Investigating tRNA abundance as a determinant of decoding speed in *Saccharomyces cerevisiae.*

Chloë Pain

MSc Cell Biology

Department of Biosciences

University of Kent

2015/2016

Declaration

No part of this thesis has been submitted in support of any other application for a degree or qualification of the University of Kent, or any other university or institution of learning.

Acknowledgments

I would like to say a huge thank you to Dr. Tobias von der Haar for his continued support throughout my masters. It was touch and go whether I could continue with academia before I joined his lab and now I am going on to do a PhD.

I could not have asked for a nicer lab to be part of and I have had a very enjoyable year. I would especially like to thank Ronan Egan, Eleanna Kazana, Gemma Staniforth, Lyne Josse and Andrew Strange; who not only put up with me for the whole year, but provided me with unconditional support.

Abstract

There are two main points that control the rate of translation, translation initiation which controls the rate of recruitment of ribosomes to the mRNA, and translation elongation which controls the speed with which ribosomes decode the mRNA. When initiation is not limiting, ribosomal queues can build up, reducing ribosomal clearance from the start codon. The dynamics of elongation control are largely determined by the decoding speed of individual codons. One way in which the system controls the rate and efficiency of translation is through codon bias, taking advantage of the differences in decoding speed. Our aim in this study is to investigate the role of tRNA abundance in determining decoding speed for the synonymous codons of threonine.

Using *Saccharomyces cerevisiae*, we build codon reporter constructs where a codon optimised *Renilla* luciferase gene was preceded by ten codon repeats of one of the four threonine codons, in such a way that the decoding speed of the codon repeats determined the expression levels of the luciferase. Our analyses showed that the observed protein levels of RLuc correlated with the predicted order of decoding speeds, based on our knowledge of threonyl-tRNA populations, in descending order ACT, ACA, ACC and ACG. However, RNA analyses suggest that only one out of the four threonine reporters, ACG^{10}_Rluc was limited at the level of translation thus able to report on decoding speed. Our current data suggest that decoding speed reporters based on the principle tested here are of limited usefulness, because of unanticipated interactions between codon usage and mRNA levels.

Contents

Chapter 2: Materials and Methods

Chapter 1: Introduction

1 Introduction to translation

Proteins are imperative to a vast array of functions in the cell including catalysis, maintenance of cellular structure, transport of materials and DNA replication to name a few. Proteins consist of monomers known as amino acids whose order help specify its 3-dimensional structure; the information being encoded within the genetic code. Each amino acid is encoded by a triplet of adenine, cytosine, guanine or thymine bases, giving rise to the 20 amino acids. The triplet codes are decoded by the ribosome during translation, which is a key step in gene expression, but the DNA must first be transcribed into messenger RNA (mRNA). There are three subdivisions of translation; (1) initiation, where the ribosomal subunits associate with the mRNA molecule, (2) elongation which involves the movement of ribosome along the mRNA binding the amino acids together into a polypeptide and (3) termination resulting in dissociation and release of the ribosome from the mRNA molecule (Sonenberg & Hinnebusch 2009). There are numerous additional steps that are essential, particularly important is the charging of transfer RNAs (tRNAs) and their transfer to the ribosome. The redundancy of the many proteins involved in the system also reduces the pressure on the translational machinery.

Protein production is costly to the cell, requiring a significant amount of resources and energy (Rodnina et al. 2005). Selection acts on the system to ensure maximum efficiency, fidelity and energy expenditure and one example of this is codon usage, where particular codons are preferentially utilised in highly expressed genes. Such codon bias is thought to reduce the requirement of ribosomes by approximately 5%, in comparison to using all codons equally (Gardin et al. 2014). The molecular determinants of codon usage are yet to be ascertained and require knowledge of each component and process involved in protein synthesis.

1.1 The ribosomal complex

One of the major players in the translational system is the ribosomal complex, which consists of the small (40S) and large (60S) subunit, along with ribosomal RNA and proteins. The two subunits require approximately 200 factors to associate and both units contribute to the arrangement of three binding sites, the aminoacyl (A), peptidyl (P) and exit (E) site (figure 1.1, Frank 2003). An mRNA transcript binds to a groove in the small subunit where it is decoded with the help of a tRNA situated in the A-site. Whereas, the large subunit is responsible for peptide bond formation by the peptidyl-transferase centre in the P-site.

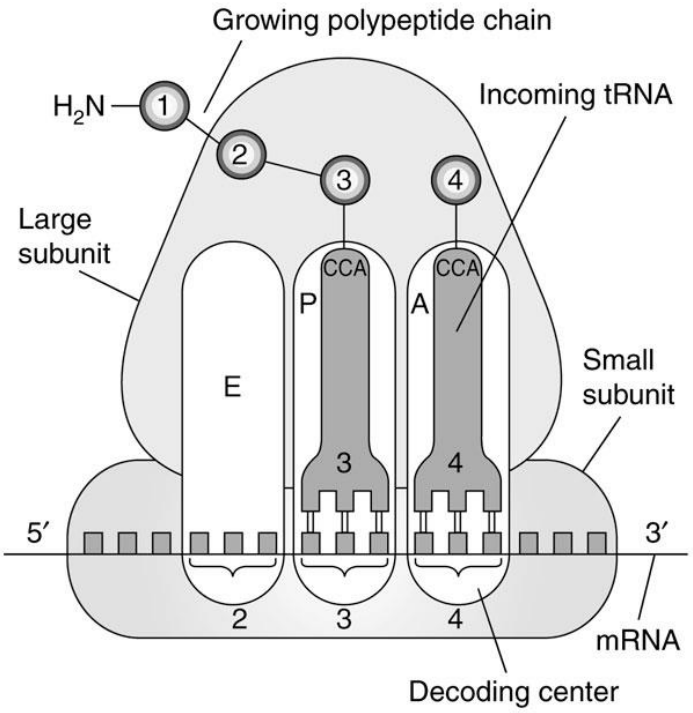


Figure 1.1- The ribosomal complex is comprised of the small (40S) and large (60S) subunit, which both contribute to the arrangement of three binding sites, the aminoacyl (A), peptidyl (P) and exit (E) site. tRNAs are decoded in the A-site (or decoding site) whilst the peptide tRNA is situated in the P-site (figure from (Frank 2003).

1.2 Outline of translation initiation

The ribosomal complex cannot bind to an mRNA transcript that is loaded with secondary structure. The eIF4F complex contains a helicase that binds to the 5' cap of the mRNA and unwinds the transcript ready for ribosome recruitment. It is thought that circularisation of the mRNA may aid initiation, which is mediated by interaction of the poly(A) binding protein (PABP) with the poly(a) tail and cap binding complex (eIF4F), figure 1.2. At this point, the 43S preinitiation complex (PIC) is recruited. The 43S PIC consists of several components, the small ribosomal subunit, several initiation factors (eIFs 1, 1A, 2, 3, and 5) and the ternary complex that anchors methionyl-tRNA (Met-tRNA) to the PIC complex. The 43S PIC binds to the mRNA sequence and the eIF4F complex forming the 48S PIC. As the complex migrates along the 5' untranslated region (UTR) of the mRNA transcript, each codon is 'scanned' as they pass through the P-site of the ribosome (Sonenberg & Hinnebusch 2009). eIF1 stabilises the open conformation of the mRNA binding groove of the small subunit until the start codon (AUG) has been identified. There is a set of mRNAs that avoid this 'scanning' process, by recruiting the PIC to an internal ribosome entry site in close vicinity of the start codon (Sonenberg & Hinnebusch 2009). Once

recognition of the start codon has taken place, GTP hydrolysis is induced resulting in the dissociation of several eIFs, the association of GTP bound eIF5B and the recruitment of the 60S large ribosome, resulting in the formation of the 80S initiation complex. Translation elongation can then begin (Sonenberg & Hinnebusch 2009).

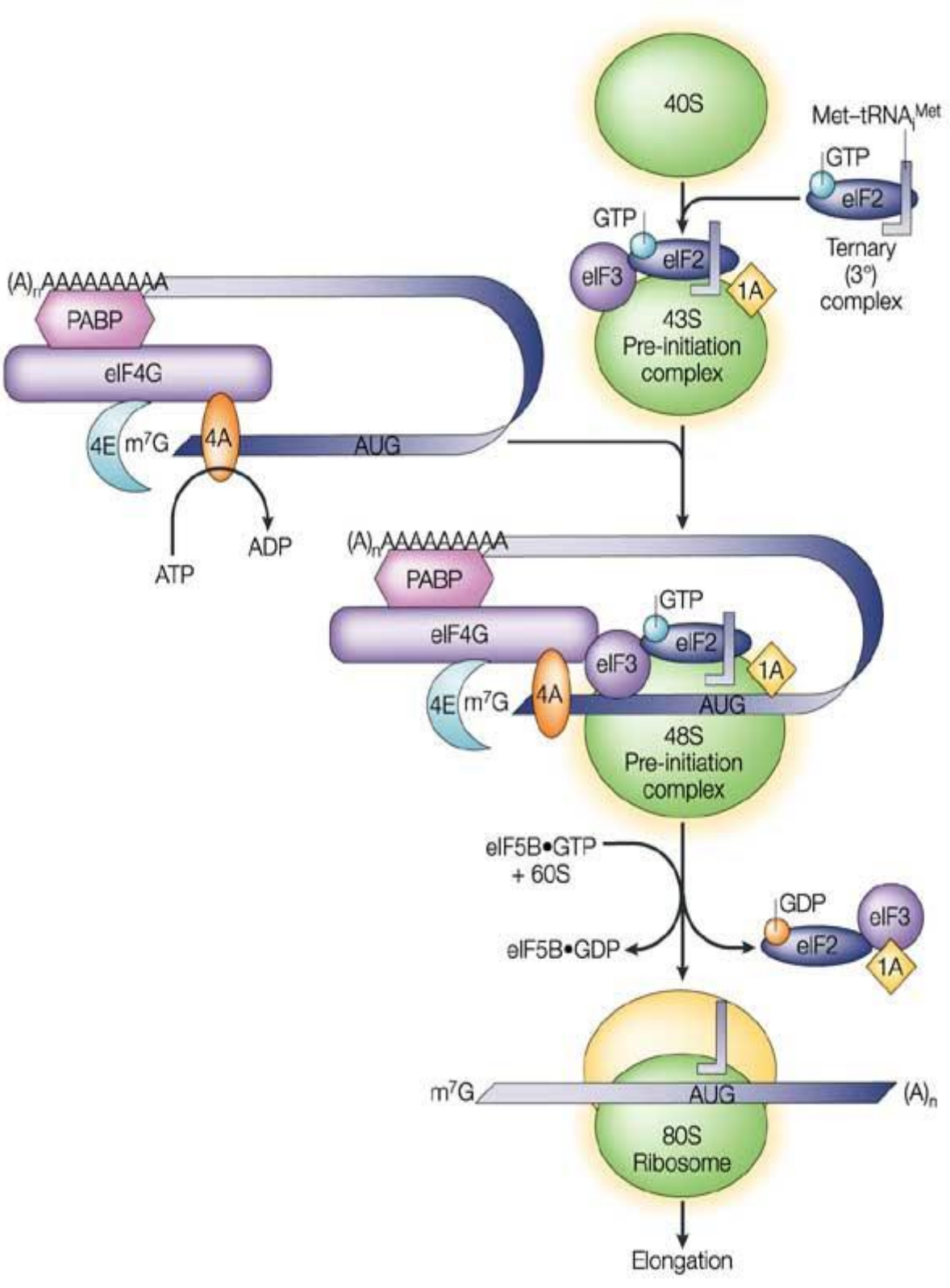


Figure 1.2- Translation initiation. The eIF4F complex (containing helicase eIF4A) and PABP bind to the 5' cap of the mRNA enabling ribosome recruitment and circularisation of the mRNA transcript. The 43S preinitiation complex (PIC) is recruited, comprised of the small ribosomal subunit, initiation factors (eIFs 1, 1A, 2, 3) and the ternary complex. Binding of the 43S PIC

to the mRNA sequence and eI4F complex, results in the 48S PIC which then 'scans' each codon of the 5' UTR until it recognises the start codon. Subsequently, recruitment of GTP bound eIF5B and the 60S large ribosome as well as dissociation of several eIFs, leads to the formation of the 80S initiation complex. At this point elongation can begin (figure from Klann & Thomas 2004).

1.3 Regulation of initiation

There are several points of regulation at the initiation level that act as checkpoints to ensure accurate identification of the start codon and termination of transcript scanning. eIF1 has scanning inhibitory elements in its N-terminal sequence which interact with the Met-tRNA, ensuring efficient halting of the ribosome at the start codon (Graifer & Karpova 2015). eIF1A promotes the release of eIF1 upon start codon identification, closing the mRNA binding groove like a lock and key system (Wilson & Cate 2015).

1.4 The role of tRNAs in translation

1.4.1 tRNA aminoacylation

tRNAs are active RNA molecules (figure 1.4.1) and are not translated into a protein. Following transcription, the tRNA is exported from the nucleus in a carrier-mediated fashion. In the cytoplasm a tRNA is altered by an aminoacyl-tRNA synthetase (aaRS) which is responsible for accurate aminoacylation of tRNAs. aaRSs are known to have several other functions, including cytokine-like activity, mitochondrial RNA splicing and histidine biosynthesis (Martinis et al. 1999). There are two classes of aaRSs based on the structure of the catalytic domains. Class 1 aaRSs are usually monomeric and aminoacylate the terminal adenosine of the tRNA at the 2'-OH, whereas class 2 aaRSs are usually multimeric and aminoacylate the terminal adenosine of the tRNA at the 3'-OH (Hausmann & Ibba 2008). These enzymes also have additional domains, including a tRNA binding domain (tRBD) which is present in human methyl-, lysyl- and valyl-tRNA synthetase, or a domain that binds a protein with a tRBD, such as ArC1p in yeast (Mirande 2010).

aaRSs bind amino acids and catalyse the binding of the amino acid to its cognate tRNA, which triggers a conformational change of the aaRS into a 'closed' state (Strom et al. 2014). The aaRS then activates the amino acid by a condensation reaction forming an aminoacyl adenylate intermediate. This intermediate is then transferred to the 3' acceptor end of the tRNA (Hausmann & Ibba 2008), upon

binding of GTP bound elongation factor 1A (eEF1A), the aminoacyl-tRNA dissociates from the aaRS. eEF1A remains bound to the tRNA until it has been recognised and fully accommodated by the ribosome (Yang et al. 2006; Taylor et al. 2007).

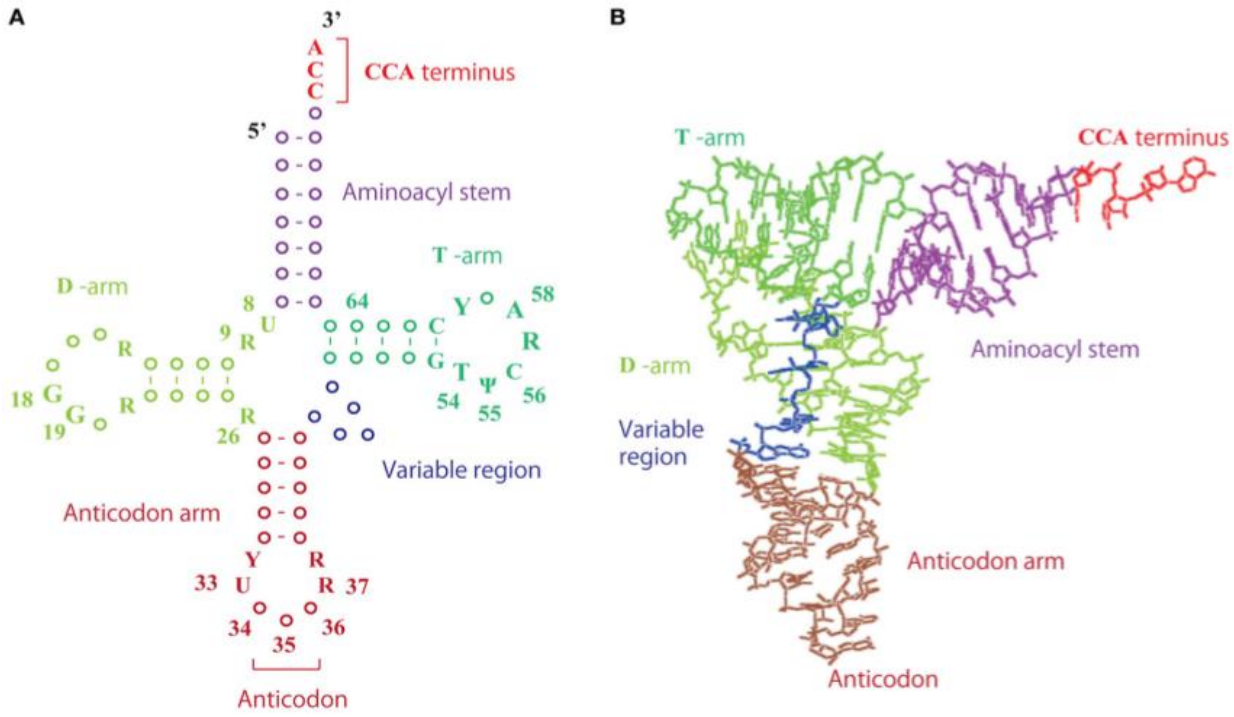


Figure 1.4.1- A) 2D structure and B) 3D structure of a deacyl-tRNA. A tRNA is aminoacylated at the 3' end and the molecule is flexible to aid binding to the ribosome (figure from Hori 2014).

1.4.2 Regulation of tRNA aminoacylation

There are domains within the aaRS that are responsible for the fidelity of tRNA aminoacylation to reduce translational error, by hydrolysing non-cognate or mis-aminoacylated tRNAs. For an aminoacylated tRNA to be transferred from the catalytic core of the aaRS to the editing domain, a channel between the two domains is opened via a conformational change in the protein. Once the aminoacylated end of the tRNA reaches the editing active site, hydrolysis of the tRNA can begin (Mirande 2010; Strom et al. 2014).

1.4.3 Delivery of aminoacyl-tRNAs

There is some ambiguity as to how tRNAs are delivered to the ribosome, it has been suggested that this may occur via random diffusion through the cytoplasm (Fluitt et al. 2007). Alternatively, it has been proposed that the tRNAs are delivered via a chaperoned processive transfer of the tRNA from the nucleus to an aaRS and onto the ribosome. The latter theory would seem more intuitive, but both theories are possible and both can account for the randomization of tRNAs trying to accommodate into the ribosome (Sang Lee et al. 2002; Yang et al. 2006).

1.4.4 tRNA cell populations

The random arrival of aminoacyl-tRNAs at the ribosome requires a selection process between cognate, near-cognate and non-cognate tRNAs. A cognate tRNA is able to form Watson-Crick base pairs with all three bases of the codon, but the tRNA can also form wobble base pairs with the third base (figure 1.4.4). Whereas, near- and non-cognate tRNAs are unable to meet these requirements. A wobble base pair forms between nucleoside 37 of the tRNA to the third codon base, due to post-transcriptional modifications at nucleoside 34 and 37 of the tRNA, (Agris et al. 2007) by tRNA-dependent adenosine deaminases (ADATs) and tRNA-dependent uridine methyltransferases (UMs). These post-transcriptional modifications maintain the correct architecture and stability of the tRNA for binding to the anticodon (Agris et al. 2007). In this way tRNAs are able to decode multiple codons (Quax et al. 2015). There can be as little as 1% of the total tRNA population that are cognate tRNAs, which results in a high level of competition at the ribosome (Rodnina et al. 2005; Novoa et al. 2012). The kinetics of each codon is different due to the varying concentrations of tRNAs in the cell and the differing abilities to form codon:anticodon interactions (Zeng et al. 2014).

