N., LANDON, M., GUELL, M., KUZNETSOV, G., TERAMOTO, J., CERVANTES, N., ZHOU, M., SINGH, K., NAPOLITANO, M. G. & MOOSBURNER, M. 2016. Design, synthesis, and testing toward a 57-codon genome. *Science,* 353**,** 819-822.

PEREZ-PINERA, P., KOCAK, D. D., VOCKLEY, C. M., ADLER, A. F., KABADI, A. M., POLSTEIN, L. R., THAKORE, P. I., GLASS, K. A., OUSTEROUT, D. G. & LEONG, K. W. 2013. RNA-guided gene activation by CRISPR-Cas9–based transcription factors. *Nature methods,* 10**,** 973.

PON, C. L. & GUALERZI, C. O. 1984. Mechanism of protein biosynthesis in prokaryotic cells. *FEBS letters,* 175**,** 203-207.

RACKHAM, O. & CHIN, J. W. 2005. A network of orthogonal ribosome· mRNA pairs. *Nature chemical biology,* 1**,** 159.

RAVIKUMAR, A., ARRIETA, A. & LIU, C. C. 2014. An orthogonal DNA replication system in yeast. *Nature chemical biology,* 10**,** 175.

SCHMIDT, M. 2012. *Synthetic biology: industrial and environmental applications*, John Wiley & Sons.

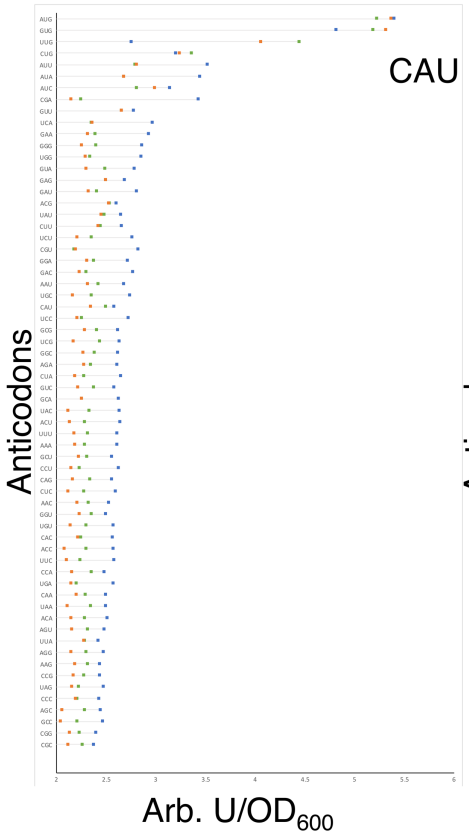
SCHULMAN, L. H. & PELKA, H. 1985. In vitro conversion of a methionine to a glutamine-acceptor tRNA. *Biochemistry,* 24**,** 7309-7314.

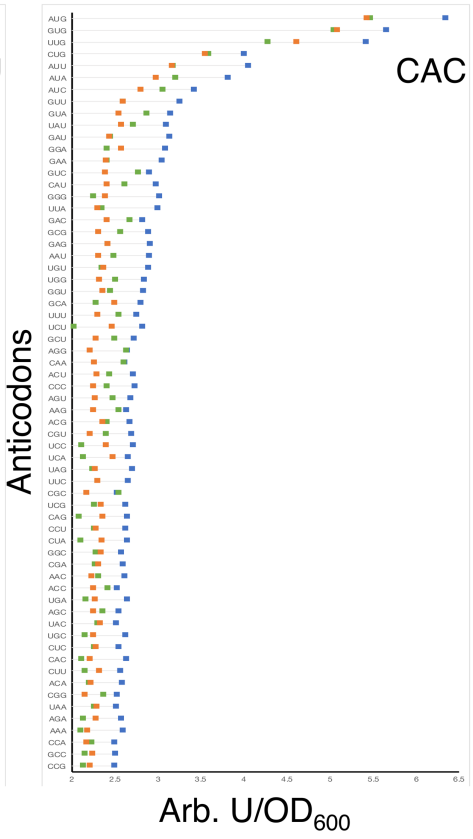
SCHULTZ, D. W. & YARUS, M. 1996. On malleability in the genetic code. *Journal of molecular evolution,* 42**,** 597-601.

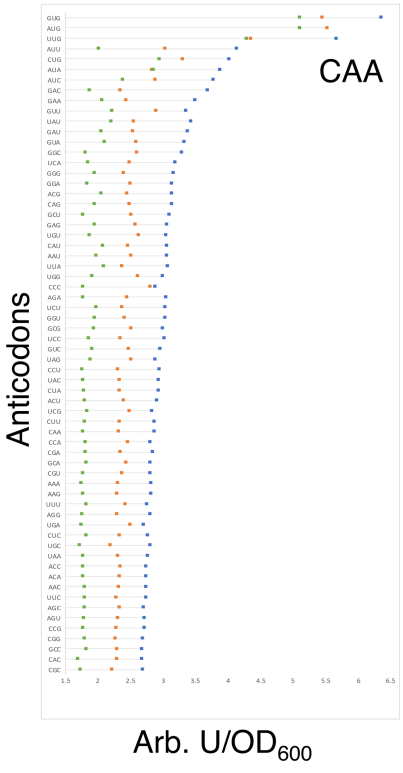
SCHWEIZER, U., BOHLEBER, S. & FRADEJAS-VILLAR, N. 2017. The modified base isopentenyladenosine and its derivatives in tRNA. *RNA biology,* 14**,** 1197-1208.
SELMER, M., DUNHAM, C. M., MURPHY, F. V., WEIXLBAUMER, A., PETRY, S., KELLEY, A. C., WEIR, J. R. & RAMAKRISHNAN, V. 2006. Structure of the 70S ribosome complexed with mRNA and tRNA. *Science,* 313**,** 1935-1942.
SHARP, P. M. & LI, W.-H. 1987. The codon adaptation index-a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic acids research,* 15**,** 1281-1295.
SHETTY, S. & VARSHNEY, U. 2016. An evolutionarily conserved element in initiator tRNAs prompts ultimate steps in ribosome maturation. *Proceedings of the National Academy of Sciences,* 113**,** E6126-E6134.
SHINE, J. & DALGARNO, L. 1974. The 3′-terminal sequence of Escherichia coli 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. *Proceedings of the National Academy of Sciences,* 71**,** 1342-1346.
SIMONETTI, A., MARZI, S., MYASNIKOV, A. G., FABBRETTI, A., YUSUPOV, M., GUALERZI, C. O. & KLAHOLZ, B. P. 2008. Structure of the 30S translation initiation complex. *Nature,* 455**,** 416.
STRINGER, E. A., SARKAR, P. & MAITRA, U. 1977. Function of initiation factor 1 in the binding and release of initiation factor 2 from ribosomal initiation complexes in Escherichia coli. *Journal of Biological Chemistry,* 252**,** 1739-1744.
STUDIER, F. W. & MOFFATT, B. A. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *Journal of molecular biology,* 189**,** 113-130.
STUDIER, F. W., ROSENBERG, A. H., DUNN, J. J. & DUBENDORFF, J. W. 1990. [6] Use of T7 RNA polymerase to direct expression of cloned genes. *Methods in enzymology*. Elsevier.
SUSSMAN, J. K., SIMONS, E. L. & SIMONS, R. W. 1996. Escherichia coli translation initiation factor 3 discriminates the initiation codon in vivo. *Molecular microbiology,* 21**,** 347-360.
TABOR, S. & RICHARDSON, C. C. 1985. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proceedings of the National Academy of Sciences,* 82**,** 1074-1078.
UDAGAWA, T., SHIMIZU, Y. & UEDA, T. 2004. Evidence for the translation initiation of leaderless mRNAs by the intact 70 S ribosome without its dissociation into subunits in eubacteria. *Journal of Biological Chemistry,* 279**,** 8539-8546.
URBONAVIČIUS, J., QIAN, Q., DURAND, J. M., HAGERVALL, T. G. & BJÖRK, G. R. 2001. Improvement of reading frame maintenance is a common function for several tRNA modifications. *The EMBO journal,* 20**,** 4863-4873.
VARANI, G. & MCCLAIN, W. H. 2000. The G· U wobble base pair: a fundamental building block of RNA structure crucial to RNA function in diverse biological systems. *EMBO reports,* 1**,** 18-23.
VARSHNEY, U. & RAJBHANDARY, U. L. 1990. Initiation of protein synthesis from a termination codon. *Proceedings of the National Academy of Sciences,* 87**,** 1586-1590.
VARSHNEY, U. & RAJBHANDARY, U. L. 1992. Role of methionine and formylation of initiator tRNA in initiation of protein synthesis in Escherichia coli. *Journal of bacteriology,* 174**,** 7819-7826.
VINCENT, R. M. 2017. *Reprogramming the genetic code: defining a new start codon in Escherichia coli.* Masters of Research Macquarie University.