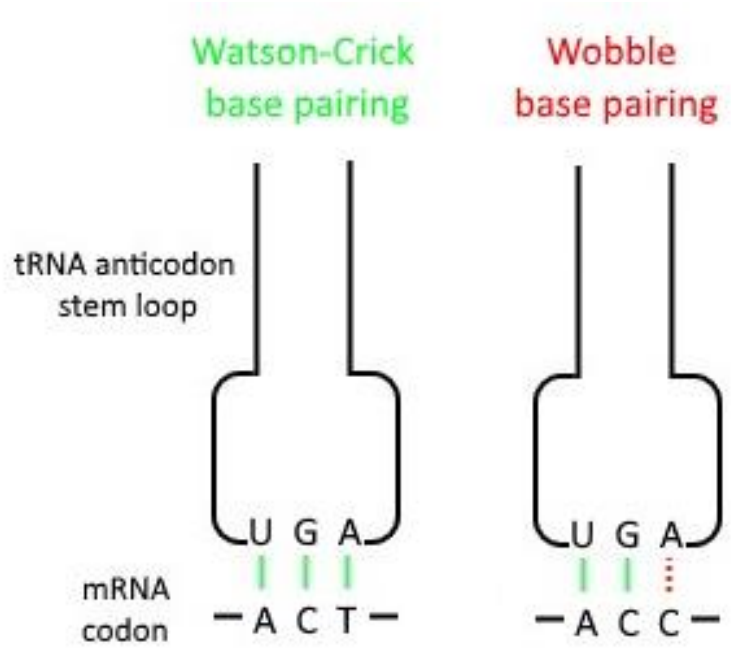


Figure 1.4.4- Decoding of a codon by Waston-Crick base pairing to all three bases (left) and wobble base pairing at the third base of the codon (right)

1.5 Translation elongation

1.5.1 Outline of translation elongation

Elongation is a highly regulated process with a fidelity of 10^{-3}-10^{-4} meaning an error occurs approximately once in every 1000/10,000 amino acid incorporations, whilst polypeptides grows with the addition of approximately 20 amino acids per second (Zeng et al. 2014; Plant et al. 2007). For the accurate incorporation of the correct amino acids into the growing polypeptide, it is essential for aminoacylation, delivery and selection of tRNAs to also be controlled to have a high level of fidelity. Following the accommodation of the aminoacyl-tRNA in the A-site (figure 1.5.1), the peptidyl transferase centre (PTC) catalyses the formation of a peptide bond between the peptidyl- and aminoacyl-tRNA. The growing polypeptide is then quickly transferred from the peptidyl tRNA to the aminoacyl-tRNA. As the polypeptide emerges from the ribosome it begins folding into its functional form. Translation of the rest of the open reading frame (ORF), requires the complex to be translocated by one codon relative to the ribosome (Frank 2003), where the elongation cycle can repeat.

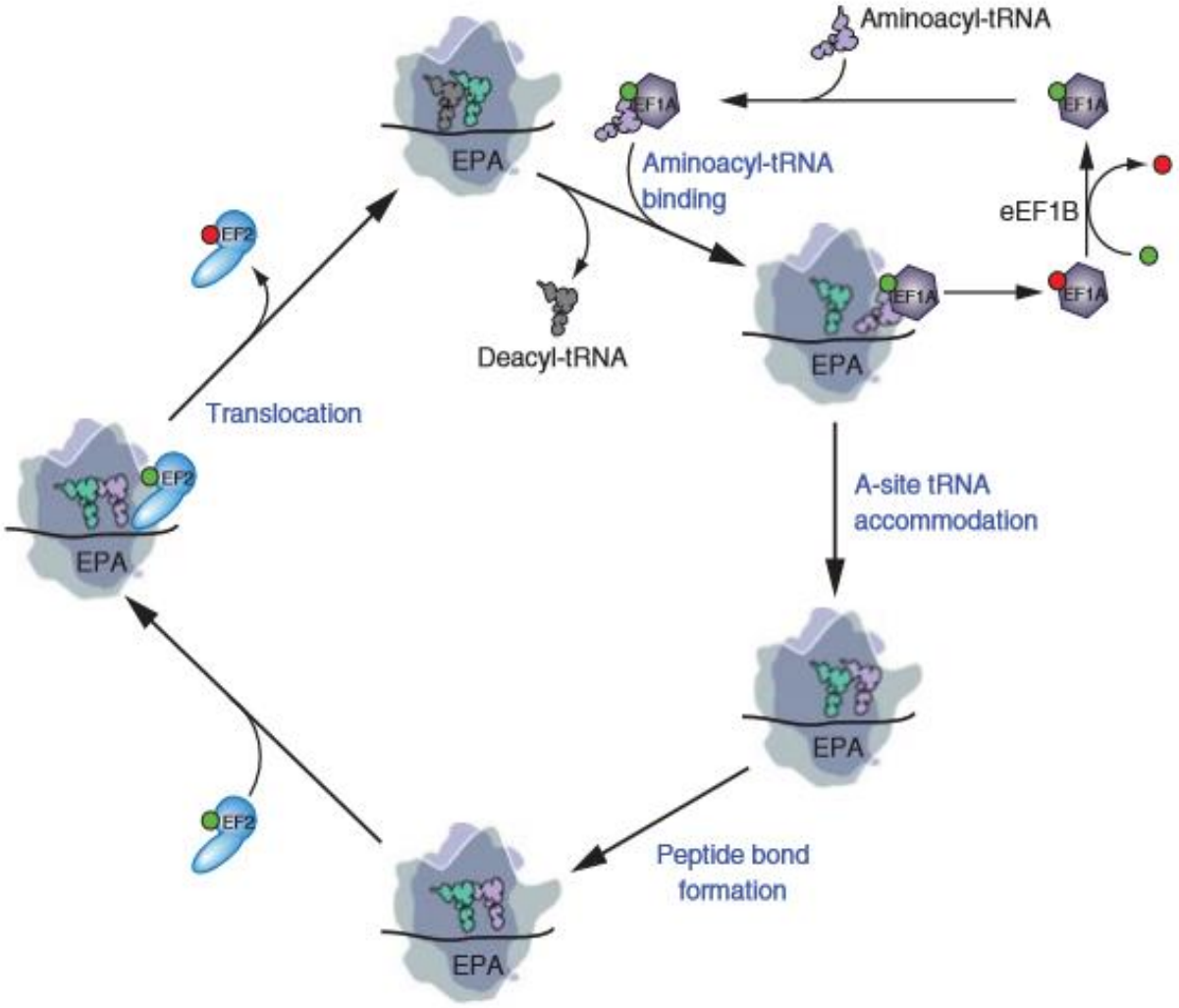


Figure 1.5.1- The elongation cycle. The appropriate aminoacylated tRNA is delivered to the ribosome bound to eEF1A and is accomodated into the A-site. A peptide bond forms between the nascent poypeptide and the new amino acid, transfering the nascent polpeptide to the aminoacyl-tRNA. The ribosome translocates by one codon relative to the ribosome and the next elongation cycle begins (figure from Dever & Green 2015).

1.5.2 The elongation cycle

1.5.2.1 tRNA selection and recognition

A decoding tRNA binds to several components of the ribosome once it has successfully been selected and accommodated into the A-site (Blanchard et al. 2004; Sanbonmatsu et al. 2005). The tRNA binds

to the decoding centre, the GTPase activation centre (GAC) of the large subunit and the PTC (Blanchard et al. 2004; Sanbonmatsu et al. 2005). Selection of a tRNA by a ribosome requires formation of a mini helix structure through base pair interactions between the tRNA anticodon and the mRNA codon. As well as this, it has been observed in bacterial systems that tRNA binding is stabilised by two universally conserved adenine residues, 1492 and 1493, and guanidine 520 of the ribosome forming hydrogen bonds to the tRNA and mRNA backbone, figure 1.5.2.1 (Agris et al. 2007; Taylor et al. 2007). tRNA recognition occurs when A1492 and A1493 flip into the mRNA binding groove of the small subunit (Zeng et al. 2014; Taylor et al. 2007). The interaction between the ribosome, tRNA and mRNA transcript are maintained throughout the elongation cycle monitoring the positioning of the tRNA, which is helped by the flexibility of its backbone (Sanbonmatsu et al. 2005; Blanchard et al. 2004).

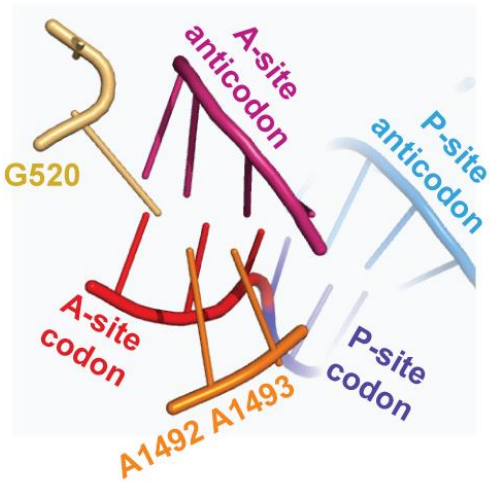


Figure 1.5.2.1- The residues of the ribosomal A-site that stabilise the tRNA binding, A1492, A1493 and G520 form hydrogen bonds with the tRNA and mRNA backbone (figure from Plant et al. (2007)).

1.5.2.2 Cognate vs near-cognate tRNAs

tRNA recruitment is the rate-limiting step in the elongation cycle and recognition of a tRNA is costly for the cell due to the conformational changes and the high levels of GTP required. The energy cost of flipping the adenosine residues is overcome relatively easily by the production of a thermodynamically stable complex with a cognate tRNA. Some near-cognate tRNAs can fully accommodate the A-site and produce a stable complex, but the hydrogen bonding of A1492, A1493 and G520 with the tRNA forms at a slower rate compared to cognate tRNA recognition (Agris et al. 2007; Plant et al. 2007). Some

near-cognate tRNAs can undergo a proportion of the reactions and kinetic processes involved in the recognition process, but are still rejected by the ribosome. If a tRNA fails to form hydrogen bonds with A1492, A1493 or G520 and induce flipping of the adenosine residues, this leads to destabilisation of the codon:anticodon interactions (Zeng et al. 2014). Also, if there is a reduction in the interaction between the aminoacyl terminal of the tRNA and the PTC (Sanbonmatsu et al. 2005), the tRNA fails to fully accommodate the A-site. If a tRNA is unsuccessful it dissociates from the ribosome, still eEF1A bound leaving the ribosome in an 'open' conformation waiting for another tRNA (Valle et al. 2002).

The competition between cognate and near-cognate tRNAs effects the rate of translation, it is generally assumed that an increased ratio of near-cognate to cognate tRNAs decreases the decoding speed of translation as a result of the competition at the ribosome (Chu et al. 2014). There is also thought to be an increase in the error frequency if the abundance of near-cognate tRNAs increases. (Fluitt et al. 2007).

1.5.2.3 tRNA accommodation within the ribosomal A-site

Subsequent to tRNA recognition, a tRNA conformational change is triggered, which activates the GAC. Hydrolysis of eEF1A bound GTP causes the factor to be released from the tRNA and is recycled by the guanine nucleotide exchange factor, eEF1B (Plant et al. 2007; Taylor et al. 2007). This means the aminoacyl terminal of the tRNA is now free to associate with the PTC, at which point the tRNA is said to be fully accommodated into the A-site. A peptide bond is formed by a nucleophilic attack by the amino group of the aminoacyl-tRNA, on the alpha group of the carbonyl carbon of the peptidyl-tRNA. The nascent polypeptide is then transferred from the peptidyl-tRNA to the aminoacyl-tRNA (Wilson & Cate 2015; Valle et al. 2002). The ribosome is put into a 'closed' or active state triggered by a conformational change in the small subunit (Graifer & Karpova 2015).

eIF5A has also been observed to stimulate the transfer of the peptidyl group to the aminoacyl-tRNA amongst numerous other functions in elongation, such as polysome disassembly and promoting the synthesis of glycine and proline containing peptides. This factor was originally thought to be involved in initiation, hence its name, promoting the methionyl-puromycin synthesis pathway (Saini et al. 2009; Li et al. 2010; Mathews & Hershey 2015). However, the molecular mechanisms behind the functions of eIF5A are not yet explicitly clear.

1.5.2.4 Translocation

Subsequent to successful translation of the codon occupying the A-site, the ribosome needs to move along the mRNA transcript to decode the rest of the ORF. This also involves the translocation of the tRNAs into different binding sites of the ribosome, which is restricted by residues of the small subunit (Frank et al. 2007). eEF2 is the major factor involved in this process, following peptide bond formation, a conformational change is triggered in eEF2, by hydrolysis of its bound GTP. eEF2 is a six domain protein that undergoes a hinge like motion between its third and fifth domain, resulting in a ratchet-like movement of the small subunit (Taylor et al. 2007). This subunit movement severs interactions between the decoding centre and the mRNA-tRNA helix, positioning the tRNAs into hybrid A/P and P/E states (Ratje et al. 2010; Frank et al. 2007). The passage between the tRNA binding sites is also widened ready for translocation of the tRNAs. Unlike tRNA selection, there is no free energy change involved in the transition of the tRNAs to the hybrid states, promoting continuation of the process (Blanchard et al. 2004). In the hybrid states, the aminoacyl terminal of the tRNAs interacts with the large subunit of its new binding site, but the anticodon loop of the tRNAs is still positioned in its former binding site of the small subunit (Ratje et al. 2010; Graifer & Karpova 2015). Rotation of the head of the small subunit finally translocates the tRNA anticodon loop into its new binding site. The deacylated-tRNA dissociates from the ribosome and the new peptidyl-tRNA resides in the P-site. Subsequently, dissociation of eEF2 from the ribosomal complex triggers the ribosome to reverse its rachet-like movements ready for the next elongation cycle (Taylor et al. 2007).

1.6 Translation termination and recycling of the translational machinery

Translation of the ORF continues until a stop codon (UAA, UGA or UAG) occupies the A-site, at which point translation is terminated. Comparatively, less protein factors are involved in termination than in initiation or elongation. The stop codon is recognised with the aid of a ternary complex, consisting of the tRNA-shaped release factors eRF1 and GTPase eRF3, figure 1.6 (Sonenberg & Hinnebusch 2009). Upon binding of the ternary complex, hydrolysis of eRF3 bound GTP results in dissociation of the factor from the ribosome (Wilson & Cate 2015). In higher eukaryotes, the release of the polypeptide and subunit dissociation is triggered by the association of the ATP binding cassette sub-family E member 1 (ABCE1/ Rli1 in yeast). It is thought that the energy produced from ATP hydrolysis by ABCE1, is converted into mechanical energy, causing the subunits to dissociate (Dever & Green 2015).

To reduce the cost of elongation to the cell, the translational machinery is recycled but the process is not yet fully understood. Dissociation of the 80S complex releases each component of the termination

complex, which are then free to form initiation complexes (Sonenberg & Hinnebusch 2009). Although, the ribosomal subunits do not always fully dissociate enabling translation of a second ORF via a process termed 'reinitiation' (Pisarev et al. 2011). Recycling of the translational machinery is likely to involve many factors, maybe initiation factors such as eIF1 and eIF1A, which would enable a smooth transition to the next translation cycle (Dever & Green 2015).

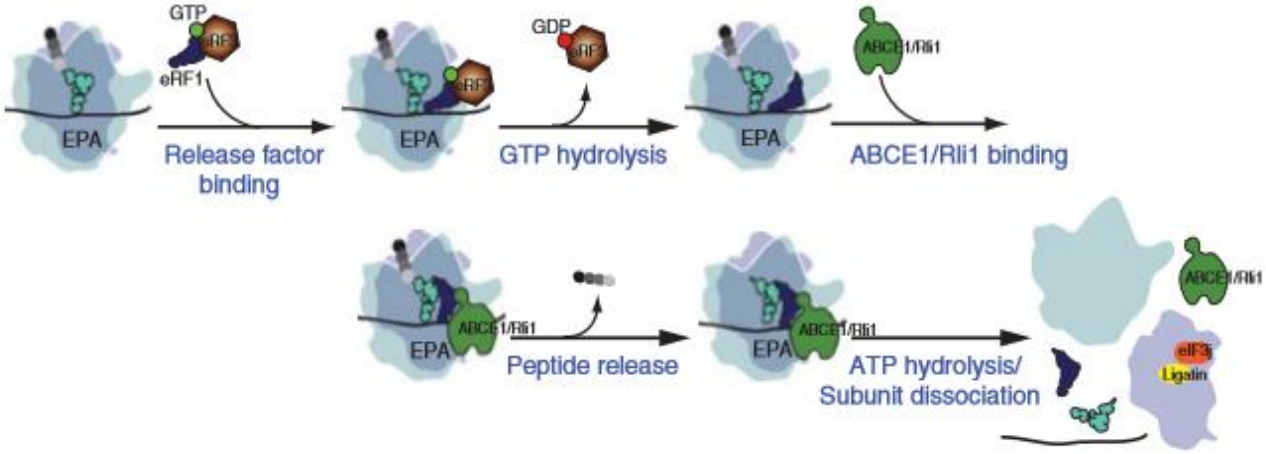


Figure 1.6- Translation termination. The ternary complex (bound eRF1 and eRF3) bind to the A-site of the ribosome and the GTP bound to eRF3 is hydrolysed, which causes dissociation of the factor from the ribosome. Polypeptide and subunit dissociation is caused by binding of ABCE1 (figure from Dever & Green 2015).

1.7 Translational control

There are numerous factors that affect translation efficiency including the abundance of initiation factors, elongation factors, aaRSs, tRNAs and ribosomal subunits. The intracellular levels of the required translational proteins are maintained at a balanced level, such that the process can occur at the maximum rate and efficiency. Enhancing the activity of one individual factor has been shown not to upregulate translation (Firczuk et al. 2013), however, in transformed mammalian cells overexpression of one factor has been shown to increase protein synthesis (Cuesta et al. 2008). This discrepancy could be due to differences in yeast and mammalian systems.

Both initiation and elongation exert a level of control on the rate of translation. In yeast, ribosomes are recruited to the mRNA within 0.2-5 seconds, but the elongation cycle can take between 0.05-1.4 seconds (von der Haar 2008). The point of control can change depending on the factors influencing

the speed of both initiation and elongation (Chu et al. 2014). If initiation is limiting, the rate at which the ribosomes are recruited to the start codon is reduced and there is a reduction in the number of active ribosomes along the mRNA transcript (figure 1.7). On the other hand, if the rate of initiation is not limiting, elongation becomes the limiting process and there is a decrease in the rate of ribosomal clearance from the start codon and a build-up of 'queues' along the mRNA transcript (Zouridis & Hatzimanikatis 2008; Chu et al. 2014). There are several possibilities that may control the decoding speed of the mRNA transcript during elongation, which will be discussed later.

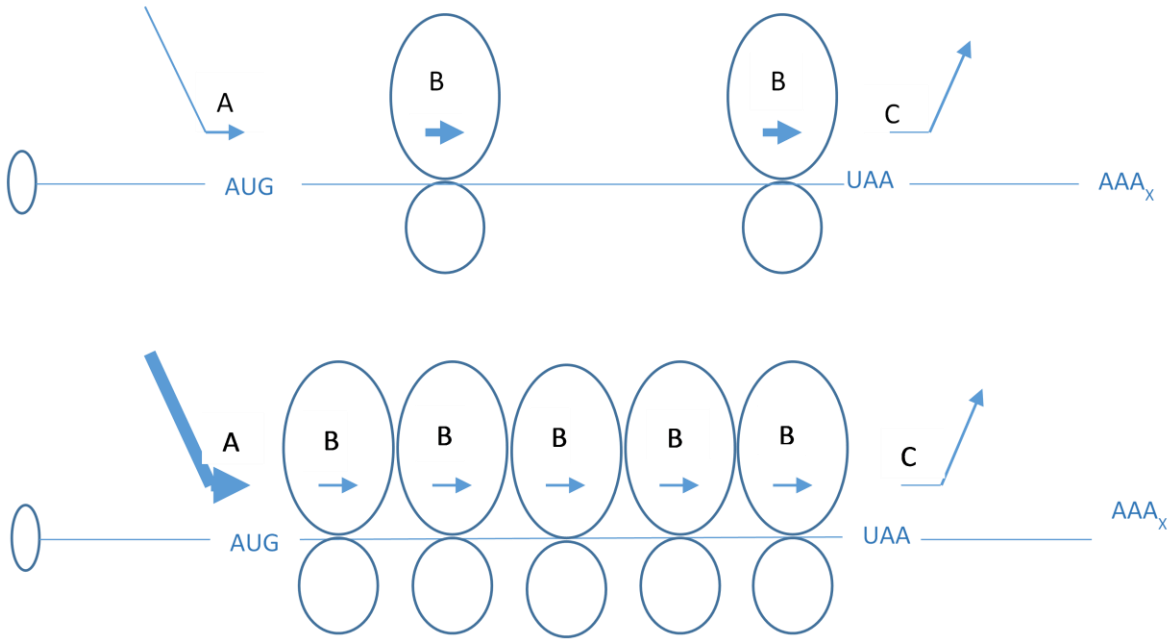


Figure 1.7- The control of the rate of translation. (A) Initiation, (B) elongation and (C) termination. (Top) When initiation is limiting there is a low ribosomal density along the mRNA transcript. (Bottom) Whereas, when the point of control lies at elongation when initiation is not limiting, there is a decrease in ribosomal start clearance and an increase in the density of ribosomes along the mRNA transcript (figure adapted from Plotkin & Kudla 2013).

1.8 Codon Bias

One way in which the system controls the rate and efficiency of translation is through codon usage. Codon usage has many effects on the system on aspects such as protein expression, protein folding and translational accuracy. Codon bias has been widely observed and studies have been able to link codon bias to translational activity, for example it has been observed that synonymous codons decoded by the same tRNA, can have significantly different decoding speeds (Cys^{TGC} and Cys^{TGT} have a

ribosome retention time (RRT) of 1.23 and 0.81, respectively) (Gardin et al. 2014). A delay in protein expression was also found to correlate with de-optimisation of a firefly luciferase construct in *Neurospora crass*a (Yu et al. 2015). However, the mechanisms as to how codon bias exerts its effects are unclear. Codon bias has been linked to genomic regulation of the cell cycle and control of oscillator proteins to name a couple of examples (Novoa & Ribas de Pouplana 2012). On the other hand, there are studies that have not found a correlation between codon bias and protein expression (Pop et al. 2014; Kudla et al. 2009). It is important to keep in mind that codon usage differs between organisms, for example an optimised Firefly luciferase gene for *N. crassa* was expressed in vitro in a *Saccharomyces cerevisiae* cell free extract and had a 2 minute reduction in translational activity (Yu et al. 2015). The difference in codon usage between organisms could possibly explain the discrepancy in observations.

1.8.1 Different forms of codon bias

As well as global codon bias there are different forms of bias at the gene level. Synonymous codon concurrence bias is common in rapidly induced highly expressed genes, such as those linked to stress conditions, where synonymous codons recognised by the same tRNA are clustered in the sequence. This enables particular tRNAs to be quickly recycled locally reducing the time required for tRNA selection (Cannarozzi et al. 2010; Novoa & Ribas de Pouplana 2012). Non-synonymous codon pair bias which is not well understood, is the non-random frequency of particular adjacent nucleotides.

1.8.2 Analysis of codon bias

The effect of codon bias on the efficiency of translation is generally observed through the use of luciferase reporter assays. This type of sensitive quantitative analysis has been used to study protein expression, gene delivery and gene silencing. Luciferases are oxidative enzymes that cause the production of light upon acting on their substrates, a well-known example is firefly luciferase. The gene of interest is fused to a luciferase reporter gene, such that the amount of reporter protein produced is directly correlated with the activity of the luciferase. When excess substrate is added to the reaction the luminescence should be proportional to the abundance of luciferase in the mixture. However, there are many interferences within the cell that can make it difficult to define a specific response, therefore a second ‘control’ reporter is often used to account for non-specific variations and to standardise the experimental reporter data (Branchini et al. 2005; Harger & Dinman 2003; de Wet et al. 1987).

1.8.3 The molecular determinants of ribosome speed

As previously mentioned, there are several determinants that could be responsible for differences in the rate of translation, such as the mRNA concentration, tRNA abundance, ribosome occupancy and ribosome density, but it is unclear as to what are the determinants of decoding speed.

1.8.3.1 Secondary structure

The secondary structure of an mRNA sequence is thought to at least contribute to the translational activity, potentially limiting the movement of ribosomes along the transcript. Several studies have found a decrease in the local sequence structure and folding energy only for the first 40 nucleotides of an mRNA sequence in *Escherichia coli* and *S. cerevisiae*, compared to the rest of the ORF. Randomly generated mRNA sequences have a higher average folding energy than the first 40 nucleotides and the lack of secondary structure near the start site ('ramp site') was probably selected to ensure efficient initiation (Kudla et al. 2009; Tuller et al. 2010). Kudla et al. (2009) correlated this lack of secondary structure in the first 40 nucleotides to the protein expression, whereas Tuller et al. (2010) found no significant correlation with protein abundance. The discrepancy in these results could be explained by different folding energies, but both datasets indicate that the overall secondary structure and codon usage are separate determinants of ribosome speed.

On the other hand, Gorochowski et al. (2015) suggested neither tRNA abundance or mRNA secondary structure is the sole molecular determinant of decoding speed, but that there is a trade-off between the two. Along a transcript of high level of secondary structure there is also increased levels of tRNAs, which could ensure a more uniform elongation rate along a transcript to prevent ribosomal collisions.

1.8.3.2 tRNA abundance and aminoacylation

The challenge to investigate the link between tRNA abundance and decoding speed, lies with quantifying the levels of tRNAs, as this is affected by fluctuating conditions and organisms generally do not express every tRNA species at a given time point (Quax et al. 2015). The tRNA gene copy number is generally used as an estimate of the abundance of tRNAs, as the two have been observed to positively correlate (Kanaya et al. 2001; Percudani et al. 1997). A link between the abundance of tRNAs and codon bias has been observed with a correlation of 95% between the relative synonymous codon usage and tRNA gene copy number, in a large scale study analysing more than 500 genomes (Novoa et al. 2012). Furthermore, taking into account codon usage as an indirect measure of tRNA

availability has been shown to improve predictions of translational activity (Brockmann et al. 2007). Tuller et al. (2010) also found a significant correlation between the local translation efficiency, protein abundance, codon bias and tRNA adaptation index in *S. cerevisiae* and *S. pombe* (Tuller et al. 2010).

Whereas, in silico simulation of fluctuating ratios of aminoacylated tRNAs indicated that only a few tRNA species were shown to have an effect on growth, including tD(GUC), tR(ACG), tE(UUC) and tL(UAG) (Chu et al. 2011). Furthermore, Pop et al. (2014) found no effect on the rate of translation having manipulated the abundance of tRNAs. Many studies have concluded that the tRNA gene copy numbers are a result of, rather than a mechanism for codon bias, having coevolved to maximise the efficiency of translation (Kahali et al. 2007; Lavner & Kotlar 2005).

1.8.4.1 The wider effects of codon bias

1.8.4.2 Protein structure and function

Codon usage along an ORF does not only affect the rate of translation but also co-translational protein folding, which is imperative for protein function. Disordered structures of evolutionary recent genes are preferentially encoded by non-optimal codons, whereas structures such as beta sheets of conserved genes are encoded by optimal codons (Yu et al. 2015; Zhou et al. 2015). At conserved sequences codon bias is at its highest, 1.7 fold higher than expected (Yu et al. 2015; Zhou et al. 2015). The circadian oscillator gene (frequency-frq) in *N. crassa* contains many non-optimal regions, optimisation of the sequence led to an increase in the rate of translation by 42 seconds and an increase in the abundance of FRQ. However, there was a complete inhibition of the circadian rhythm, where FRQ is significant in both the positive and negative feedback loops (Zhou et al. 2013; Zhou et al. 2015). Evidently, the use of 'rare' codons is important for co-translational folding and protein function (Yona et al. 2013) which can also be assessed by the change in susceptibility to enzymatic action (Yu et al. 2015).

1.8.4.3 Translational accuracy

Another effect of codon bias is thought to be a decrease in translational error, with a positive selection pressure for the bias of optimal codons to improve cellular fitness (Quax et al. 2015; Sachs & Liu 2013; Stoletzki & Eyre-Walker 2007; Plotkin & Kudla 2013). If this was the case, we would expect codon bias to be high if selection was acting to reduce missense errors in a long ORF. We would also expect the

strength of codon bias to increase with gene length, if selection was acting to reduce nonsense errors. The longer an ORF the more essential it is to prevent the incorporation of errors as this would waste energy and resources of the cell. A positive correlation has been observed between codon bias and gene length, but codon bias increased weakly along the length of the gene. In this case, selection seems to be more bias towards decreasing the presence of missense errors (Stoletzki & Eyre-Walker 2007). On the contrary, a decrease in error was only observed when nonsense errors were accounted for in a study that analysed 73 genes from 50 species (Shah & Gilchrist 2010).

1.8.5 Origin of codon bias: natural selection or mutational biases?

There are two major theories on the origin of codon bias, being as a result of natural selection or a neutral process such as mutational bias. These theories are not thought as mutually exclusive events and it could be a balance between the two that was responsible for shaping codon usage (Plotkin & Kudla 2013; Quax et al. 2015). A combination of the two, the selection-mutation-drift theory states that there is codon bias due to mutational selection of G and C bases in sequences that have high recombination rates. Indeed, the frequency of optimal codons ending in G or C in *Caenorhabditis elegans* and *Drosophila melanogaster* positively correlates with the rate of recombination (Rocha 2004; Marais et al. 2001). Additional supporting evidence for this has also been observed in yeast (Gerton et al. 2000) and prokaryotes (Palidwor et al. 2010).

1.9 Summary of objectives

As discussed, codon bias has been widely observed and different decoding speeds between codons are thought to be one of the reasons for the observed bias. However, there is still contradictory evidence as to what molecular determinant(s) govern the differences in decoding speed. In this study, we focus on investigating whether predictions of protein expression can be made based on tRNA abundance. We compared reporter protein expression levels in two species of yeast, *S. cerevisiae* and *S. uvarum,* that differed in their tRNA content. We also perturbed the tRNA pool of *S. cerevisiae* by individually deleting tRNA genes and overexpressing individual tRNAs. A better understanding of the mechanism(s) underpinning codon bias could have positive implications in the biotechnological field, having the potential to improve production rates and yields of recombinant proteins.

Chapter 2: Materials and methods

2.1 Strains of *E. coli* and *S. cerevisiae*

Table 2.1- Genotype of the strains used in this study. The following *S. cerevisiae* tRNA knockdown strains ΔtT(AGU)B, tT(AGU)H, ΔtT(UGU)G1 and ΔtT(UGU)P were provided by Dr Daniela Delneri and Dr Ray O'Keefe, University of Manchester.

***E. coli* strains**	**Genotype**
T10 *E. coli*	genotype F– *mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(ara leu) 7697 galU galK rpsL (StrR) endA1 nupG.*
***Saccharomyces* strains**	**Genotype**
BY4741 (Brachmann et al. 1998)	*MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*
S. uvarum (Scannell et al. 2011)	*MATα hoΔ:loxP his3 lys2 ura3*
ΔtT(AGU)B	*MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 ΔtT(AGU)B*
ΔtT(AGU)H	*MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 ΔtT(AGU)H*
ΔtT(UGU)G1	*MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 ΔtT(UGU)G1*
ΔtT(UGU)P	*MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 ΔtT(UGU)P*

2.2 Growth conditions

Transformants of BY4741 were selected by the uracil (*URA3*) (table 2.2) marker and grown at 30°C and those of *E. col*i were selected by the ampicillin resistance (ampR) marker and grown at 37°C. BY4741 and the knock down tRNA strains were grown in liquid or on solid YPD/ YPD inoculated with geneticin disulphate salt (G418, A1720 Sigma) for 2-4 days at 30°C. The media composition is shown in table 2.2.A and was autoclaved before use.

Table 2.2.A- Media composition

Media	Composition (w/v)
LB + ampicillin	1% tryptone (211705, Becton, Dickinson and company (BD)) 0.5% yeast extract (212750, BD) 1% NaCl (S/3160/60, Fisher scientific) ampicillin (100mg/ml) (2% agar (214530, BD))
-ura selective media	2% glucose (G/0500/61, Fisher scientific) 0.67% Yeast nitrogen base (291940, BD) 0.19% synthetic complete mixture drop-out:-ura (DSCK1009, Formedium LTD) (2% agar (214530, BD))
YPD	2% glucose (G/0500/61, Fisher scientific) 1% yeast extract (212750, BD) 2% bactopeptone (211677, BD) (+Geneticin 200µg/ml) (2% agar (214530, BD))

2.3 Cloning

2.3.1 Polymerase chain reaction (PCR)

An optimised version of *Renilla* luciferase (*rluc*) was cloned from pTH818 using the forward oligonucleotides, (ACA)10_max, (ACC)10_max, (ACG)10_max, (ACT)10_max and C_maxR (table 2.3.1.A) and M13r as the reverse oligo, to produce the following constructs, ACA^{10}_Rluc, ACC^{10}_Rluc, ACG^{10}_Rluc, ACT^{10}_Rluc and maxRluc. All oligos were synthesised by Eurofins MWG operon.

Table 2.3.1.A- The oligonucleotides used to clone an optimised *rluc* gene to produce the ACA^{10}_Rluc, ACC^{10}_Rluc, ACG^{10}_Rluc, ACT^{10}_Rluc and maxRLuc constructs.

Construct	**Forward oligonucleotide**	**Sequence 5'- 3'**
ACA^{10}_Rluc	(ACA)10_max	GCGCGCCCCGGGATGACAACAACAACAACAACAACAACAACAACAA CTTCCAAAGTCTACGACCCGGAAC
ACC^{10}_Rluc	(ACC)10_max	GCGCGCCCCGGGATGACCACCACCACCACCACCACCACCACCAC TTCCAAAGTCTACGACCCGGAAC
ACG^{10}_Rluc	(ACG)10_max	GCGCGCCCCGGGATGACGACGACGACGACGACGACGACGACGA CTTCCAAAGTCTACGACCCGGAAC
ACT^{10}_Rluc	(ACT)10_max	GCGCGCCCCGGGATGACTACTACTACTACTACTACTACTACTACT TCCAAAGTCTACGACCCGGAAC
maxRluc	C_maxR	GCGCGCCCCGGGATGACTTCCAAAGTCTACGACCCGGAAC
Reverse oligonucleotide		**Sequence 5'-3'**
M13r		AGCGGATAACAATTTCACACAGGA

The polymerase chain reaction (PCR, table 2.3.1.B) was performed in a reaction volume of 100µl consisting of 20µl 5x green GoTaq Flexi buffer (M8918, Promega), 10µl dNTPs 2mM (diluted #N0447S, New England Biolabs (NEB)), 5µl $MgCl_2$ 25mM (A351B, Promega) 1µl of 5' and 3' oligos, 1µl template DNA, 1µl GoTaq G2 Flexi DNA polymerase 5u/µl (M780A, Promega) and 61µl distilled H_2O (dH_2O).

Table 2.3.1.B- PCR programs for the cloning of *rluc*, tT(AGU) and tT(UGU).

	Rluc	**tT(AGU)**	**tT(UGU)**
Cycles	30	30	30
Initialization	94°C, 2 mins	94°C, 2 mins	94°C, 2 mins
Denaturation	94°C, 45 secs	94°C, 45 secs	94°C, 45 secs
Annealing	51°C, 45 secs	51°C, 45 secs	55°C, 45 secs
Extension	72°C, 1 min 10 secs	72°C, 45 secs	72°C, 45 secs
Final extension	72°C, 10 mins	72°C, 10 mins	72°C, 10 mins
Total run time	1 hour 55 minutes	1 hour 55 minutes	1 hour 55 minutes

The tRNA genes tT(AGU) and tT(UGU) were cloned from 2µl of BY4741 genomic DNA using oligos as shown in table 2.3.1.C and the PCR programs are shown in table 2.3.1.B.

Table 2.3.1.C- Oligonucleotides to clone tT(AGU) and tT(UGU)

Oligonucleotide	**Sequence 5'-3'**
tT(AGU)H_f	GGCCGGCTCGAGGTTCAGAAGAGCCCAAGTATGTAATTATTTTTTGC
tT(AGU)H_r	CGCGCGCCCGGGAGTTCTTTAGAGAGCTTGCTCTTGTACG
tT(UGU)H_f	GGCCGGCTCGAGCTCATTCTCGCATTCCAACAGTTATG
tT(UGU)H_r	CGCGCGCCCGGGTTTCTCACTTGTCAACTATATGTTTTTAG

2.3.2 DNA Sequencing

Plasmid DNA was sent to Eurofins MWG operon for sequencing. The ACA^{10}_Rluc, ACC^{10}_Rluc, ACG^{10}_Rluc, ACT^{10}_Rluc and maxRluc constructs were sequenced using the PADH1Seq and M13r oligos (table 2.3.2). Sequencing of tT(AGU) and tT(UGU) utilised oligonucleotides M13f and M13r.

Table 2.3.2- The PADH1seq and M13r oligonucleotides to sequence the ACA^{10}_Rluc, ACC^{10}_Rluc, ACG^{10}_Rluc, ACT^{10}_Rluc and maxRluc constructs. The M13f and M13r oligonucleotides to sequence tT(AGU) and tT(UGU).

Oligonucleotide	**Sequence 5'-3'**
PADH1Seq	TTTTGTTTCCTCGTCATTGTTCTCGTTCCC
M13f	CGCCAGGGTTTTCCCAGTCACGAC
M13r	AGCGGATAACAATTTCACACAGGA

2.3.3 Restriction enzyme digest

The cloned *rluc* genes, the tRNA genes (tT(AGU) and tT(UGU)) and plasmids pTH727, pRS(313) and pRS(423) were subjected to digest by restriction enzymes *Xma*I (#R01806, NEB) and *Xho*I (#R01465, NEB). Plasmids pRS313 and pRS423 as well as pTH485 containing tT(CGU) underwent digestion by BamHI-HF (#R3136S, NEB). Each reaction mixture was composed of 4µl cutsmart buffer (#B72045, NEB), 1µl of restriction enzyme, 15µl DNA (5µl for the site directed fragment) and 20µl dH_2O and incubated at 37°C for 2 hours. tT(CGU), pRS313 and pRS423 were then incubated with 2µl rAPID alkaline phosphatase (1U/µl, Roche) and 5µl 10X rAPID alkaline phosphatase buffer (Roche) for 30 minutes at 37°C.