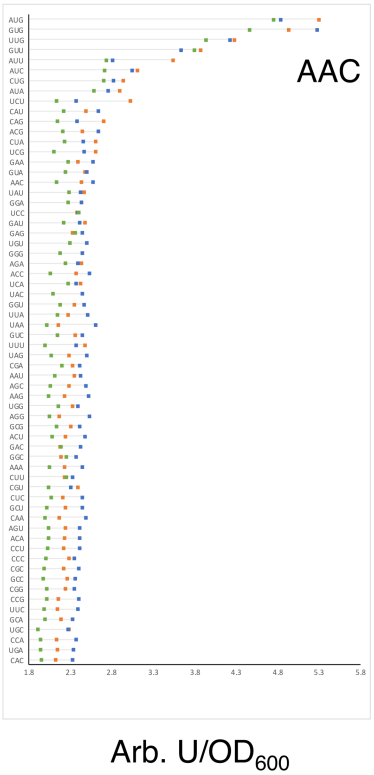
WANG, K., NEUMANN, H., PEAK-CHEW, S. Y. & CHIN, J. W. 2007. Evolved orthogonal ribosomes enhance the efficiency of synthetic genetic code expansion. *Nature biotechnology,* 25**,** 770.
WANG, K., SACHDEVA, A., COX, D. J., WILF, N. M., LANG, K., WALLACE, S., MEHL, R. A. & CHIN, J. W. 2014. Optimized orthogonal translation of unnatural amino acids enables spontaneous protein double-labelling and FRET. *Nature chemistry,* 6**,** 393.
WATSON, J. D. & CRICK, F. H. 1953. Molecular structure of nucleic acids. *Nature,* 171**,** 737-738.
WEBER, W. & FUSSENEGGER, M. 2012. Emerging biomedical applications of synthetic biology. *Nature Reviews Genetics,* 13**,** 21.
WINTERMEYER, W. & GUALERZI, C. 1983. Effect of Escherichia coli initiation factors on the kinetics of N-Acphe-tRNAPhe binding to 30S ribosomal subunits. A fluorescence stopped-flow study. *Biochemistry,* 22**,** 690-694.
WU, X.-Q., IYENGAR, P. & RAJBHANDARY, U. 1996. Ribosome-initiator tRNA complex as an intermediate in translation initiation in Escherichia coli revealed by use of mutant initiator tRNAs and specialized ribosomes. *The EMBO journal,* 15**,** 4734-4739.
WU, X.-Q. & RAJBHANDARY, U. L. 1997. Effect of the amino acid attached to Escherichia coli initiator tRNA on its affinity for the initiation factor IF2 and on the IF2 dependence of its binding to the ribosome. *Journal of Biological Chemistry,* 272**,** 1891-1895.
YAMAMOTO, H., WITTEK, D., GUPTA, R., QIN, B., UEDA, T., KRAUSE, R., YAMAMOTO, K., ALBRECHT, R., PECH, M. & NIERHAUS, K. H. 2016. 70S-scanning initiation is a novel and frequent initiation mode of ribosomal translation in bacteria. *Proceedings of the National Academy of Sciences,* 113**,** E1180-E1189.
ZHOU, M., LONG, T., FANG, Z.-P., ZHOU, X.-L., LIU, R.-J. & WANG, E.-D. 2015. Identification of determinants for tRNA substrate recognition by Escherichia coli C/U34 2′-O-methyltransferase. *RNA biology,* 12**,** 900-911.