2.3.4 DNA purification

As per manufacturer's instructions using the GeneJET PCR purification kit (#K0702, Thermo scientific).

2.3.5 Ligation

The reaction mixture consisted of calculated volumes of insert and vector DNA, then made up to 10µl with dH_2O using the following equations:

$$\frac{50}{concentration\ of\ DNA\ (\frac{ng}{\mu l})} = volume\ of\ vector\ DNA\ (\mu l)$$

$$\frac{150x(\frac{length\ of\ insert\ DNA\ (bp)}{length\ of\ vector\ DNA\ (bp)})}{concentration\ of\ insert\ DNA\ (\frac{ng}{\mu l})} = volume\ of\ insert\ DNA\ (\mu l)$$

Subsequently, 10µl of 2X Quick ligase reaction buffer (#B2200S, NEB) and 1µl Quick ligase (#M2200S, NEB) was added and thoroughly mixed. For the control reaction, dH_2O was used instead of insert DNA. Each mixture was centrifuged briefly and incubated at room temperature for 5 minutes and kept on ice.

2.3.6 DNA agarose gel electrophoresis

A 1% agarose (MB1200, Melford Biolaboratories LTD) gel was made as per Sambrook & Russell (2001) for all analysis except in the detection of tRNA genes which required a 2% gel. Ethidium bromide staining (E1510, Sigma) was added to the gel preparation at 2% and the gel was run at 70 volts/ 200 amps.

2.3.7 Gel extraction

As per manufacturer's instructions using the GeneJET Gel extraction kit (#K0692, Thermo scientific).

2.4 Genetic methods for *E. coli*

2.4.1 Production of competent cells

Firstly, the SOB medium and CCM80 buffer were prepared in an autoclaved conical and 1L flask, respectively. The SOB medium consisted of 1.25g yeast extract, 5g tryptone, 0.146g NaCl, 0.05g KCl, 0.6g $MgSO_4$ and 250 ml water, autoclaved and stored at room temperature. The CCMB80 buffer (0.98g KOAc (10mM), 11.8g $CaCl_2*2H_2O$ (80mM), 4.0g $MnCl_2*4H_2O$ (20mM), 2g $MgCl_2*6H_2O$ (10mM), 100g glycerol (10%) and the addition of water up to 1 litre) was pH adjusted to 6.4, filter sterilised and stored at 4°C.

A vial of seed stock of T10 *E. coli* cells was inoculated into the SOB media and grown at 30°C to an oD_{600} of 0.3, measuring regularly. Whilst the cells were growing the CCMB80 buffer was placed on ice. Once the cells reached the specified OD, they were centrifuged for 10 minutes at 3000rpm, to obtain a cell pellet. The cells were resuspended in 80 ml of ice cold CCMB80 buffer and incubated on ice for 20 minutes. The mixture was then centrifuged again to obtain a cell pellet once more, which was then resuspended in 10ml of ice cold CCMB80 buffer. The cell mixture was aliquoted into microcentrifuge tubes and frozen at -80°C.

2.4.2 Transformation of *E. coli* cells

Plasmid DNA was chilled on ice and 9µl of ligation reaction or 1µl of plasmid was mixed with 100µl of competent *E. coli* cells. Following a 20 minute incubation on ice, the reaction was subject to a 42°C 60 second heat shock and then immediately placed back on ice until the next step. To the cells, 1ml LB was then added, followed by 30 minutes shaking at 37°C. For transformation of a ligation reaction, the mixture was centrifuged for 5 minutes at 3000rpm. 900µl of the supernatant was removed and the pellet was resuspended in the last 100µl of supernatant. All of the cells were then plated onto LB containing ampicillin agar plates and incubated at 37°C overnight. For transformation of an existing plasmid 100µl of the LB mixture was plated.

2.4.3 DNA miniprep of *E. coli* cells

Overnight cultures of transformed *E. coli* cells grown at 37°C in LB ampicillin broth were subject to GeneJET plasmid miniprep Kit (#K0503, Thermo scientific) as per manufacturer's instructions.

2.5 Genetic methods for *S. cerevisiae* cells

2.5.1 Obtaining yeast genomic DNA

Two OD_{600} units from an overnight culture of BY4741 were harvested and centrifuged for 1 minute at 12000rpm, disregarding the supernatant. The cell pellet was resuspended in 200µl of lysis buffer (170µl sorbitol (>98% purity, sigma), 20µl Na_3PO_4 (sigma) and 10µl lyticase (L2524, sigma)) and incubated at 37°C for 20 minutes, followed by cycling of boiling and freezing (95°C for 15 minutes,-80°C for 15 minutes and 95°C for 15 minutes). The cells were then centrifuged for 5 minutes at 12000rpm and the supernatant obtained.

2.5.2 Obtaining total RNA from yeast cells

Two OD_{600} units from an overnight culture of BY4741 were inoculated into 10ml of –ura liquid media and grown to an OD_{600} of 0.6-0.8. From this culture 2 OD_{600} units were harvested and centrifuged at 4000rpm for 5 minutes. All of the supernatant was removed and the cell pellets frozen at -80°C. RNA extraction was performed using Qiagen RNeasy mini kit (ref 74104) with the on column *DNase*I digestion and the final RNA yield was not diluted.

2.5.3 Quantitative Real Time-PCR (qRT-PCR)

To amplify *rluc* from ACA^{10}_Rluc, ACC^{10}_Rluc, ACG^{10}_Rluc, ACT^{10}_Rluc and maxRluc RNA cell extracts, the Quantifast SYBR Green RT-PCR protocol (Qiagen) was followed using 1µl of the oligos qmaxR_f and qmaxR_r and 11.75µl of a 25x diluted RNA sample. To amplify the control gene, firefly luciferase (stacFluc), oligos qFLucsf and qFlucsr were used (table 2.5.2.A). The reaction was performed in a Bio-Rad CFX connect optics module real-time system using a two step-cycling program (table 2.5.2.B), and a melt curve was obtained.

Table 2.5.2.A- Oligonucleotides qmaxR_f and qmaxR_r to clone *rluc* from RNA extracts obtained from cells transformed with the ACA[10]_Rluc, ACC[10]_Rluc, ACG[10]_Rluc, ACT[10]_Rluc and maxRluc constructs.

Oligonucleotide	**Sequence 5'-3'**
qmaxR_f	GAAGAATTTGCCCGCTACTT
qmaxR_r	ACCTTTGACCAACGGAATTT
qFLucsf	TGCAAGCTTTGGACTTCTTC
qFlucsr	CAAGGTAGACAAGCCGACAA

Table 2.5.2.B- Quantitative real-time PCR twostep cycling program to amplify *rluc.*

Step	**Time**	**Temperature (°C)**
Reverse transcription	10 min	50
PCR activation step	5 min	95
Two step cycling		
39 cycles		
Denaturation	10 secs	95
Combined annealing and extension	30 secs	60

To analyse the data Δct (cycle threshold- number of cycles required for the fluorescent signal to cross the threshold), ΔΔct (as below) and $2^{-\Delta\Delta ct}$ were calculated.

$$\Delta\text{ct} = \frac{rluc\ ct}{fluc\ ct}$$

$$\Delta\Delta\text{ct} = \frac{\Delta\text{ct}\ rluc}{\Delta\text{ct maxRluc}}$$

2.5.4 Transformation of yeast cells

Firstly, single stranded DNA 10mg/ml was heated for 10 minutes at 95°C and placed on ice immediately. From an overnight yeast culture, 1 ml was centrifuged for 30 seconds and the supernatant removed. To the cell pellet the following reagents were added in order, 240µl of 50% PEG

(sigma), 36µl 1M Lithium acetate (sigma), 10µl ssDNA, 2.5µl β-mercaptoethanol (M6250, sigma), 2µl plasmid DNA and 69.5µl dH_2O, this mix was then vortexed for 1 minute. The reaction was incubated at room temperature for 20 minutes followed by incubation at 42°C for 20 minutes. A cell pellet was then obtained via centrifugation at 2000rpm for 2 minutes. The supernatant was removed and the cells resuspended in 200µl sterile dH_2O. All of the cells were plated onto –ura agar plates and incubated at 30°C for 2-4 days.

2.6 Molecular biology techniques for *S. cerevisiae*

2.6.1 Growth analysis

In a 24-well suspension culture plate, 1 ml of –ura media was added to each well inoculated with culture to 0.1 OD_{600} units. The growth of the culture was measured until stationary growth phase at 600nm in a BMG labtech SPECTROstar^{nano}. The absorbance was measured in 30minute cycles following 29 minutes of double orbital shaking at 400rpm. The mean doubling time, and percentage growth rate was calculated (as below), the specific growth rate was calculated to work out the percentage growth compared to the control. The exponential value of growth was obtained from a logarithmic graph of the optical density.

$$Mean\ doubling\ time = \frac{Ln2}{exponential\ value}$$

$$Specific\ growth\ rate\ (h^{-1}) = \frac{1}{doubling\ time} \times 60$$

$$\%\ growth\ \frac{specific\ growth\ rate\ of\ maxRluc}{specfific\ growth\ rate\ of\ experimental\ construct}$$

2.6.2 Dual-reporter luciferase assay

Single cell colonies were grown in 150µl of –ura media overnight at 30°C in a 96 well microtitre cell culture plate. The following morning, 10µl of the culture was transferred into a fresh media plate as above and incubated for 3-4 hours. Using reagents from the Dual-Glo Luciferase Assay system (ref

9E2940, Promega) 20µl of Passive Lysis Buffer, 20µl of culture and 40µl of luciferase assay reagent was added to the reader plate. The *firefly luciferase* fluorescence values at 600nm were obtained by the Omega BMG fluostar following 10 minutes of incubation at 30°C. Before each well reading, there was a 20 second orbital shake of 500rpm with a settling time of 0.2 seconds.

To the same reader plate, Dual-Glo stop and Glo buffer and Dual-Glo Stop and Glo substrate were added as a mix, calculated as follows:

$$\text{Number of wells} \times 40 = \text{A}$$

$$\text{A} + (\text{A} \times 0.15) = \text{Volume of Dual Glo stop and Glo buffer (B)}$$

$$\frac{(B)}{100} = \text{Volume of Dual Glo Stop and Glo substrate}$$

To each well 40µl of the buffer and substrate mix was added followed by a 10 minute incubation at 30°C, before readings for the Rluc fluorescence at 600nm were taken.

2.6.3 Alkaline protein extraction

An overnight culture was used to inoculate 25ml of –ura media to an OD_{600} of 0.2. The culture was incubated for approximately 4 hours to reach exponential growth at an OD_{600} of 0.6-0.8. A total of 5 OD_{600} units were extracted and centrifuged at 3000rpm for 5 minutes and the supernatant disregarded. The cell pellet was resuspended in 1ml of ice cold dH_2O and centrifuged for 30 seconds at 12000rpm, removing the supernatant. The cell pellet was resuspended in 200µl lysis buffer ((100mM NAOH (S/4880, FSA Laboratory supplies), 50mM EDTA (D/0700/53, Fisher scientific), 2% SDS (S1030, Melford biolabs) and 2% β-mercaptoethanol (M6250, sigma)) and heated for 10 minutes at 95°C. To the mixture 5µl of 4M acetic acid (A/0400/PB17, Fisher scientific) and 50µl Blue buffer (80mM Tris pH 6.8 (T/P630/60, Fisher scientific) 20% glycerol (G/0650/17, Fisher scientific), 2% SDS (S1030, Melford biolabs) and 0.05% bromophenol blue (B/4630/46, Fisher scientific), supplemented just before use with 5% β-mercaptopethanol (M6250, sigma)) was added and thoroughly vortexed. The protein extracts were stored at -20°C for short term storage, or -80°C for long-term storage.

2.7 Protein analyses

2.7.1 SDS polyacrylamide (SDS-PAGE) gel electrophoresis

The components to produce a 12.5% resolving gel and stacking gel are shown in table 2.7.1, the gel and electrophoresis components were set up according to Sambrook & Russell (2001) using 10X running buffer (30g Tris-Base (T/P630/60, Fisher scientific), 144g glycine (G8898, sigma) and 15g SDS (S1030, Melford Biolaboratories)). Sample were loaded onto the gel and run at 180V until the samples had resolved sufficiently.

Table 2.7.1- Media composition of the resolving and stacking gel for SDS-PAGE electrophoresis.

Component	**Resolving gel (12.5%)**	**Stacking gel**
30% acrylamide:bisacrylamide 29:1 (#161-0156, Bio-Rad)	4.5ml	1ml
4X lower tris (1.5M Tris pH8.8 (T/P630/60, Fisher scientific), 0.4% SDS (S1030, Melford biolabs)	2.7ml	-
4X upper tris (1M Tris pH6.8 (T/P630/60, Fisher scientific), 0.4% SDS (S1030, Melford biolabs)	-	1.75ml
dH_2O	3.6ml	4.2
Ammonium persulphate (40% ammonium persulphate (A3678, sigma) in H_2O)	40µl	40µl
TEMED (N,N,N',N'-Tetramethylethylenediamine (T8133, sigma))	5µl	5µl

2.7.2 Western blot

The blot was assembled according to Sambrook & Russell (2001) and run at 9V, 200mA for 30 minutes. Following this, the PVDF membrane was incubated in TBS-M ((10ml of 10X TBS (500mM Tris pH 7.8 (T/P630/60, Fisher scientific), 1.5M NaCl (S/3160/60, Fisher scientific), 50µl Tween 20 (P1379, sigma) and H_2O up to 100ml)) for 10 minutes. The primary anti-Rluc polyclonal antibody (Caltag Medsystems)

was diluted by 1 in 1000 in TBS-M and was incubated with the membrane overnight at 4°C. The membrane was then rinsed in TBS-M and incubated with the secondary antibody, anti-rabbit IgG-HRP ((F9887, sigma) diluted by 1 in 10,000 in TBS-M), for 1 hour at room temperature. The membrane was washed four times with TBS-T for a total of 30 minutes.

2.7.3 ECL detection method

To image the western blot, solutions 1 (100µl Luminol (09253, sigma), 44µl Coumaric acid (HPLC grade, sigma), 1ml Tris pH 8.5 (T/P630/60, Fisher scientific) with the addition of water to 10ml) and solution 2 (6.4 µl hydrogen peroxide (H1009, sigma), 1ml Tris pH 8.5 (T/P630/60, Fisher scientific) with the addition of water to 10ml) were made immediately prior to use. The solutions were mixed and incubated with the membrane for one minute. The membrane was then placed into the imager and exposed to UV for a series from 30 seconds to 2 minutes to image the membrane.

2.8 Secondary structure prediction

The RNAfold server (Lorenz et al. 2011) from ViennaRNA Web Services was used to predict the free energy of the thermodynamic ensemble and the centroid secondary structure of the *rluc* RNA.

2.10 Statistics

To analyse data for significant differences, which was classed at a 95% confidence level, two sample t-tests and one-way analysis of variance (ANOVA) followed by a Tukey's 95% analysis was performed.

Chapter 3: Results

3.1 Design of the reporter constructs

To begin to delineate the mechanism(s) behind elongation control of translation and to assess the effect of tRNA levels on the decoding speed of mRNA transcripts, a dual reporter luciferase expression system was created. Four *rluc* reporters were constructed such that the level of Rluc expression could be directly linked to the decoding speed of a particular set of codons. This was also achieved by optimising the gene sequences for efficient translation and situating the genes of interest, under the control of constitutive promoters, resulting in a high level of gene expression. In four of the five experimental reporters generated, ten codon repeats of a particular codon were inserted by PCR, at the beginning of the ORF of an optimised version of *rluc* (maxRluc), figure 3.1. The use of a codon optimised version of *rluc* provided efficient translation of the sequences such that, the level of protein production by the reporters would only be governed by the decoding speed of the codon repeats. The control reporter, consisted of maxRluc with no additional codons. To normalise the experimental data a non-optimised firefly luciferase (staCFluc) reporter was also expressed in the luciferase system, to effectively compare the *rluc* reporter expression levels. It has been observed that optimised versions of *fluc* can cause recombination problems in bacteria hence the use of a non-optimised sequence (TvdH-personal communication). This version of *fluc* also had a truncated terminal to maintain its presence in the cytosol for efficient expression.

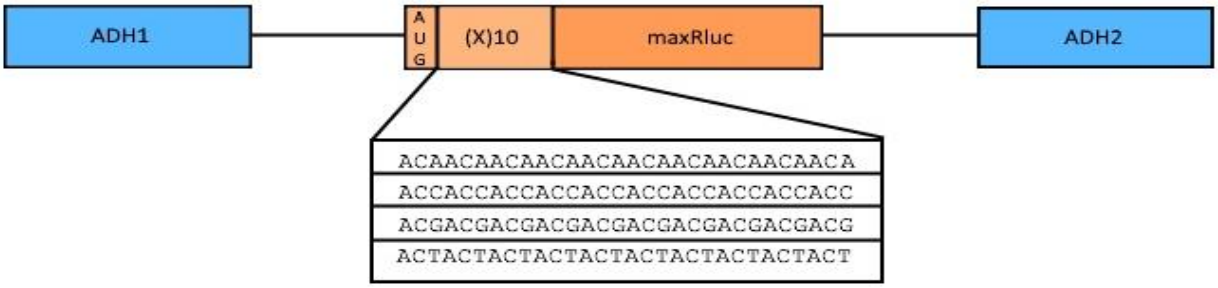


Figure 3.1- Schematic of four *rluc* reporter constructs where ten codon repeats of ACA, ACC, ACG and ACT were inserted at position 2 of the coding region. *rluc* is under the control of the ADH1 promoter and ADH2 terminator.

3.2 Secondary structure analyses of the reporter constructs

Using the aforementioned approach, introducing codon runs into the experimental reporters could have resulted in the formation of secondary structure, which itself could alter the decoding speed. However, computational analysis published by Letzring et al. (2010), provided secondary structure data for firefly luciferase reporters, also with the introduction of 10 codon repeats. This was

performed for 59 of the 61 sense codons in the genetic code. The predicted free energy of the RNA secondary structure for the first 50 nucleotides of each reporter was published. The data showed that threonine was the only amino acids whose majority of synonymous codons (ACA, ACC, ACG and ACT) did not have a strong tendency to form secondary structures, lending these threonine codons as the best candidates for the reporters in this study.

In contrast, to the present investigation, Letzring et al. (2010) introduced the codon repeats at the 4th position of the ORF, the experimental reporters were constructed using firefly luciferase and the gene sequence was not codon optimised. Therefore, in the context of this study, the predicted free energy of the thermodynamic ensemble was re-analysed for the threonine codon runs preceding the optimised *rluc* (figure 3.2). The predicted free energy values of the constructs ACA^{10}_Rluc, ACC^{10}_Rluc and ACT^{10}_Rluc and maxRluc varied from -254.21 to -249.61kcal/mol. The free energy was slightly larger for the ACG^{10}_Rluc construct at -265.60kcal/mol, correlating with the data published by Letzring et al. (2010), although only the first 50 nucleotides were analysed in their study. We compared the likelihood of the codon repeats in each construct to form base pairs and it did vary. The repeats of ACT seemed most likely to form base pairs, whereas the repeats of ACC seemed least likely to form base pairs. The predictions of secondary structure indicated that there was unlikely to be a large effect on protein production, but it still needed to be taken into consideration.