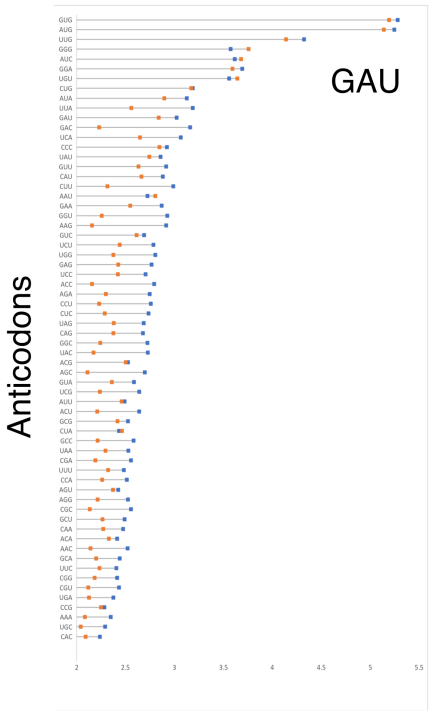
# Supplementary Figures

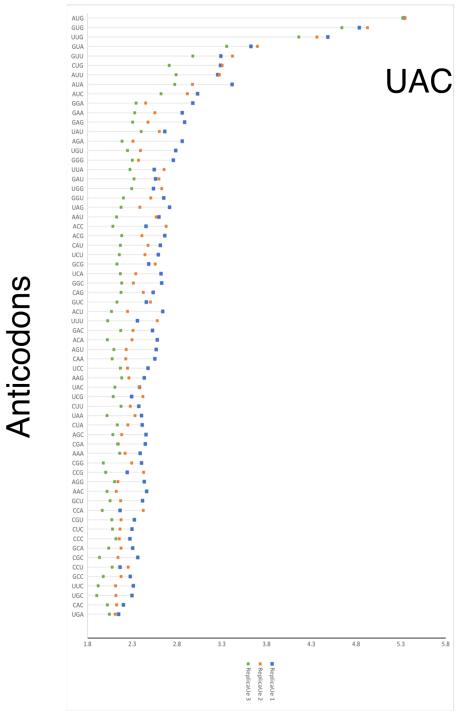


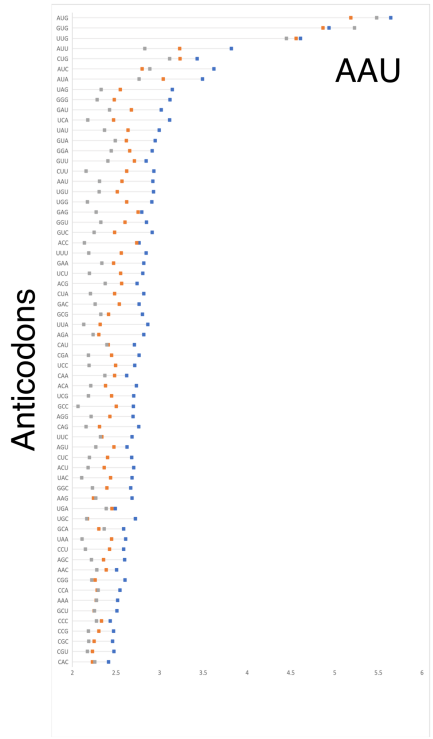



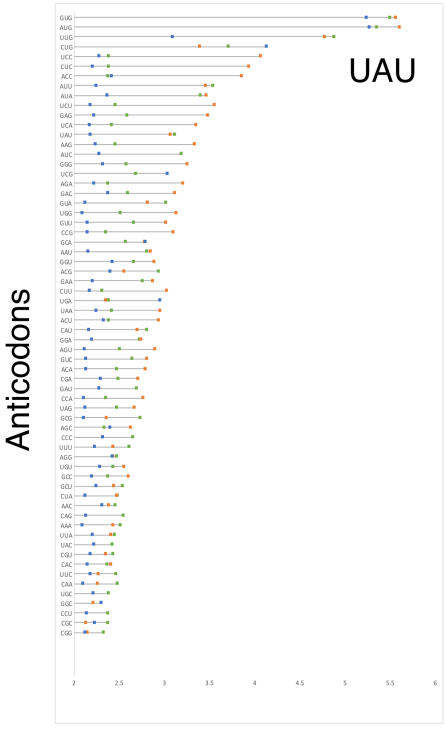



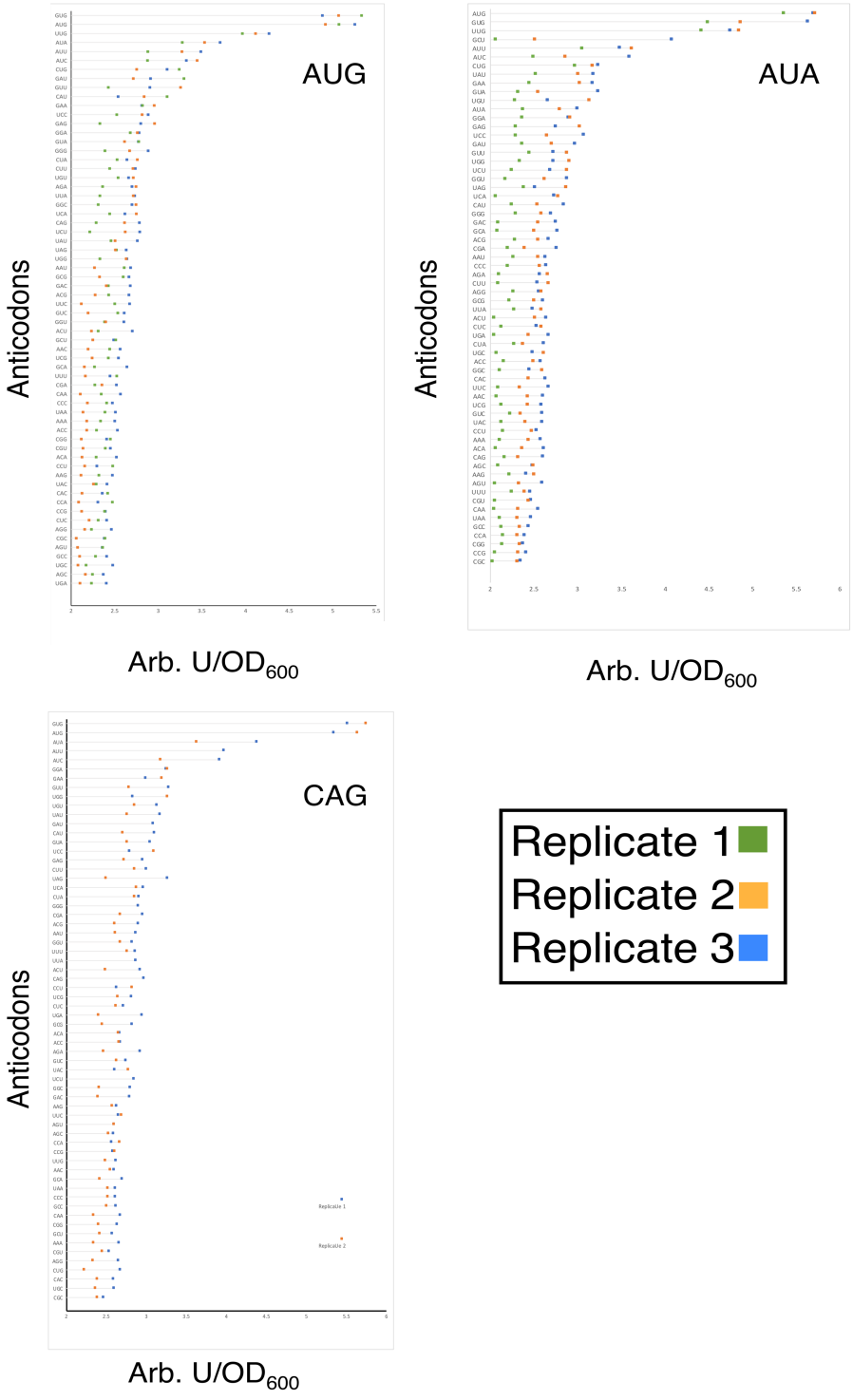




GAU
Anticodons
Arb. U/OD600


UAC
Anticodons
Arb. U/OD600


AAU
Anticodons
Arb. U/OD600


UAU
Anticodons
Arb. U/OD600




**Supplementary Figure 1. Bulk fluorescence results for each mutant i-tRNA.** Normalised bulk fluorescence data (Arb.U/$OD_{600}$) of the interaction between induced mutant i-tRNAs and the mutant sfGFP reporters with altered start codons.



13/12/2017

Dear Dr Paul Jaschke ,

**Re: 5201600567 - Reprogramming the genetic code to have a unique signal for gene start sites - 0135- [ 520170135260 ]**

Your Personnel Amendment request for the above mentioned project has been approved effective 13 December 2017

Please note this amendment will not change the project expiration date 09/09/2021 .

Kind Regards,
Claudia Huang

Biosafety Secretariat
Research Office
Level 3, Research Hub, Building C5C East
Macquarie University
NSW 2109 Australia
T: +61 2 9850 4063
F: +61 2 9850 4465
http://www.mq.edu.au/research

# Animal Ethics Application

| | |
|---|---|
| Application ID : | 5201600567 |
| Application Title : | Reprogramming the genetic code to have a unique signal for gene start sites |
| Date of Submission : | N/A |
| Primary Investigator : | Dr Paul Jaschke |
| Other Investigators : | Mr Dominic Logel |

## Instruction

### For Applicants

**Ethics category code**
**For biosafety applications change code to "biosafety"***

Biosafety

A Macquarie University Biosafety Application is required for all research and teaching projects that involve the use of biological materials.
The Macquarie University Institutional Biosafety Committee reviews research and teaching proposals involving any of the below categories. Work involving these categories must not commence until IBC approval has been received.

- Microorganisms classified as risk group 2 and above
- Genetically modified organisms
- Security sensitive biological agents
- Agents requiring quarantine containment
- Animals with the potential to carry zoonotic agents classified as risk group 2 and above
- Human and animal clinical and diagnostic samples

For work involving the use of risk group 1 microorganisms (including animals with the potential to carry risk group 1 zoonoses), animal or human tissue, or invertebrates completion of this form is still required, however it is treated as a notification whereby IBC review is not required. All applications must be approved by the Chief Investigator and all are subject to IBC review. Work involving the use of animals will also require animal ethics approval.

The Biosafety Applicant Guide is available to assist you with completing this online form.

- This online application form applies to all work involving biological materials. For all other forms and templates (e.g. progress/final reports, amendments, etc.), please visit the Managing Approved Projects page on the IBC website.
- For more information on the application process, please consult the IBC website.
- You may begin this form anytime, save, and continue the form, however, this form will not be considered until it is submitted.
- If you wish to email a copy of your draft application to the Biosafety Advisor or any of your MQ or external colleagues, you will need to perform this action *before* submitting your application for signoff.
- Please complete all mandatory questions before attempting to submit your application for signoff.
- Your application will be considered via an online expedited review process
- If this form has not been submitted within six (6) months of the creation date, your draft form may be removed from our system.
- If you incur problems with your browser while completing this application, please attempt to complete the form with another browser before contacting MQ for help.
- For all other issues, please contact:

**For technical issues:**
(web) OneHelp Enquiry
(email) IRIS Help
(phone) +61 2 9850 4455

For issues specific to your biosafety application:
(contact person): Biosafety Advisor: Liette Vandine
(email) Biosafety
(phone) +61 2 9850 4063

## 1. Administration

### 1.1 Project Details

Application ID

5201600567

1.1 **Title of Proposed Project***

Reprogramming the genetic code to have a unique signal for gene start sites

1.2 **Please indicate the purpose(s) of this application**
*Tick all applicable**

- [x] Research
- [ ] Teaching

1.3 **Reason for application** *

- (•) New Project
- ( ) Resubmission
- ( ) Continuation of an expired, or soon to expire, project
- ( ) Update a current project

1.4 **Chief Investigator:** - Primary Contact
Please click on your name below and complete all mandatory questions.
If the project is to be undertaken by a student the supervisor will be considered the Chief Investigator.
**Note: The Chief Investigator is responsible for ensuring that all project personnel receive appropriate training.**
If you are not the Chief Investigator, please add their details in the search box below and remove your name. Please click on the CI's name and complete all mandatory questions.
*

| | | |
|---|---|---|
| 1 | Given Name | Paul |
| | Surname | Jaschke |
| | Full name | Dr Paul Jaschke |
| | MQ ID | MQ20154149 |
| | Department: | 4301 |
| | **Qualifications** | B.Sc Biochemistry (Honours), Ph.D. Microbiology & Immunology, Postdoctoral Training Bioengineering |
| | **Position in this project** | Chief Investigator |
| | **Primary Contact?**<br>*Only Chief Investigator can be Primary Contact* | Yes |
| | **Full mailing address:** | Department of Chemistry and Biomolecular Sciences Building E8A, Room 357 Macquarie University Balaclava Road North Ryde, NSW 2109 |
| | Work number: | +61-2-9850-8295 |
| | Email address: | paul.jaschke@mq.edu.au |

**Department***

Chemistry & Biomolecular Sciences

1.5

**Project Category**
The following links will provide information on risk group and GMO classifications

***Notification**: Risk group 1 microorganisms and zoonoses, animal or human tissue, or invertebrates*

***Risk group 2 and above**: Risk group 2 and above microorganisms and zoonoses, animals with the potential to carry risk group 2 and above zoonoses, clinical and diagnostic samples*

*

- ○ Notification
- ○ Risk group 2 and above
- ◉ GMO Exempt
- ○ GMO NLRD
- ○ GMO DNIR/DIR

1.7 **Do you require quarantine containment or a Permit for Movement?***

- ○ Yes
- ◉ No

1.8 **Does your project involve the use of a security sensitive biological agent (SSBA)?***