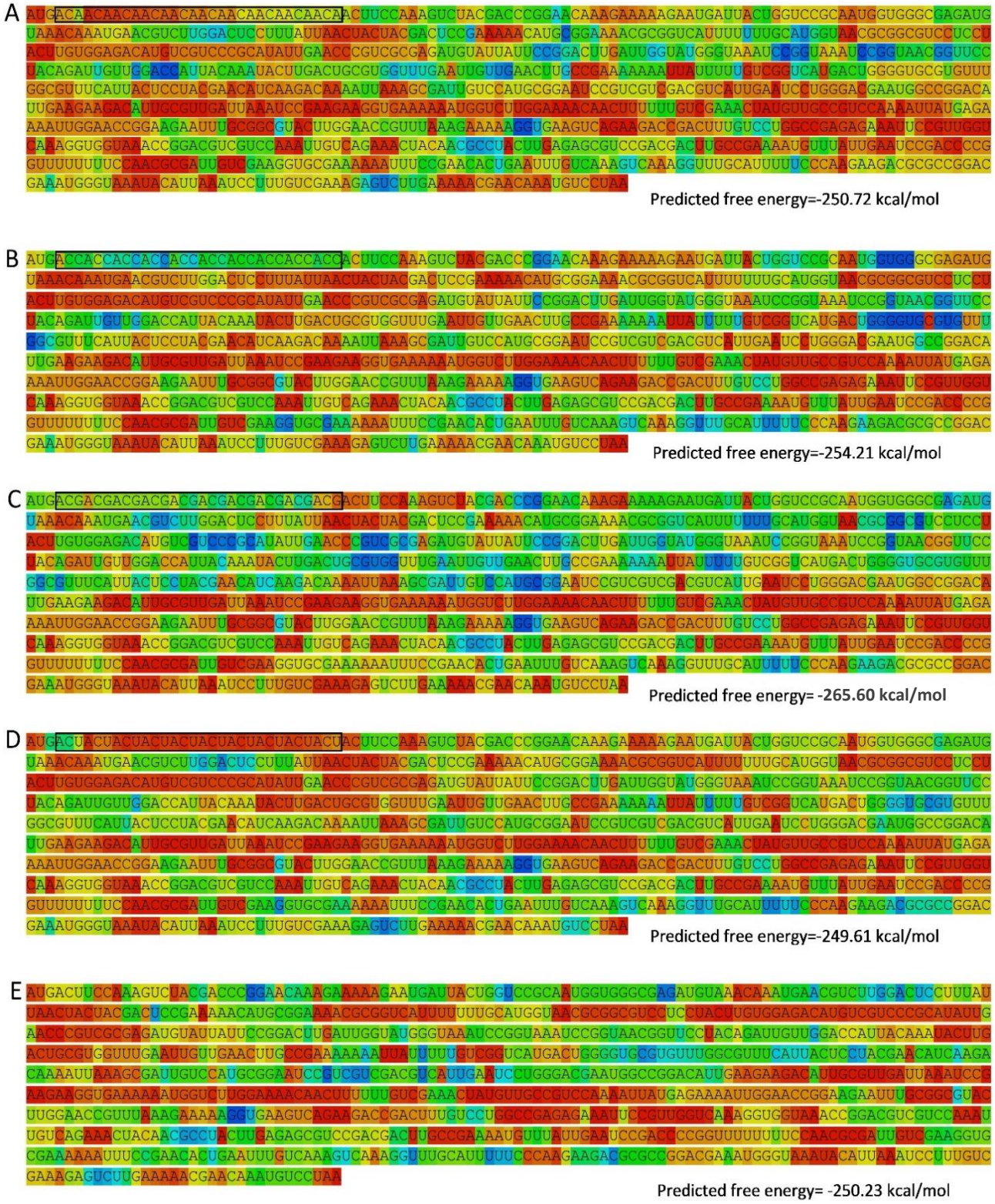


Figure 3.2- A-E represent constructs ACA^{10}_Rluc, ACC^{10}_Rluc, ACG^{10}_Rluc, ACT^{10}_Rluc and maxRluc where the base pairing probabilities and the predicted free energy of the thermodynamic ensemble is shown. The introduced ten codon repeats are highlighted in a black box. The red colour indicates a high probability of forming base pairs, whereas a blue colour indicates a low probability of forming base pairs.

3.3 Reporter and plasmid construction

The *rluc* reporters were constructed by amplification of the codon optimised *rluc* gene and ten threonine codons were introduced; via primers during PCR, specifically between the start codon and the second codon of the *rluc* ORF. The primers also introduced *XmaI* and *XhoI* restriction sites, which were used to replace the non-optimised *rluc* gene in pTH727 (an existing centromeric dual-luciferase expression plasmid with ampicillin and uracil selectable markers, Chu et al. (2014)) with the PCR products. Each plasmid is referred to by the construct that is encoded. The full map of the resulting expression plasmid for ACA^{10}_Rluc is shown in figure 3.3A. The correct integration of the PCR products, was initially confirmed by re-excising the cloned fragment of 1317 bp, with *XmaI* and *XhoI* (figure 3.3.B), which was then followed by Sanger sequencing.

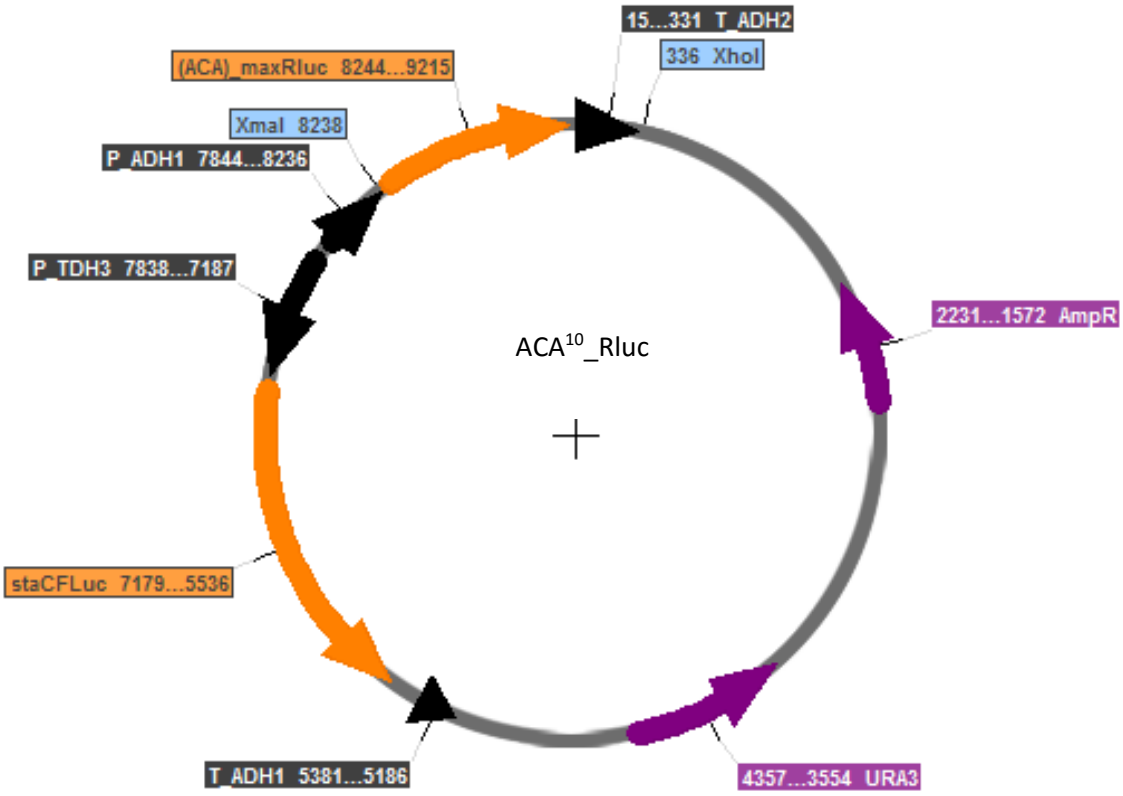
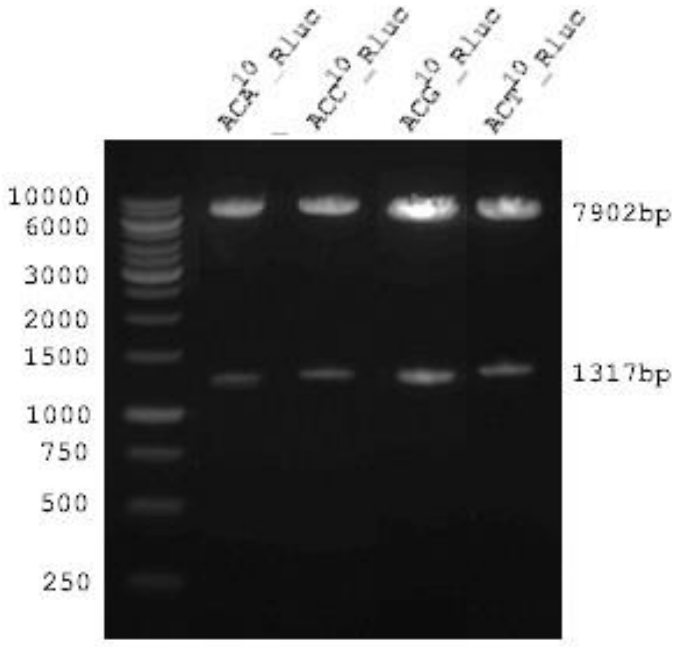


Figure 3.3.A- Schematic illustration of the ACA^{10}_Rluc plasmid. An insert of 1317bp, incorporating the *rluc* reporter construct was introduced into vector pTH727 of 7902bp, using *Xma*I and *Xho*I restriction sites. The vector encoded staCFluc and the selectable markers, uracil (*URA3*) and ampicillin (AmpR). *rluc* was under control of the *ADH1* promoter and the *ADH2* terminator, whereas, *fluc* was under the control of promoter *TDH3* and terminator *ADH1*.

Figure 3.3.B- A DNA electrophoresis gel displaying the re-excision of the insert of plasmids ACA^{10}_Rluc, ACC^{10}_Rluc, ACG^{10}_Rluc, ACT^{10}_Rluc and maxRluc, *Xma*I and *Xho*I to confirm correct integration. The top band is the vector backbone of 7902bp and the second band is the insert of 1317 bp.



3.4 Protein analyses of the reporter constructs

To initially observe differences in the rate of translation between the *rluc* reporters by assessing the protein production, protein extracts of ACA^{10}_Rluc, ACC^{10}_Rluc, ACG^{10}_Rluc, ACT^{10}_Rluc and maxRluc *S. cerevisiae* transformants were prepared. In parallel, extracts were prepared from BY4741 to serve as a negative control. Rluc was detected within these extracts by western blotting, using polyclonal antibodies labelled with HRP and visualised using ECL. The Rluc protein resolved at 37kDa for the control reporter maxRluc (figure 3.4.A) and was slightly larger for the ACA^{10}_Rluc, ACC^{10}_Rluc, ACG^{10}_Rluc, ACT^{10}_Rluc reporters, due to the introduced threonine codons. As expected, there was no band visible in the BY4741 extract as Rluc is not produced in this strain. It was clear that the ACT^{10}_Rluc reporter produced the highest level of Rluc and that the ACG^{10}_Rluc reporter produced the lowest. However, there were no visible differences between the production of Rluc from the ACA^{10}_Rluc, ACC^{10}_Rluc or maxRluc reporters. It was also important to assess whether the introduced threonine codons affected the stability of the Rluc protein, but we observed no degradation products on the western blot that would indicate such an effect.

Furthermore, we needed to ensure that the differences observed in the expression of Rluc were caused by differences in the cellular protein content, rather than technical error in loading of the gel. In order to assess total protein expression, the same protein extracts used for the western blot were also run on an SDS-PAGE gel and stained with coomassie blue (figure 3.4.B). There were no observed meaningful differences between the samples, indicating that each extract was loaded equally onto the gel. Therefore, we confirmed that there were indeed differences in protein production between the Rluc reporters. It was visibly clear that there was an increase in total protein production of the negative control, but this was of no concern as this strain does not express *rluc*.

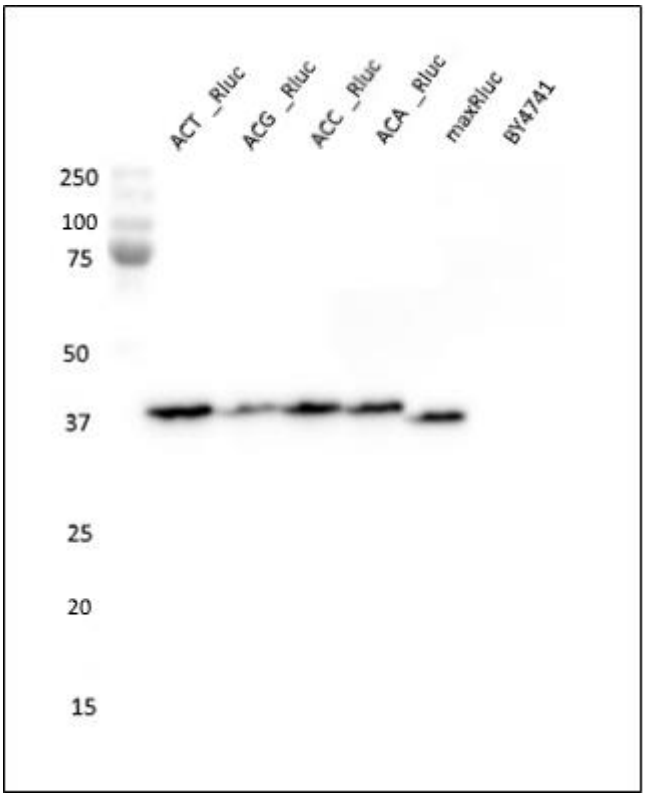


Figure 3.4.A- Western blot of ACA^{10}_Rluc, ACC^{10}_Rluc, ACG^{10}_Rluc, ACT^{10}_Rluc and maxRluc reporters with BY4741 as the control. The blot was incubated with anti-*Renilla* luciferase polyclonal antibody and specific binding of the antibody was detected by ECL. The Rluc protein resolved at 37kDa, but slightly larger for the Rluc reporters with the additional threonine codons. No band was visible for the control.

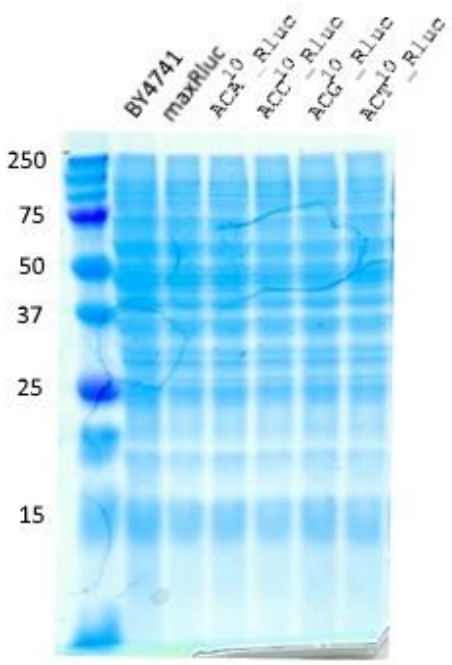


Figure 3.4.B- Coomassie blue staining of an SDS-polyacrylamide gel showing total protein expression for the ACA^{10}_Rluc, ACC^{10}_Rluc, ACG^{10}_Rluc, ACT^{10}_Rluc and maxRluc reporters and the negative control, BY4741.

3.5 Growth analyses of S. *cerevisiae* transformed with the reporter constructs

Introducing highly expressed reporters with an increased number of threonine codons into yeast cells, had the potential to deplete the threonine and/or threonyl-tRNA pools within the cell. If this happened, the transformed cells may have experienced a reduction in growth and may have shown signs of toxicity. To investigate this, the optical density of transformed cell cultures was measured over a period of time (21.5 hours) until the cells were in the stationary phase of growth. The mean doubling time (mdt) was calculated and percentage growth relative to the control (maxRluc) was calculated, revealing no significant difference (p-value 0.652) between the mdts of any of the transformants (figure 3.5 and table 3.5). Therefore, we concluded that the introduction of the threonine rich reporters did not deplete cellular threonine levels or threonyl-tRNA levels, to the point at which this became toxic to the cells.

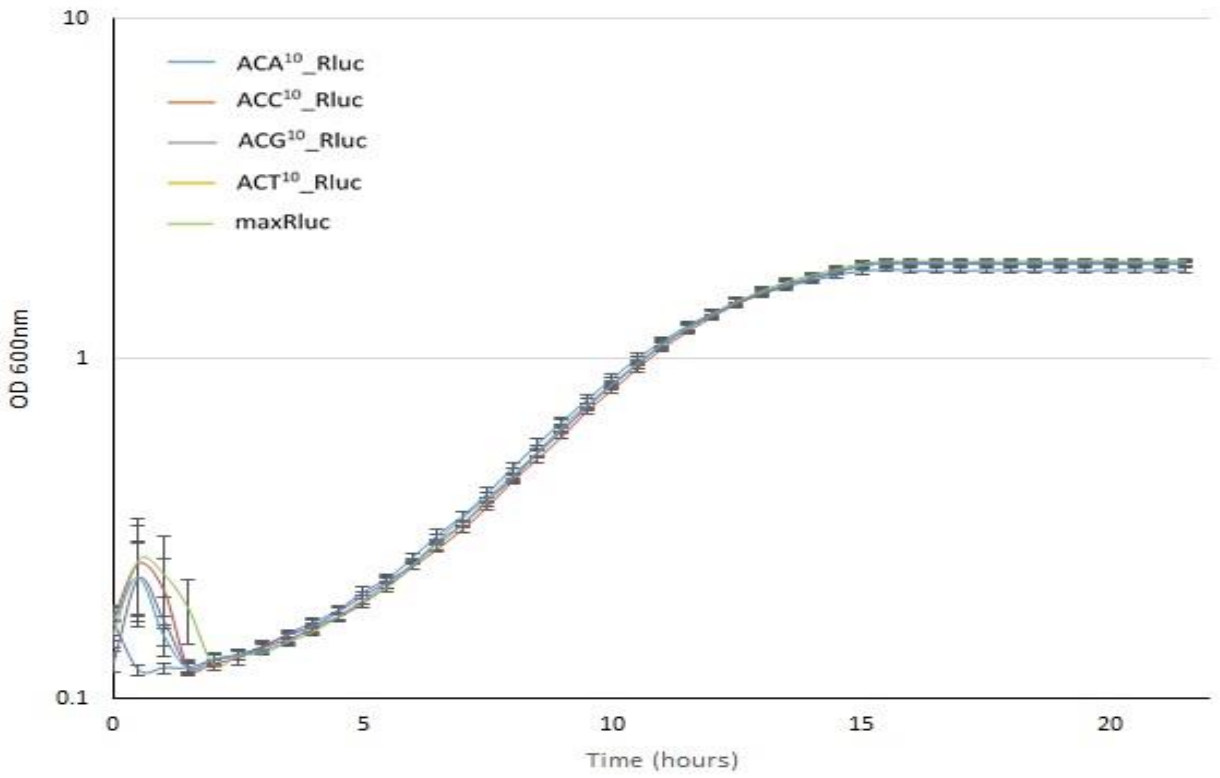


Figure 3.5- Mean logarithmic (base 10) growth curve of yeast cells transformed with the ACA^{10}_Rluc, ACC^{10}_Rluc, ACG^{10}_Rluc, ACT^{10}_Rluc and maxRluc plasmids grown for 21.5 hours, until the cells were in stationary growth phase. The optical density was measured at 600nm and standard error bars are shown.