- ○ Yes
- ◉ No

1.9 **Proposed start date***

05/09/2016

1.10 **Proposed end date***

06/09/2021

1.11 **Identify all facilities required for your work***

- ☐ General laboratory
- ☑ PC1
- ☐ PC2
- ☐ QC1
- ☐ QC2
- ☐ CAF
- ☐ Zebrafish Facility
- ☐ Fauna Park
- ☑ Other facilities

1.11.2 **PC1 Laboratory**
Please provide exact location*

PC1 Laboratory - E8C 230

1.11.8 **Other Facilities**
Please provide exact locations*

Flow Cytometry Facility - E8C 264

1.12 **Are additional approvals required?** (Tick all appropriate)

- [ ] Animal Ethics
- [ ] Human Ethics
- [ ] Fieldwork
- [ ] Other

*This question is not answered.*

## 2. Project Information

### 2.1 Biological agents and control measures

2.1.1 **What is the type of biological material?***

- [ ] Animal
- [x] Bacteria
- [ ] Virus
- [ ] Other
- [ ] Cell Line
- [ ] Fungus
- [x] Nucleic Acid
- [ ] Plant
- [ ] Prions
- [ ] Soil
- [ ] Tissue
- [ ] Toxins

2.1.2 **What is the name of the biological agent(s)? (separate by comma if multiple)***

Escherichia coli C321.deltaA.exp, E. coli DH10B, E. coli DH5alpha, E. coli Stbl4, pEVOL, pULTRA, M13 bacteriophage

2.1.3 **Control measures**
(researchers must also consider the control of aerosols)*

- [ ] Eliminate risk
- [ ] Substitute the hazard
- [ ] Isolate the hazard
- [ ] Implement engineering controls
- [x] Administration (eg. SWP/Training)
- [x] PPE

**Administration**
Please provide details*

We will perform laboratory inductions for all personnel including reviewing SWP and risk assessments. We will ensure personnel have undergone the online biosafety course.

**PPE**
Please provide details*

We will ensure personnel use the appropriate PPE listed below in 2.1.4 and are trained in the correct use of these items.

2.1.4 **What personal protective equipment is required?***

- [x] Gloves (e.g. chemical resistant)
- [x] Clothing (e.g. button up lab coat/coveralls/apron)
- [ ] Respiratory Protection (e.g. PF2 face mask)
- [x] Eye protection (e.g. safety glasses/goggles)
- [x] Footwear (e.g. Enclosed/Gumboots/overshoe covers)
- [ ] Other

2.1.5 **Attach any supporting documents which must be read in conjunction with the application (e.g. Safe Working Procedures, Safety Data Sheets, Guidelines/Protocols) in Section 7 - Attachment**

### 2.2 Risks associated with the biological agents

2.2.1 **Which risk group does your biological agent belong to?**
(Note: there can be more than one risk group depending on method. Tick all applicable)*

- [x] Risk Group 1 - Low Individual and community risk
- [ ] Risk Group 2 - Moderate individual risk, limited community risk
- [ ] Risk Group 3 - High individual risk, limited community risk

2.2.2 **Risk Group 1- Low Individual and community risk**
(Microorganism that is unlikely to cause human, plant or animal disease)

**Risk reduction methods that must be followed by the researcher**
1 - Standard laboratory procedures will be followed in accordance with Laboratory Microbiological Standards AS/NZ 2243:3:2010 and University guidelines and include spillage and emergency response.
2 - Chief Investigator has completed the University Biosafety training course.
3 - The Chief Investigator confirms that personnel working on this project will receive appropriate biosafety training and instruction or has adequate supervision and understands safe laboratory practice according to AS/NZ2243:3:2010 and University guidelines.

2.2.3 **Detail the biohazards and risks associated with biological agent(s) under Risk Group 1***

E. coli K12 strains, including C321.DeltaA.exp, DH10B, DH5alpha, and Stbl4 are generally recognised as safe (GRAS), and are not known to cause disease in healthy adults.

All strains identified in this application contain antibiotic resistance genes that if transferred to other bacteria could pose a hazard. The probability of this event is low. The impact of this event is also low.

2.2.4 **How will you reduce any associated risks associated with biological agent(s) under this Risk Group?***

We will follow standard laboratory procedures in accordance with Laboratory Microbiological Standards AS/NZ 2243:3:2010 and university guidelines include spillage.

All personnel working with the GMOs will have attended university Biosafety training course and had a full laboratory induction prior to work commencement.

### 2.3 Summary of Project

2.3.1 **What are the aims and purpose of this project?** (limit to 4000 characters)*

Through the creation of a recoded strain of bacteria this project aims to enable scientists and engineers to precisely program where genes start and provide understanding of this fundamental biological process.

2.3.2 **Provide details of the procedures used in this project and the exact locations where experiments will be performed.** (limit to 4000 characters)*

E. coli cultures will be grown at 37C on LB agar plates and in flasks in LB liquid medium with appropriate antibiotics (zeomycin and ampicillin). These cultures will be grown in the lab in room 230 of building E8C and in the controlled temperature room adjacent to this lab (no room number available).

pEVOL and pULTRA vectors will be modified by assembling methionine tRNA genes and metR genes from E. coli K12 origin into it. Additionally, libraries of metR gene (of E. coli K12 origin) variants will be generated using site-directed mutagenesis and will be assembled with the pEVOL and pULTRA vectors. These procedures will be performed in the lab in E8C 230.

Cells from cultures will be analyzed in the flow cytometry facility in E8C 264.

2.3.3 **How are samples transported, stored and disposed? Include exact room and disposal locations.***

Bacterial samples will be transported between E8C 230 and the adjacent temperature controlled room in agar petri dishes or plastic tubes within a tube rack. Purified protein samples will be stored in 1.5 mL plastic tubes within a cardboard box in a 80C freezer in E8C 230 and transported to APAF by walking the samples over. Disposal will take place in the autoclave room in E8C 202. Samples collected in the biological waste bins in E8C 230 lab will be disposed of by autoclaving in E8C 202 followed by disposal into the yellow Biological Hazard waste bins. These bins will then be taken to the E5A compound for disposal.

2.3.4 **Describe the clean-up procedures that will be conducted after experiments are completed and in the event of a spill**
Refer to the Macquarie University Clean-up biological spills safe work procedure*

Clean up procedures after experiments are completed will involve collecting the liquid cultures and agar petri dishes followed by autoclaving and disposed of as above (2.3.3).
In the event of a spill cleanup will follow the Clean up Biological Spills SWP. Briefly, appropriate PPE is donned and the biological material is contained and wiped towards a central point using absorbent material (e.g. paper towel). The area following removal of the contaminated paper towels will be covered with a disinfectant and after the appropriate contact time the area will be wiped with more paper towel. The paper towels will be disposed of in a biological water bin that are periodically autoclaved and disposed of as above (2.3.3).

## 3. GMO - Exempt Dealing

### 3.1 Exempt Dealing Category

**Visit the Dealing Classification webpage for assistance in determining the appropriate category for your research.**
**Please tick the appropriate box/s:**

**OGTR Item 2 - A dealing with a genetically modified *Caenorhabditis elegans*, whereby:**

2a) an advantage is not conferred on the animal by the genetic modification; and

☐ Yes

*This question is not answered.*

2b) as a result of the genetic modification, the animal is not capable of secreting or producing an infectious agent.

☐ Yes

*This question is not answered.*

**OGTR Item 3 - A dealing with an animal into which genetically modified somatic cells have been introduced, if:**

3a) the somatic cells are not capable of giving rise to infectious agents as a result of the genetic modification; and

☐ Yes

*This question is not answered.*

3b) the animal is not infected with a virus that is capable of recombining with the genetically modified nucleic acid in the somatic cells.

☐ Yes

*This question is not answered.*

**OGTR Item 3A - A dealing with an animal whose somatic cells have been genetically modified in vivo by a replication defective viral vector, if:**

3Aa) the in vivo modification occurred as part of a previous dealing; and

☐ Yes

*This question is not answered.*

3Ab) the replication defective viral vector is no longer in the animal; and

☐ Yes

*This question is not answered.*

3Ac) no germ line cells have been genetically modified; and

☐ Yes

*This question is not answered.*

3Ad) the somatic cells cannot give rise to infectious agents as a result of the genetic modification; and

☐ Yes

*This question is not answered.*

3Ae) the animal is not infected with a virus that can recombine with the genetically modified nucleic acid in the somatic cells of the animal.

☐ Yes

*This question is not answered.*

**OGTR - Item 4**

4.1. Subject to subitem 2) below, a dealing involving a host/vector system mentioned in Part 2 of Schedule 2 and producing no more than 25 litres of GMO culture in each vessel containing the resultant culture.

☑ Yes

4.2) The donor nucleic acid:
a) must meet either of the following requirements:
i) it must not be derived from organisms implicated in, or with a history of causing, disease in otherwise healthy human beings, animal, plants or fungi;

☑ Yes

ii) it must be characterised and the information derived from its characterisation show that it is unlikely to increase the capacity of the host or vector to cause harm; and

☑ Yes

4.2. b) must not code for a toxin with an LD50 of less than 100µg/kg; and

☑ Yes

4.2. c) must not code for a toxin with an LD50 of 100µg/kg or more, if the intention is to express the toxin at high levels; and

☑ Yes

4.2. d) must not be uncharacterised nucleic acid from a toxin-producing organism; and

☑ Yes

4.2. e) must not include a viral sequence, unless the donor nucleic acid:
i) is missing at least 1 gene essential for viral multiplication that:

- A) is not available in the cell into which the nucleic acid is introduced; and
- B) will not become available during the dealing; and

ii) cannot restore replication competence to the vector.

☑ Yes

**OGTR - Item 5**
A dealing involving shot-gun cloning, or the preparation of a cDNA library, in a host/vector system mentioned in item 1 of Part 2 of Schedule 2, if the donor nucleic acid is not derived from either:
a) a pathogen; or
b) a toxin-producing organism.

☐ Yes

*This question is not answered.*

### 3.2 Donor DNA details

3.2.1 **Describe the biological source of the donor DNA to be used - include the genus, species and strain or organ/tissue as applicable. Include the specific genes to be involved in the dealing.***

The metY, metZ, and metG gene source DNA are all of E. coli K12 strain MG1655 origin

### 3.3 Host Organism

3.3.1 **Describe the host organism or tissue to be used - include the genus, species and strain where applicable. If not a commonly used laboratory strain, include the name of the strain from which it is derived.***

E. coli K12 strains

### 3.4 Vectors

3.4.1 Describe the vectors or methods to be used to transfer donor DNA to the host. Include information regarding the origin and properties of the vector and confirm that all bacterial plasmid vectors are non-conjugative. If your project involves the use of a replication defective viral vector (unable to transduce human cells), please provide an explicit description of the assay you intend to perform to exclude the presence of replication competent virus. *

pEVOL and pULTRA vectors will be used to transfer donor DNA to the E. coli K12 strains. Bacterial transformation will be accomplished through chemically competence. These vectors are non-conjugative with p15a origins and ampicillin and chloramphenicol resistance markers.