Table 3.5- Growth analysis of ACA^{10}_Rluc, ACC^{10}_Rluc, ACG^{10}_Rluc, ACT^{10}_Rluc and maxRluc transformants.

Transformant	Mean doubling time (hours)	% growth relative to control
ACA^{10}_Rluc	2.21	100
ACA^{10}_Rluc	2.19	100
ACG^{10}_Rluc	2.17	101
ACT^{10}_Rluc	2.20	100
MaxRluc	2.20	100

.

3.6 Investigating the link between tRNA abundance and decoding speed

The cellular abundance of tRNAs is thought to be one of the determinants in decoding speed responsible for the speed differences seen in decoding synonymous codons (Novoa & Ribas de Pouplana 2012; Brockmann et al. 2007). As previously discussed, quantification of tRNAs in the cell is challenging, but it has been shown that tRNA gene copy number can be used as an estimation of the

tRNA levels (Kanaya et al. 1999). In *S. cerevisiae* the variability in the tRNA gene copy numbers of threonine is quite high; tT(AGU), tT(CGU) and tT(UGU) have 11, 1 and 5 gene copies respectively (figure 3.6.A). The three tRNAs above decode the four synonymous codons of threonine, with ACC undergoing wobble decoding. If the decoding speed did depend on the tRNA abundance and protein production was slowly limited by the decoding speed of the threonine codons, we would predict that the ACC^{10}_Rluc and ACT^{10}_Rluc reporters would have the highest level of protein production and the ACG^{10}_Rluc reporter the lowest. We have already shown initial differences in the protein production of Rluc by the reporters, indicating differences in decoding speeds of the threonine codons. To quantify this observation using a more sensitive analysis, we performed a dual-reporter luciferase assay using the activity of Fluc to standardise the data. We assume that the luciferase activity is directly correlated with the amount of reporter protein produced.

Statistical analysis of the luciferase data using Tukey's 5% comparison, showed that the luciferase activity of the ACT^{10}_Rluc reporter was significantly higher than the other Rluc reporters (figure 3.6.B). However, the same level of luciferase activity was not observed for the ACC^{10}_Rluc reporter as predicted. Although ACC and ACT are decoded by the same tRNA (tT(AGU)), the wobble decoding of ACC is thought to occur at a slower rate than Watson-Crick base pairing of ACT, which could explain the difference in activity. The luciferase activity of the ACA^{10}_Rluc and ACG^{10}_Rluc reporters are significantly different, but the activity of these two reporters were not significantly different from the activity of ACC^{10}_Rluc. The luciferase data had extremely low variability with a standard error of 0.001-0.005. Surprisingly, the maxRluc control reporter had a significant reduction in luciferase activity than the other reporters, yet the western blot did not reflect this. We cannot explain the lack of activity for maxRluc in the luciferase assay and for this reason this reporter was disregarded in further analyses. The reporters were constructed in such a way that the luciferase activity should be governed by the decoding speed of the threonine repeats. If we only consider the codons that undergo Watson-Crick base pairing, the observed protein levels of RLuc correlate with the predicted order of decoding speeds, in descending order ACT, ACA and ACG. Therefore, in principle this confirms that the tRNA levels could drive the decoding speed of threonine.

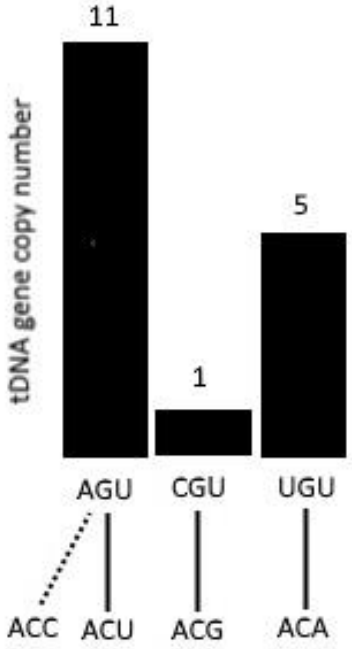


Figure 3.6.A- The four synonymous codons of threonine ACA, ACC, ACG and ACT are decoded by three tRNAs tT(AGU), tT(CGU) and tT(UGU) as shown. Watson crick base pairing is indicated by a solid black line, whereas wobble base pairing is indicated by a dotted black line. The tRNA gene copy number for each tRNA is indicated.

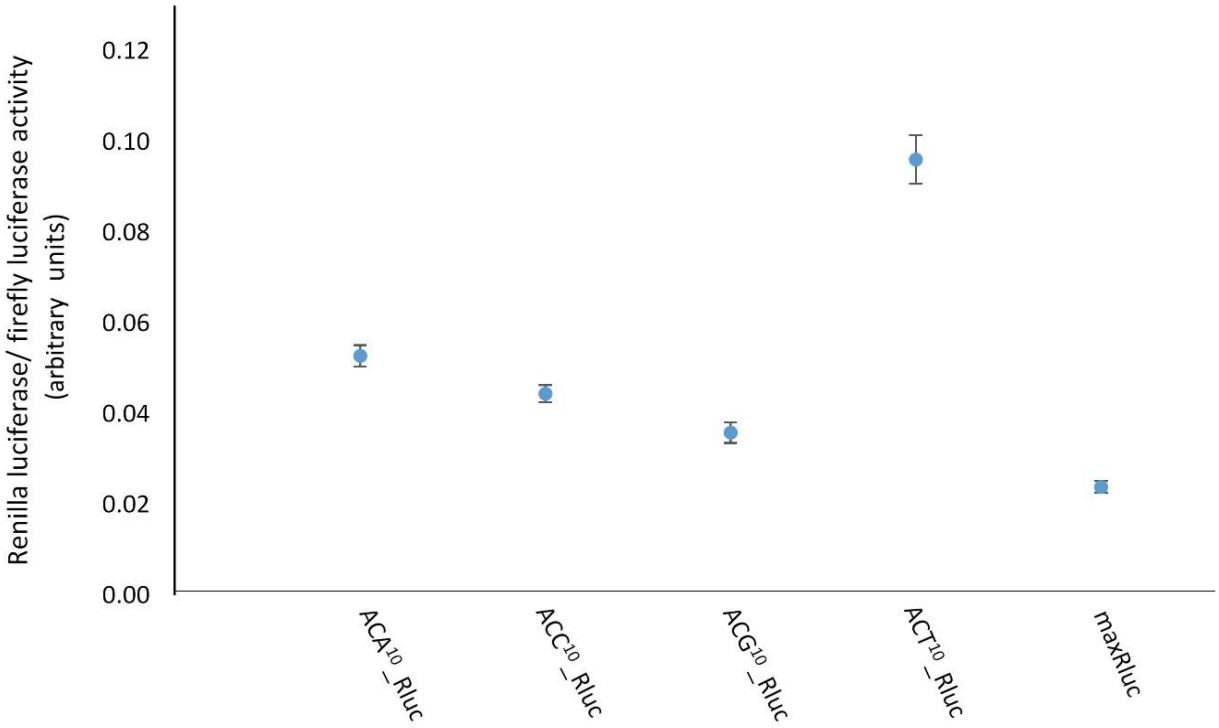


Figure 3.6.B- Mean Rluc/ Fluc activity ratios of *S. cerevisiae* cells transformed with ACA^{10}_Rluc, ACC^{10}_Rluc, ACG^{10}_Rluc, ACT^{10}_Rluc and maxRluc reporters. Standard error bars are indicated.

3.7 mRNA expression of *rluc* by the threonine codon reporters

We hypothesised that the differences in protein production were caused by differences in the rate of translation, caused by the decoding speeds of the threonine repeats in the Rluc reporters. However, codon usage has recently been shown to affect mRNA levels (Presnyak et al. 2015). Therefore, we needed to determine whether protein production was limited at the level of translation, changes to mRNA levels, or both. We carried out a two-step qRT-PCR using extracts of total cellular RNA, prepared from ACA^{10}_Rluc, ACC^{10}_Rluc, ACG^{10}_Rluc, ACT^{10}_Rluc and maxRluc transformants. To verify that the *rluc* primers amplified one PCR product, we performed a melt curve analysis and all samples had a single native-to-denatured transition at 79.5°C. In this way the fluorescence levels should have directly related to the quantity of *rluc* mRNA in the extracts. *rluc* expression levels were standardised to *fluc*, where the oligonucleotides had been previously validated (Chu et al. 2014).

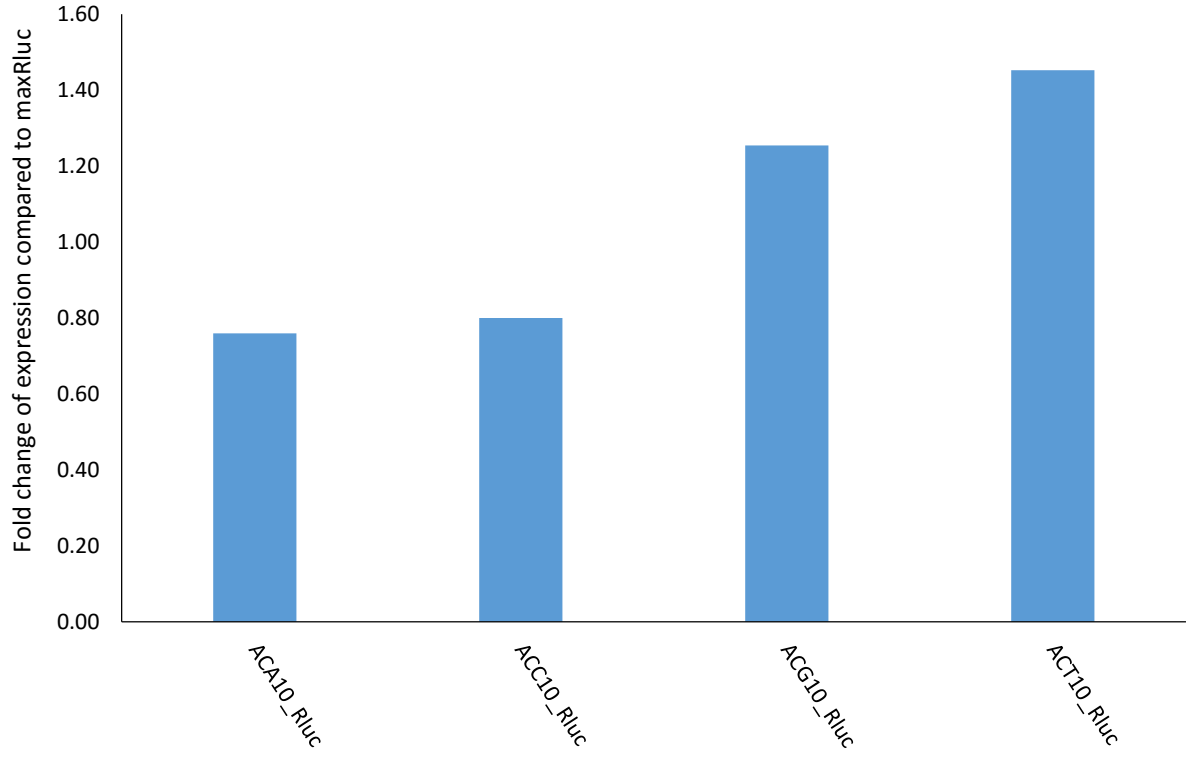


Figure 3.7- Fold change of RNA expression compared to the control, maxRluc, by the ACA^{10}_Rluc, ACC^{10}_Rluc, ACG^{10}_Rluc and ACT^{10}_Rluc reporters.

The fold change in RNA expression compared to the control ($2^{-(-\Delta\Delta ct)}$), maxRluc, was calculated to normalise the *rluc* data to be able to effectively compare the dataset. The ACT^{10}_Rluc reporter had the highest level of mRNA expression with a 1.45-fold increase whereas; the ACC^{10}_Rluc reporter had a 0.2-fold reduction (figure 3.7). The mRNA expression levels of *rluc* by the ACA^{10}_Rluc reporter had a

decrease of 0.24-fold, similar to the ACC^{10}_Rluc reporter. The dataset thus far, correlates to the observed protein levels for each reporter except ACG^{10}_Rluc. ACT^{10}_Rluc has the highest level of expression whereas; ACA^{10}_Rluc and ACC^{10}_RLuc both have a lower but similar level of expression. The mRNA expression of *rluc* by the ACG^{10}_Rluc reporter had a 1.25-fold increase, whereas this reporter had the lowest protein level.

Although, the expression dataset has quite a wide variation, an ANOVA of the ct values, showed no significant differences between the *rluc* mRNA levels with a p-value of 0.219. With a p-value of this size there was still a chance, albeit small, that the data were meaningful. Furthermore, qRT-PCR generally has higher variability in its results compared to other techniques, which made the interpretation of this dataset more challenging. We concluded, that only the ACG^{10}_Rluc reporter is likely to be controlled at the level of translation and so able to report on decoding speed, having observed high mRNA levels but low protein levels. However, we could not draw strong conclusions on the point of control for the other threonine codon reporters based on our data.

3.8 Assessment of changes in tRNA abundance on decoding speed

Although we concluded that only the ACG^{10}_Rluc reporter is likely to be controlled at the level of translation, it was still interesting to see whether we could make predictions on Rluc synthesis, based on changes to the abundance of tRNAs for the other threonine reporters. To do this we performed a dual-reporter luciferase assay on a closely related species of *S. cerevisiae, Saccharomyces uvarum*. There are subtle differences in the tRNA gene copy numbers in *S. cerevisiae* compared to *S. uvarum*, with an additional gene copy for tT(AGU) and one fewer for tT(UGU), figure 3.8.A. Based on the theoretical understanding of the decoding system, this should have meant that ACU and ACC (decoded by tT(AGU)) would have been decoded faster, whereas ACA (decoded by tT(UGU)) would have been decoded more slowly. These predictions should have been reflected in the luciferase activity of the *rluc* reporters and were based on the proportional change in tRNA gene copies of 9.1% for tT(AGU) and 20% for tT(UGU), however both effects might have been expected to be small.

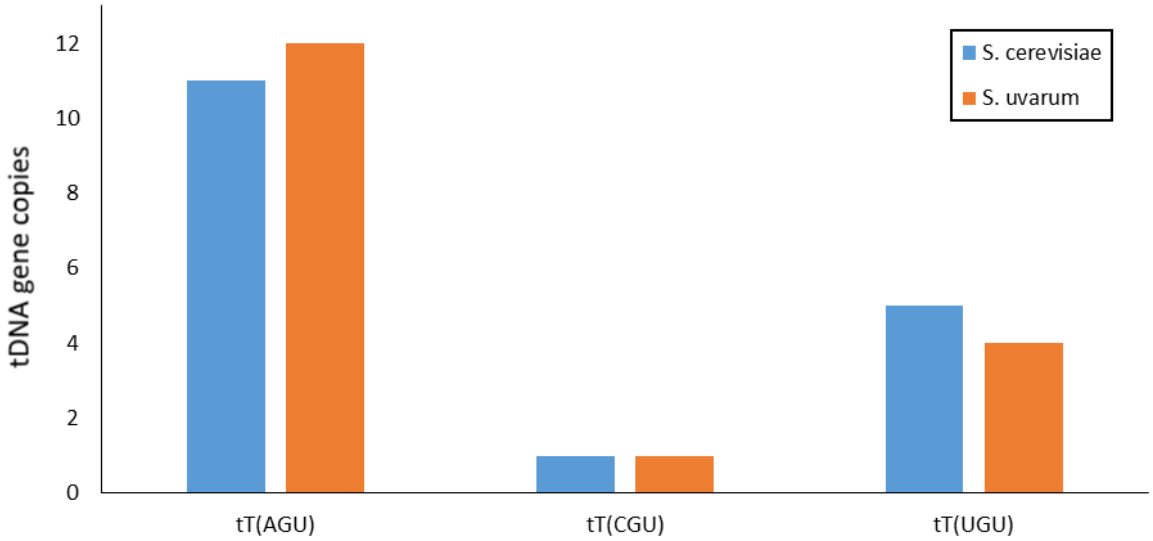


Figure 3.8.A- Gene copy numbers for the three threonyl-tRNAs, tT(AGU), tT(CGU) and tT(UGU) in *S. cerevisiae* and *S. uvarum*.

In comparison to *S. cerevisiae*, there was 28% less variation between the protein levels produced by the threonine reporters in *S. uvarum* (figure 3.8.B) and the luciferase activity of each reporter was significantly increased based on the observed Rluc/Fluc ratio (p-value 0.002-0.000). There was a significant difference between the activity of the ACC^{10}_Rluc and ACT^{10}_Rluc reporters in *S. uvarum* (p-value 0.001), as previously observed in *S. cerevisiae* (p-value 0.000). There was no change in the gene copy levels of tT(CGU) between the two species and so we did not expect to see a change in the activity of the ACG^{10}_Rluc reporter. Therefore, we normalised the Rluc/Fluc ratios of *S. cerevisiae* and *S. uvarum* to this reporter, figure 3.8.C. As predicted there was a significant decrease in the activity of the ACA^{10}_Rluc reporter (p-value 0.024). We cannot explain the effects of the increase in gene copies of tT(AGU), with no significant change in the activity of the ACC^{10}_Rluc reporter, but a significant decrease in the activity of the ACT^{10}_Rluc reporter. In conclusion, we were unable to make predictions on luciferase activity of our reporters based on the abundance of tRNAs. The unpredictability is likely due to the differences in the mRNA levels of *S. uvarum* compared to *S. cerevisiae* and/or supporting our mRNA analyses that the ACA^{10}_Rluc, ACC^{10}_Rluc and ACT^{10}_Rluc reporters may not be controlled at the level of translation.

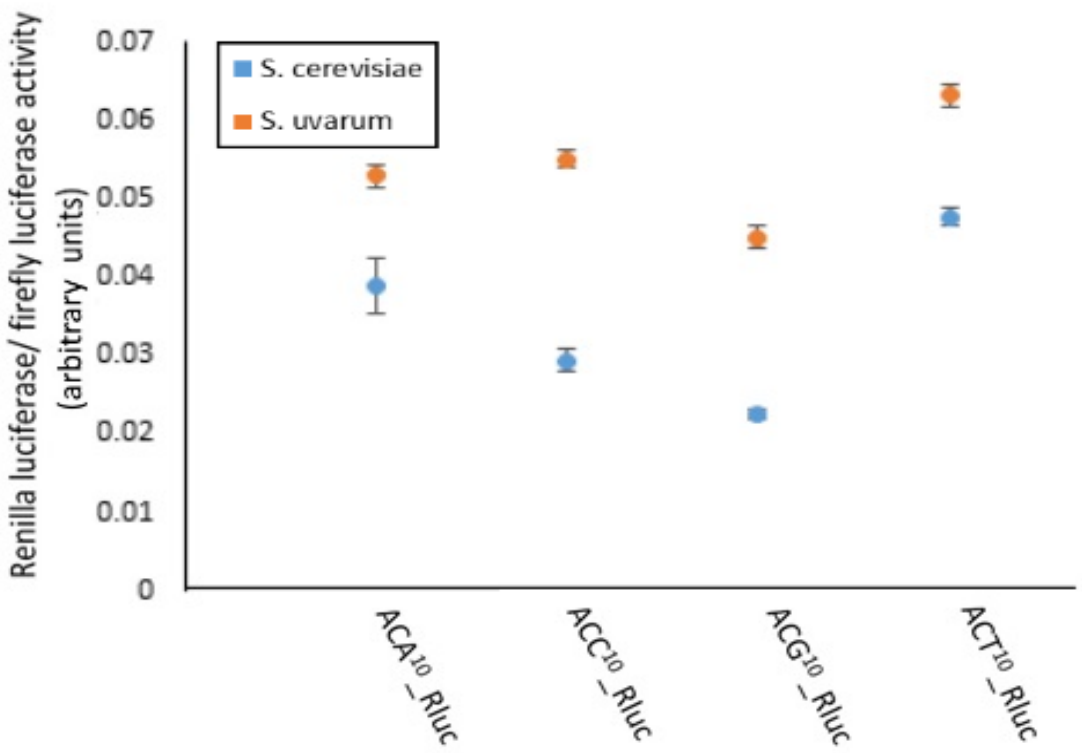


Figure 3.8.B- (above) Mean *Renilla* luciferase/firefly luciferase activity of *S. cerevisiae* and *S. uvarum* transformed with the ACA[10]_Rluc, ACC[10]_Rluc, ACG[10]_Rluc, ACT[10]_Rluc and maxRluc reporters. Standard error bars are indicated.

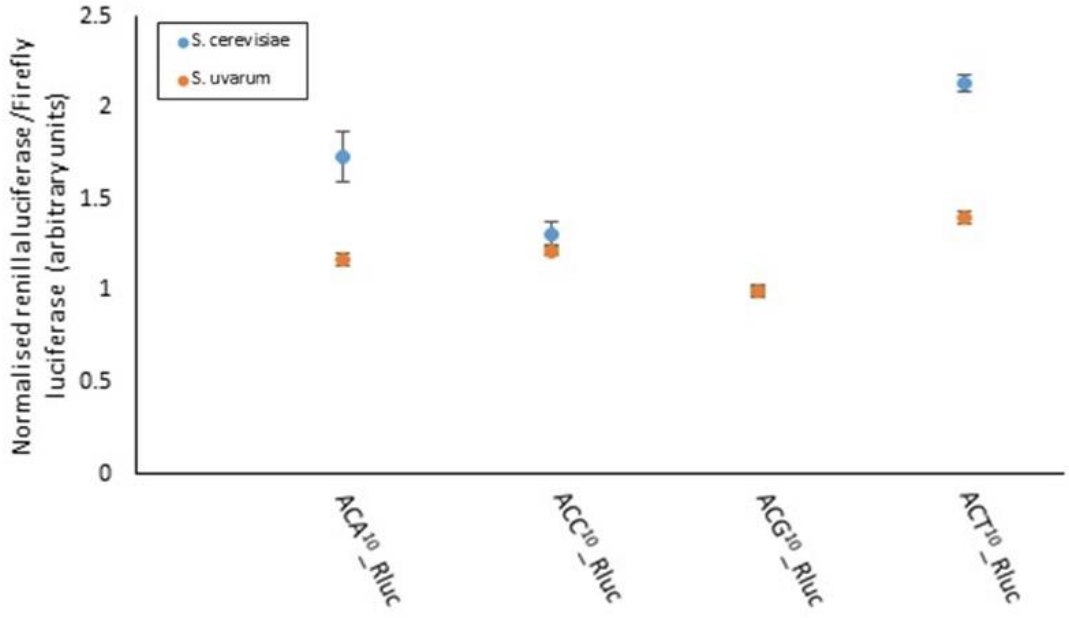


Figure 3.8.C- Mean *Renilla* luciferase/firefly luciferase activity normalised to the ACG[10]_Rluc reporter, of *S. cerevisiae* and *S. uvarum* transformed with the ACA[10]_Rluc, ACC[10]_Rluc, ACG[10]_Rluc, ACT[10]_Rluc and maxRluc reporters. Standard error bars are indicated.

3.9 Perturbation of the tRNA pool and the effect on protein production of the ACG^{10}_Rluc reporter construct

So far, mRNA analysis of the *rluc* reporters indicated that only the ACG^{10}_Rluc may be controlled at the level of translation. To investigate whether tRNA abundance is a determinant of decoding speed of the ACG codon, we examined Rluc production in *S. cerevisiae* strains in which tRNA levels had been manipulated, either by deleting individual tRNA genes, or by overexpressing tRNAs by single and multi-copy plasmids.

3.9.1 Construction of tRNA expression plasmids and tRNA knockdown strains

A gene copy of each tRNA was cloned into a single copy pRS313 and a multi-copy pRS423 plasmid. Cells transformed with the single copy plasmid would gain an additional copy of the tRNA gene per cell, whereas those transformed with the multi-copy plasmid would gain numerous copies per cell. An effect observed by the introduction of additional tRNAs from a single copy plasmid should be enhanced further when introduced by a multi-copy plasmid. One gene copy of both tT(AGU) and tT(UGU) was cloned from genomic DNA of BY4741 into pRS313 and pRS423 using *Xma*I and *Xho*I. The other tRNA, tT(CGU)K, was cloned from an existing plasmid (pTH485) using BamHI. To assess the knockdown in tRNA expression we had access to the following *S. cerevisiae* strains, $\Delta^{tT(AGU)B}$, $\Delta^{tT(AGU)H}$, $\Delta^{tT(UGU)G1}$ and $\Delta^{tT(UGU)P}$. We did not have a deletion for tT(CGU) because the single gene for this tRNA is essential.

3.9.2 Growth analyses of the tRNA knockdown strains and cells transformed with the tRNA expression plasmids

Transformants of the single-copy plasmids did not have significantly different (p-value 0.104) mdts relative to the control (transformants of the empty pRS313 vector), table 3.9.2. Overexpressing a tRNA gene could potentially be toxic to the cell due to subsequent depletion of the threonine pool, but this does not seem to be the case. Transformants of the tRNA knockdown strains did exhibit a significant increase in growth compared to BY4741 (p-value 0.017), but the increase was relatively small and as we were performing dual-reporter luciferase assays it was unlikely to affect our results. We concluded that neither the knockdown or overexpression of individual tRNAs, caused a decrease in growth indicating no interference with cell fitness.

Table 3.9.2- Growth summary of tRNA knockdown strains and BY4741 as the control, as well as cells transformed with the single-copy vectors overexpressing tT(AGU), tT(CGU) and tT(UGU) and BY4741 transformed with an empty single-copy plasmid as the control.

Transformant	Mean doubling time (hours)	Specific growth rate (h^{-1})	% growth (relative to control)
BY4741 + pRS313	2.27	0.44	100
tT(AGU)	2.23	0.45	97
tT(CGU)	2.30	0.43	102
tT(UGU)	2.20	0.45	97
ΔtT(UGU)G1	2.31	0.43	102
ΔtT(AGU)B	1.89	0.51	103
ΔtT(AGU)H	1.94	0.51	103
ΔtT(UGU)P	1.95	0.51	100
BY4741	1.89	0.53	100

3.9.3 Assessment of protein production by the ACG^{10}_Rluc reporter using the dual-reporter luciferase assay

The decoding tRNA for the ACG codon is tT(CGU)K, for which there is a single gene, and introduction of a single copy plasmid of the tRNA gene should double its cellular content and positively affect the decoding speed of ACG codons. The other two tRNAs should be competing with tT(CGU)K at the ribosome to access the ACG codons and introduction of plasmid borne copies, could therefore be expected to decrease the decoding speed of ACG codons. tT(AGU) and tT(UGU) have higher gene copy numbers and any effect might only be detectable using multi-copy plasmids.

The expression of ACG^{10}_Rluc in the presence of tRNA overexpression, was analysed by a dual-reporter luciferase activity where the Rluc/Fluc activity was standardised relative to the control (BY4741 transformed with an empty pRS313 vector and the ACG^{10}_Rluc reporter), figure 3.9.3. Surprisingly, significant differences in Rluc production was only seen in cells transformed with the multi-copy plasmids (p-values <0.05). There was a significant increase in luciferase activity of the ACG^{10}_Rluc reporter when tT(CGU)K (p-value 0.011) and the non-cognate tRNA tT(AGU) (p-value 0.05) were overexpressed. Whereas, no meaningful difference was observed upon overexpressing tT(UGU). The

differential production of Rluc of ACG^{10}_Rluc seen between the two non-cognate tRNAs may be due to the presence of different copy numbers of the tRNAs. Haploid yeast has approximately 20 copies of multi-copy plasmid per cell, but these copies are lost through mitotic segregation during cell division (Christianson et al. 1992). Therefore, each transformant was likely to have varying copy numbers which could explain the observed differences in luciferase activity.

A knockdown in the abundance of the non-cognate tRNAs (tT(AGU) and tT(UGU)) was thought to lead to less competition for access to the ACG codons by tT(CGU)K. Therefore, we predicted that the luciferase activity of the ACG^{10}_Rluc reporter would increase in the knockdown strains, which indeed was the case producing a significant result (p-value <0.015). Our data indicates that there is unpredictability in the protein production of Rluc from the ACG^{10}_Rluc reporter, based on the abundance of cognate and non-cognate tRNAs.

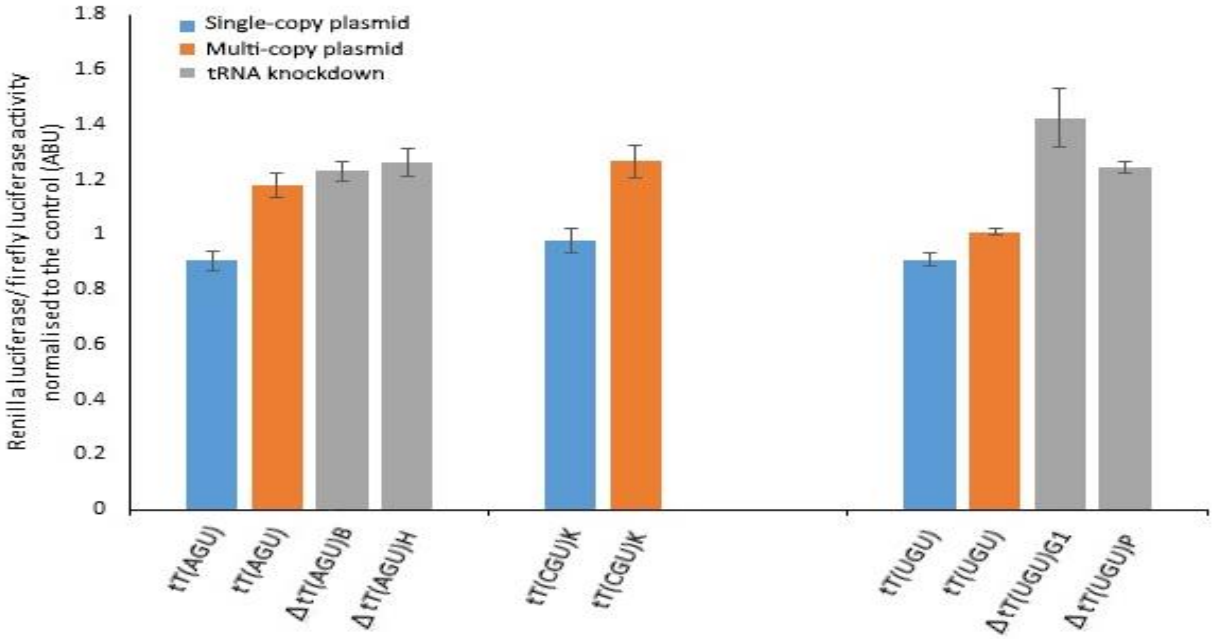


Figure 3.9.3- Mean *Renilla* luciferase/ firefly luciferase activity of *S. cerevisiae* cells all transformed with the ACG^{10}_Rluc reporter. The graph shows cells also overexpressing the three tRNAs (tT(AGU), tT(CGU)K and tT(UGU) by both single-copy and multi-copy plasmids. The mean *Renilla* luciferase/ firefly luciferase activity is also shown for strains with individual tRNAs knocked down, $\Delta^{tT(AGU)B}$, $\Delta^{tT(AGU)H}$, $\Delta^{tT(UGU)G1}$ and $\Delta^{tT(UGU)P}$. The luciferase activity was standardised to the appropriate control and standard error bars are indicated.

3.9.4 Summary of results

- The threonine reporters exhibited different expression levels, indicating the influence of codon bias on gene expression
- mRNA analyses showed that the point of control is likely to lie at the translational level for only one of the reporters, ACG^{10}_Rluc. However, the expression levels of the reporters were consistent with our knowledge of threonyl-tRNA populations
- Investigating the effect of tRNA abundance on Rluc production in two yeast species as well as experimental analyses, did not produce results as expected

Chapter 4: Discussion

4 Discussion

The purpose of this study was to generate and characterise reporter constructs for measuring codon decoding speeds, with the aim of investigating current theories on the role of tRNA abundance and competition in determining decoding speeds. Ten codon repeats of the four synonymous threonine codons were introduced at the beginning of the ORF of an optimised version of *rluc*. This ORF should have provided efficient translation of the sequences, in such a way that the codon repeats were the only determinant limiting the rate of elongation. Letzring et al (2010), provided secondary structure analysis suggesting the synonymous threonine codons to be the best candidates for the reporters, as the majority did not have a strong tendency to form secondary structures in a similar context.

Synthesis of RLuc from each experimental reporter appeared to be stable as we observed no evidence of protein degradation on the western blots, similar to a control reporter which consisted only of the codon optimised *rluc,* with no additional threonine codons. Interestingly, protein synthesis by the control was reduced compared to the threonine codon reporters. For this reason, we disregarded the control reporter in further analyses. According to the N-end rule, proteins with a threonine at the N-terminal have one of the highest *in vivo* half-lives in *S. cerevisiae* (Varshavsky 1996). The repeat of this amino acid may have increased the protein half-life, leading to an increase in protein stability explaining the difference in luciferase activity between the threonine codon and control reporters.

RLuc production was initially assessed via western blots, providing a qualitative analysis of the differences in protein synthesis between the reporters. Using a more sensitive dual-reporter luciferase assay, the luciferase activity correlates with the level of protein synthesis and so we use this assay as an indicator of the rate of translation of each reporter. The resulting data initially appeared consistent with the predicted decoding speeds of the four threonine codons, in descending order ACT, ACA, ACC and ACG, based on our knowledge of threoninyl-tRNA levels. tT(AGU) decodes both ACC and ACT, but the difference in protein levels suggested a clear reduction in the decoding speed of ACC compared to ACT. It is thought that selection and accommodation of a wobble-decoding tRNA into the A-site of the ribosome, induces flipping of the conserved adenosine resides (A1492 and A1293), stabilisation of codon:anticodon interactions and interactions with the PTC occurs at a slower rate compared to Watson-Crick base pairing (Zeng et al. 2014; Taylor et al. 2007; Agris et al. 2007; Plant et al. 2007), which is likely to explain the observed differences. While our observed order of protein levels was consistent with the decoding speeds proposed by Chu et al. (2014), other publications suggested different decoding speeds. For example, Shah & Gilchrist (2011) identified ACC and ACT *in silico*, to

have the shortest elongation times in descending order, which would be less consistent with our observations.

The discussion in the preceding paragraphs assumes that the Rluc production is limited by the decoding speeds of the threonine codons. To ensure that protein synthesis from the four threonine codon reporters was indeed limited at the translational level, we also assessed mRNA levels for these reporters. The results suggested that much of the differences in protein levels was not due to translational control but caused by changes to the mRNA levels. ACG^{10}_RLuc was the only reporter where the mRNA levels could not explain the luciferase activity and so this reporter looks likely to be translationally controlled, and protein levels are therefore likely to depend on the decoding speed of the ACG codons. We found no significant difference between the ct values from the qRT-PCR dataset (p-value 0.219), however the fold change in expression of each reporter compared to the control exhibited meaningful differences, making the interpretation challenging. Marín et al. (2003) suggested that transcripts that have a high GC content have a higher mRNA concentration, but our data does not seem to support this, ACT^{10}_Rluc has the highest concentration of mRNA and ACC^{10}_Rluc has a significantly lower concentration. This trend may be as a consequence of selection, which would explain why it would not apply to our synthetic reporters.

Recently it has been suggested that codon usage affects mRNA stability, in such a way that transcripts with a higher proportion of optimal codons have been observed to have an increase in half-life and a reduced turnover rate (Presnyak et al. 2015). The mechanistic link between codon usage and mRNA stability is unclear. Furthermore, codon optimality has been shown to correlate with the abundance of tRNAs (Novoa et al. 2012; Brockmann et al. 2007). Therefore, in the context of this study, for codons undergoing Watson-Crick base pairing (all except ACC), the more tRNA gene copies and so abundance of tRNAs, the more optimal the codon and the higher the expected concentration of mRNA according to the findings from Presnyak et al. (2015). ACG is decoded by tT(CGU) for which there is only one gene copy, so you would expect the mRNA levels for the ACG^{10}_Rluc reporter to be the lowest, but this was not the case.

Presnyak et al. (2015) determined the order of mRNA stability by the occurrence of specific codons to be, in descending order ACT, ACC, ACA and ACG. They found ACT and ACC to be stable codons, but ACA and ACG to be unstable. Our data did not support these results, rather that ACA and ACC may be unstable but ACG and ACT stable. Using codon content as a determinant of mRNA stability does not apply globally, such as histone components for example (Presnyak et al. 2015) and it is possible that it

also does not apply for *rluc.* Indeed, the correlation between mRNA levels and codon optimality is unclear, as a correlation can be present if two variables interact, if one is a cause of the other or if they are both effects of a third variable (Coghlan & Wolfe 2000).

Of the four threonine codon reporters, ACG^{10}_Rluc was the only one where the reduced protein levels could clearly not be explained by reduced mRNA levels (figures 3.6.B & 3.7). Therefore, protein synthesis from this reporter is very likely limited at the translational level, and this might be caused by the predicted low decoding speed of ACG codons. However, this reporter was also the only one showing a slightly stronger tendency to form secondary structure compared to the other three reporters (figure 3.2). This difference was suggested both by our own analyses, and reported by Letzring et al (2010). There are studies that identify when secondary structure may affect the rate of translation, for example, an increase in secondary structure of the 'ramp site' following the start codon is thought to decrease the efficiency of initiation (Kudla et al. 2009; Tuller et al. 2010). It may be that the slight increase in predicted secondary structure caused by introduction of ACG repeats to the reporter may have reduced the rate of initiation and elongation, decreasing protein production. According to Sagliocco et al. (1993), an increase in secondary structure in the 5'-UTR by 10kcal/mol can cause a reduction in the rate of translation initiation by as much as 50%. Therefore, the influence of secondary structure on the rate of translation rather than the decoding speed of the ACG codons cannot be excluded. Although, our structural analysis predicted that the ACG repeats were not the most likely to form base pairs out of the synonymous codon runs, indicating that the introduction of the ACG codons may have had more of a global effect on mRNA secondary structure.