### 3.5 Justification

3.5.1 If you believe the protein/gene is characterised, and unlikely to increase the capacity of the host or vector to cause harm, or secrete a toxin, briefly explain why, referring to what is known about its structure, function and/or genetics.*

The genes code for transfer RNA and a methionyl-transferase enzyme, neither of whic is associated with any toxin or pathogenic process. The genes are sourced from the same E. coli K12 strain that we are transferring the gene into. Therefore, it is highly unlikely that an additional copy of a slightly modified gene from the host will confer undesirable characteristics on the host strain.

### 3.6 Research Location

3.6.1 **Where will the research be conducted (building, room, facility)** Please provide details of all facilities to be used for this dealing (include room number, facility type and physical containment level)*

¿Research will be conducted using the live bacteria in the E8C 230 (PC1) laboratory and adjacent warm room. Flow cytometry experiments on live cells will be conducted in E8C 264.

## 6. Other project personnel

### 6.1 Categories of project personnel

**Use categories for proposed project personnel. If categories of personnel are chosen and approved by the IBC it is under the provision that:**
**- the IBC secretariat will be notified of the details of these people before they start work on the project. Chief Investigators' can notify the IBC Secretary by submitting an Amendment request - Change of Personnel**
**- New personnel have been suitable trained, inducted and completed the University's biosafety workshop.**
**Note: The Chief Investigator is responsible for ensuring that all project personnel receive appropriate training.**

6.1.1 Please tick appropriate boxes

- [x] Co-Investigator
- [ ] Research Assistant
- [ ] Technical Staff
- [x] Masters Students
- [x] PhD Students
- [ ] PostDocs
- [ ] Visiting Academics
- [x] Other

6.1.2 If other, please specify:

Undergraduate and postgraduate volunteers

### 6.2 Individual project staff

6.2.1 **Other Personnel**
Please add details of other personnel involved in this project. At least one person needs to be included in case the Chief Investigator is absent or uncontactable*

| | | |
|---|---|---|
| 1 | Given Name | Dominic |
| | Surname | Logel |
| | Full name | Mr Dominic Logel |
| | MQ ID | MQ20141424 |
| | Department: | 4301 |
| | **Qualifications** | B.Sc. Biomolecular Sciences, MRes |
| | **Position in this project** | PhD Student |
| | **If student, please specify degree, course in which enrolled** | Ph.D. |
| | **Full mailing address:** | Department of Chemistry and Biomolecular Sciences Building E8A, Room 357 Macquarie University Balaclava Road |
| | Work number: | |
| | *We currently do not have your work phone number in our database. Please provide one below.* **Work number:** | +61298508295 |
| | Email address: | dominic.logel@mq.edu.au |

## 7. Attachments

### 7.1 Documents

*Attachments may include, for example, Safe Work Procedures, Standard Operating Procedures, GMO table, Import Permit, Permit for Movement, laboratory induction sheet, etc. Please note that the max size limit is **40MB** for an attachment. If any of your attachments is bigger than 40MB then split it into two or more parts and attach, just clearly mark those as part1, part2 etc...*

**To begin attaching items:**

1. Click **Add New Document**
2. Place the cursor in the textbox and **type the name of the attachment** (as listed above)
3. Click on the **green tick** to confirm the name
4. Click on the **Soft copy icon** to open the browsing window and select a file
5. Press **OK** to attach (and repeat the process as necessary).

| Description | Reference | Soft copy | Hard copy |
|---|---|---|---|
| SWP - Handling Sharps and Syringes | handling_sharps_syringes.pdf | ✓ | |
| SWP - High Speed Centrifuges | use_of_ultra_high_speed_centrifuges_sorval_beckman.pdf | ✓ | |
| SWP-Autoclave | swp-autoclave.pdf | ✓ | |
| SWP-Centrifuges | use_of_laboratory_centrifuges.pdf | ✓ | |
| SWP-Lab Coat Laundry | lab_coat_laundry_swp.pdf | ✓ | |
| SWP-Microwave | use_of_microwave_oven_to_melt_agarose.pdf | ✓ | |
| SWP-Spill Cleanup | clean_up_biological_spills_swp.pdf | ✓ | |

## 8. Sign off

### 8.1 Signoff

By submitting this assessment the Chief Investigator identified in Section A, confirms that any supporting documents, training, guidance, instruction or protocols issued by the University will been followed so far as reasonably practicable to ensure the work is carried out without risk to health, safety or the environment. The Chief investigator is responsible for ensuring, so far as reasonably practicable the safety of researchers and others who may be affected by the work described within this document.*

☑ I agree

Signoff date:*

18/07/2016

**Conflict of Interest***

- ○ I am also a member of MQ IBC
- ○ Other conflict of interest
- ◉ I foresee no potential conflict of interest in submitting this research proposal to MQ

Submit Application