A negative correlation between tRNA abundance and mRNA secondary structure has been suggested as a mechanism to maintain a uniform level of translation (Zur & Tuller 2012). In this way, a transcript with mostly optimal codons would have a higher level of secondary structure, whereas a transcript with mostly non-optimal codons, would have a lower level of secondary structure. If the correlation between the mRNA structure and tRNA abundance is a consequence of selection alone, this would not apply to our synthetic constructs. There is only one gene copy of tT(CGU) decoding ACG, so we might expect the mRNA to have a lower level of secondary structure, but our data did not fit this trend.

We could not be sure as to the point of control for each of the reporter constructs, for those that seemed to be governed by changes to mRNA levels (ACA^{10}_Rluc, ACC^{10}_Rluc and ACT^{10}_RLuc), it was still interesting to see whether we could make predictions on protein levels of RLuc based on changes to the abundance of tRNAs. We looked at the closely related species, *S. uvarum* that differed only by

an additional tT(AGU) gene copy and a reduction of one tT(UGU) gene copy. Based on the theoretical understanding of the decoding system, we expected ACU and ACC (decoded by tT(AGU)) to have been decoded faster, whereas ACA (decoded by tT(UGU)) would have been decoded more slowly. There was unpredictability in the results, indicating that there may be no control at the level of translation for the three aforementioned threonine codon reporters, or there are multiple determinants involved in controlling decoding speed. It would be interesting to look at expression of the reporters in *S. eubayanus* which has the same tRNA gene copy numbers as *S. cerevisiae,* to observe the phenotypic differences in a similar translational system (Scannell et al. 2011). This would involve generating genetically tractable strains of this organism.

The most promising of the threonine reporters, for which there was clear evidence that protein levels may be limited translationally, was the ACG^{10}_Rluc reporter. ACG is decoded by tT(CGU), which had identical gene copy numbers in *S. cerevisiae* and *S. uvarum*. Within the *Saccharomyces* genus many of the species have one tT(CGU) gene copy which is essential and cannot be deleted. In order to assess changes to tRNA levels on the protein production by the ACG^{10}_Rluc reporters, we experimentally altered the abundance of tRNAs by the use of overexpression vectors and knockdown strains and assessed the effects by luciferase assays. Effects of tRNA overexpression was only seen in cells transformed with multi-copy plasmids, where effects on the cell may be enhanced to an observable level. The introduction of a multi-copy plasmid into a cell, generally exerts a level of stress (Yona et al. 2013), these plasmids are inherited stochastically and copies are lost through mitotic segregation at 4.3±1.3% of progeny per doubling time (Christianson et al. 1992). As gene copy numbers for tT(CGU) appear strictly limited in yeasts, we expected the growth rates of cells transformed with the multi-copy plasmids encoding additional tT(CGU) genes to have a negative effect on the cell, but this was not the case. Overexpression of non-cognate tRNAs (tT(AGU) and tT(UGU)) effected protein synthesis of the ACG^{10}_Rluc reporter, but the results did not follow the patterns expected from our understanding of the decoding system. As multi-copy plasmids are inherited in a non-mendelian fashion, there was likely to be different copy numbers per transformant, which would result in different luciferase activities. We observed an increase in protein levels of ACG^{10}_Rluc when one out of two non-cognate threonyl-tRNAs were overexpressed, but a knockdown in the abundance of non-cognate tRNAs, that do not decode ACG codons, led to an expected increase in protein synthesis from ACG^{10}_Rluc. It may be that we do not fully understand how tRNAs compete and also interact with the ribosome. To further understand our tRNA overexpression data, RNA analysis by qRTPCR or northern blot would enable an estimate of the plasmid copy numbers between the transformants. Our data suggest that either the “speedometer” constructs do not report on codon decoding speeds, or that

there are more determinants other than the abundance of tRNAs that govern decoding speed. Chu et al. (2011) supports the theory that there may be multiple determinants of decoding speed rather than or as well as tRNA abundance, as they observed that only the abundance of a few tRNAs, including (tD(GUC), tR(ACG), tE(UUC) and tL(UAG)) affect the rate of translation.

The process of translation and machinery are complex and challenging to understand. Our reporter constructs did not work as expected and it seems that ACG^{10}_Rluc, may be the only reporter that is able to report on decoding speeds. We were unable to determine whether tRNA is one of the molecular determinants of decoding speed, but this may be due to the inefficiency of our reporters. In such an intricate system, it seems intuitive that there is likely to be a synergy of determinants acting to control decoding speed. Several studies have found that the correlation between tRNA abundance and protein expression is stronger when other factors are taken into account (Tuller et al. 2010). Therefore, when investigating the rate of elongation, all of the possible contributing factors and machinery must be incorporated and accounted for. Although this would be an extensive challenge, it seems necessary to be able to observe and understand all of the relationships that work to control the rate of elongation. We, like many others, use tRNA gene copy number as an estimate of tRNA abundance, but directly measuring the tRNA levels by novel next generation sequencing (e.g Illumina tRNA-seq) would further enable us to understand the regulation and role of tRNAs (Pang et al. 2014; Zheng et al. 2015). Future work would require an alternative approach to identify the determinants governing decoding speed and we suggest, that working with an *in vitro* system may provide a simpler platform to work with.

5 References

Agris, P.F., Vendeix, F.A.P. & Graham, W.D., 2007. tRNA's Wobble Decoding of the Genome: 40 Years of Modification. *Journal of Molecular Biology*, 366(1), pp.1–13.

Blanchard, S.C. et al., 2004. tRNA dynamics on the ribosome during translation. *Proceedings of the National Academy of Sciences of the United States of America*, 101(35), pp.12893–12898.

Brachmann, C.B. et al., 1998. Designer deletion strains derived from Saccharomyces cerevisiae S288C: A useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast*, 14(2), pp.115–132.

Branchini, B.R. et al., 2005. Red- and green-emitting firefly luciferase mutants for bioluminescent reporter applications. *Analytical Biochemistry*, 345(1), pp.140–148.

Brockmann, R. et al., 2007. Posttranscriptional expression regulation: What determines translation rates? *PLoS Computational Biology*, 3(3), pp.0531–0539.

Cannarozzi, G. et al., 2010. A Role for Codon Order in Translation Dynamics. *Cell*, 141(2), pp.355–367.

Christianson, T.W. et al., 1992. Multifunctional yeast high-copy-number shuttle vectors. *Gene*, 110(1), pp.119–122.

Chu, D. et al., 2014. Translation elongation can control translation initiation on eukaryotic mRNAs. *EMBO Journal*, 33(1), pp.21–34.

Chu, D., Barnes, D.J. & von der Haar, T., 2011. The role of tRNA and ribosome competition in coupling the expression of different mRNAs in Saccharomyces cerevisiae. *Nucleic Acids Research*, 39(15), pp.6705–6714.

Coghlan, A. & Wolfe, K.H., 2000. Relationship of codon bias to mRNA and concentration protein length in Saccharomyces cerevisiae. *Yeast*, 16(12), pp.1131–1145.

Cuesta, R. & Gupta, M., 2009. The regulation of protein synthesis in cancer. *Progress in molecular biology and translational science,*90, pp.255-92.

Dever, T.E. & Klann, E., 2004. Biochemical mechanisms for translational regulation in synaptic plasticity. *Nature Reviews* Neuroscience, 5, pp.931-94.

Dever, T.E. & Green, R., 2015. Phases of Translation in Eukaryotes. *Cold Spring Harbor Laboratory Press*, 4(7).

Firczuk, H. et al., 2013. An in vivo control map for the eukaryotic mRNA translation machinery. *Molecular systems biology*, 9(635), p.635.

Fluitt, A., Pienaar, E. & Viljoen, H., 2007. Ribosome kinetics and aa-tRNA competition determine rate and fidelity of peptide synthesis. *Computational Biology and Chemistry*, 31(5-6), pp.335–346.

Frank, J. et al., 2007. The process of mRNA – tRNA translocation. *Proceedings of the National*

Academy of Sciences of the United States of America, 104(50), pp.19671–19678.

Frank, J., 2003. Toward an understanding of the structural basis of translation. *Genome biology*, 4(12), p.237.

Gardin, J. et al., 2014. Measurement of average decoding rates of the 61 sense codons in vivo. *eLife*, 3, pp.1–20.

Gerton, J.L. et al., 2000. Global mapping of meiotic recombination hotspots and coldspots in the yeast Saccharomyces cerevisiae. *Proceedings of the National Academy of Sciences of the United States of America*, 97(21), pp.11383–90.

Gorochowski, T.E. et al., 2015. Trade-offs between tRNA abundance and mRNA secondary structure support smoothing of translation elongation rate. *Nucleic Acids Research*, 43(6), pp.3022–3032.

Graifer, D. & Karpova, G., 2015. Interaction of tRNA with Eukaryotic Ribosome. *International Journal of Molecular Sciences*, 16(4), pp.7173–7194

von der Haar, T., 2008. A quantitative estimation of the global translational activity in logarithmically growing yeast cells. *BMC systems biology*, 2(1), p.87

Harger, J.W. & Dinman, J.D., 2003. An in vivo dual-luciferase assay system for studying translational recoding in the yeast Saccharomyces cerevisiae. *RNA (New York, N.Y.)*, 9(8), pp.1019–24.

Hausmann, C.D. & Ibba, M., 2008. Aminoacyl-tRNA synthetase complexes: molecular multitasking revealed. *FEMS microbiology reviews*, 32(4), pp.705–21.

Hori, H., 2014. Methylated nucleosides in tRNA and tRNA methyltransferases. *Frontiers in Genetics*, 5, pp.1–26.

Kahali, B., Basak, S. & Ghosh, T.C., 2007. Reinvestigating the codon and amino acid usage of S. cerevisiae genome: A new insight from protein secondary structure analysis. *Biochemical and Biophysical Research Communications*, 354, pp.693–699.

Kanaya, S. et al., 2001. Codon usage and tRNA genes in eukaryotes: Correlation of codon usage diversity with translation efficiency and with CG-dinucleotide usage as assessed by multivariate analysis. *Journal of Molecular Evolution*, 53(4-5), pp.290–298.

Kanaya, S. et al., 1999. Studies of codon usage and tRNA genes of 18 unicellular organisms and quantification of Bacillus subtilis tRNAs: Gene expression level and species-specific diversity of codon usage based on multivariate analysis. *Gene*, 238(1), pp.143–155.

Kudla, G. et al., 2009. Coding-sequence determinants of gene expression in Escherichia coli. *Science (New York, N.Y.)*, 324(5924), pp.255–8.

Lavner, Y. & Kotlar, D., 2005. Codon bias as a factor in regulating expression via translation rate in the human genome. *Gene*, 345(1), pp.127–38.

Letzring, D.P., Dean, K.M. & Grayhack, E.J., 2010. Control of translation efficiency in yeast by codon-

anticodon interactions. *RNA (New York, N.Y.)*, 16(12), pp.2516–2528.

Li, C.H. et al., 2010. eIF5A promotes translation elongation, polysome disassembly and stress granule assembly. *PloS one*, 5(4), p.e9942.

Lorenz, R. et al., 2011. ViennaRNA Package 2.0. *Algorithms for Molecular Biology*, 6, p.26.

Marais, G., Mouchiroud, D. & Duret, L., 2001. Does recombination improve selection on codon usage? Lessons from nematode and fly complete genomes. *Proceedings of the National Academy of Sciences of the United States of America*, 98(10), pp.5688–5692.

Marín, A. et al., 2003. Relationship between G+C content, ORF-length and mRNA concentration in Saccharomyces cerevisiae. *Yeast*, 20(8), pp.703–711.

Martinis, S.A. et al., 1999. Aminoacyl-tRNA synthetases: A new image for a classical family. *Biochimie*, 81(7), pp.683–700.

Mathews, M.B. & Hershey, J.W.B., 2015. The translation factor eIF5A and human cancer. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, 1849(7), pp.836–844.

Mirande, M., 2010. Processivity of translation in the eukaryote cell : Role of aminoacyl-tRNA synthetases. *FEBS Letters*, 584(2), pp.443–447.

Novoa, E.M. et al., 2012. A role for tRNA modifications in genome structure and codon usage. *Cell*, 149(1), pp.202–13.

Novoa, E.M. & Ribas de Pouplana, L., 2012. Speeding with control: Codon usage, tRNAs, and ribosomes. *Trends in Genetics*, 28(11), pp.574–581.

Palidwor, G. a, Perkins, T.J. & Xia, X., 2010. A general model of codon bias due to GC mutational bias. *PloS one*, 5(10), p.e13431.

Pang, Y.L.J. et al., 2014. Diverse cell stresses induce unique patterns of tRNA up- and down-regulation: tRNA-seq for quantifying changes in tRNA copy number. *Nucleic Acids Research*, 42(22), pp.1–10.

Percudani, R., Pavesi, A. & Ottonello, S., 1997. Transfer RNA gene redundancy and translational selection in Saccharomyces cerevisiae. *Journal of molecular biology*, 268(2), pp.322–330.

Pisarev, A. V et al., 2011. The role of ABCE1 in eukaryotic post-termination ribosomal recycling. , 37(2), pp.196–210.

Plant, E.P. et al., 2007. Differentiating between near- and non-cognate codons in Saccharomyces cerevisiae. *PloS one*, 2(6), p.e517.

Plotkin, J. & Kudla, G., 2013. Synonymous but not the same: the causes and consequences of codon bias. *Nature Reviews Genetics*, 18(9), pp.1199–1216.

Pop, C. et al., 2014. Causal signals between codon bias, mRNA structure, and the efficiency of translation and elongation. *Molecular systems biology*, 10, p.770.

Presnyak, V. et al., 2015. Codon optimality is a major determinant of mRNA stability. *Cell*, 160(6), pp.1111–1124.

Quax, T.E.F. et al., 2015. Codon Bias as a Means to Fine-Tune Gene Expression. *Molecular Cell*, 59(2), pp.149–161.

Ratje, A.H. et al., 2010. Head swivel on the ribosome facilitates translocation by means of intra-subunit tRNA hybrid sites. *Nature*, 468(7324), pp.713–716.

Rocha, E.P.C., 2004. Codon usage bias from tRNA's point of view: redundancy, specialization, and efficient decoding for translation optimization. *Genome research*, 14(11), pp.2279–86..

Rodnina, M. V et al., 2005. Recognition and selection of tRNA in translation. *FEBS Letters*, 579(4 SPEC. ISS.), pp.938–942.

Sachs, M.S. & Liu, Y., 2013. Non-optimal codon usage affects expression, structure and function of FRQ clock protei. , 495(7439), pp.111–115.

Sagliocco, F. a et al., 1993. The Influence of 5 ' -Secondary Structures upon Ribosome Binding to mRNA during Translation in Yeast *. *Journal of Biological Chemistry*, 268(35), pp.26522–26530.

Saini, P. et al., 2009. Hypusine-containgin protein eIF5A promotes translation elongation. *Nature*, 459(7243), pp.118–121.

Sambrook, J. & Russell, D.W., 2001. *Molecular cloning: a laboratory manual*,

Sanbonmatsu, K.Y., Joseph, S. & Tung, C., 2005. Simulating movement of tRNA into the ribosome during decoding. *Proceedings of the National Academy of Sciences of the United States of America*, 102(44), pp.15854–9.

Sang Lee, J. et al., 2002. Interaction network of human aminoacyl-tRNA synthetases and subunits of elongation factor 1 complex. *Biochemical and biophysical research communications*, 291(1), pp.158–164.

Scannell, D.R. et al., 2011. The Awesome Power of Yeast Evolutionary Genetics: New Genome Sequences and Strain Resources for the Saccharomyces sensu stricto Genus. *G3 (Bethesda, Md.)*, 1(1), pp.11–25.

Shah, P. & Gilchrist, M. a, 2010. Effect of correlated tRNA abundances on translation errors and evolution of codon usage bias. *PLoS genetics*, 6(9), p.e1001128.

Shah, P. & Gilchrist, M. a, 2011. Explaining complex codon usage patterns with selection for translational efficiency, mutation bias, and genetic drift. *Proceedings of the National Academy of Sciences of the United States of America*, 108(25), pp.10231–10236.

Sonenberg, N. & Hinnebusch, A.G., 2009. Regulation of Translation Initiation in Eukaryotes: Mechanisms and Biological Targets. *Cell*, 136(4), pp.731–745.

Stoletzki, N. & Eyre-Walker, A., 2007. Synonymous codon usage in Escherichia coli: Selection for

translational accuracy. *Molecular Biology and Evolution*, 24(2), pp.374–381.

Strom, A.M. et al., 2014. Probing the global and local dynamics of aminoacyl-tRNA synthetases using all-atom and coarse-grained simulations. *Journal of molecular modeling*, 20(5), p.2245.

Taylor, D.J. et al., 2007. Structures of modified eEF2 80S ribosome complexes reveal the role of GTP hydrolysis in translocation. *EMBO J*, 26(9), pp.2421–2431.

Tuller, T. et al., 2010. Translation efficiency is determined by both codon bias and folding energy. *Proceedings of the National Academy of Sciences*, 107(8), pp.3645–3650.

Valle, M. et al., 2002. Cryo-EM reveals an active role for aminoacyl-tRNA in the accommodation process. *EMBO J*, 21(13), pp.3557–3567.

Varshavsky, a, 1996. The N-end rule: functions, mysteries, uses. *Proceedings of the National Academy of Sciences of the United States of America*, 93(October), pp.12142–12149.

de Wet, J.R. et al., 1987. Firefly luciferase gene: structure and expression in mammalian cells. *Molecular and cellular biology*, 7(2), pp.725–737.

Wilson, D.N. & Cate, J.H.D., 2015. The Structure and Function of the Eukaryotic. *Cold Spring Harbor Perspect Biol Perspect Biol*, 4:a011536, pp.1–18.

Yang, X. et al., 2006. Two conformations of a crystalline human tRNA synthetase-tRNA complex: implications for protein synthesis. *The EMBO journal*, 25(12), pp.2919–2929.

Yona, A.H. et al., 2013. Trna genes rapidly change in evolution to meet novel translational demands. *eLife*, 2013(2), pp.1–17.

Yu, C. et al., 2015. Codon Usage Influences the Local Rate of Translation Elongation to Regulate Co-translational Protein Folding. *Molecular cell*, 59(5), pp.744–754.

Zeng, X. et al., 2014. Flipping of the Ribosomal A-Site Adenines Provides a Basis for tRNA Selection. *Journal of Molecular Biology*, 426(19), pp.3201–3213.

Zheng, G. et al., 2015. Efficient and quantitative high-throughput transfer RNA sequencing. *Nature methods*, 12(9), pp.835–837.

Zhou, M. et al., 2013. Non-optimal codon usage affects expression, structure and function of clock protein FRQ. *Nature*, 495(7439), pp.111–5.

Zhou, M. et al., 2015. Nonoptimal codon usage influences protein structure in intrinsically disordered regions. *Molecular Microbiology*, 97(5), pp.974–987.

Zouridis, H. & Hatzimanikatis, V., 2008. Effects of Codon Distributions and tRNA Competition on Protein Translation. *Biophysical Journal*, 95(3), pp.1018–1033.

Zur, H. & Tuller, T., 2012. Strong association between mRNA folding strength and protein abundance in S. cerevisiae. *Embo Reports*, 13(3), pp.272–277.