12-21-2017

Transcriptome-Wide Analysis of Roles for Transfer RNA Modifications in Translational Regulation

Hsin-Jung Chou
University of Massachusetts Medical School

Follow this and additional works at: https://escholarship.umassmed.edu/gsbs_diss

Part of the Bioinformatics Commons, and the Molecular Biology Commons

This work is licensed under a Creative Commons Attribution-Noncommercial 4.0 License

Recommended Citation
https://escholarship.umassmed.edu/gsbs_diss/943

This material is brought to you by eScholarship@UMMS. It has been accepted for inclusion in GSBS Dissertations and Theses by an authorized administrator of eScholarship@UMMS. For more information, please contact Lisa.Palmer@umassmed.edu.
TRANSCRIPTOME-WIDE ANALYSIS OF ROLES FOR TRANSFER RNA MODIFICATIONS IN TRANSLATIONAL REGULATION

A Dissertation Presented

By

HSIN-JUNG CHOU

Submitted to the Faculty of the University of Massachusetts Graduate School of Biomedical Sciences, Worcester in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

December 21st, 2017

Interdisciplinary Graduate Program
TRANSCRIPTOME-WIDE ANALYSIS OF ROLES FOR TRANSFER RNA MODIFICATIONS IN TRANSLATIONAL REGULATION

A Dissertation Presented
By
HSIN-JUNG CHOU

The signatures of the Dissertation Defense Committee signify completion and approval as to style and content of the Dissertation

Oliver J. Rando, MD, PhD, Thesis Advisor

Paul Kaufman, PhD, Member of Committee

Andrei Korostelev, PhD, Member of Committee

Marian Walhout, PhD, Member of Committee

Peter Dedon, MD, PhD, External Member of Committee

The signature of the Chair of the Committee signifies that the written dissertation meets the requirements of the Dissertation Committee

Allan Jacobson, PhD, Chair of Committee

The signature of the Dean of the Graduate School of Biomedical Sciences signifies that the student has met all graduation requirements of the school

Anthony Carruthers, PhD, Dean of the Graduate School of Biomedical Sciences

Interdisciplinary Graduate Program
December 21st, 2017
This work is dedicated to
my beloved husband, best friend, and colleague

Stanley Hsieh 謝宗翰
Acknowledgments

This work was contributed greatly from the assistance and discussion of mentors, colleagues, and collaborators. First, I acknowledge my thesis advisor Oliver Rando, who gave me the opportunity to join this magical lab in my 4th year. He has been extremely reachable and helpful, no matter for what kinds of questions or problems. He is encouraging to my idea and decision, and incredibly motivating when things did not go well. Also, his idea and work contribute substantially to this study. Next, I thank members in the Rando lab for their help and scientific discussion, especially Stanley, Hsiuyi, Amanda, Caitlin, and David for their instructions of yeast genetic, deep sequencing, and ribosome profiling, and Tobias for assisting with yeast strain construction. I must thank my bioinformatics collaborator Elisa Donnard in Garber lab for turning analysis idea into functional codes. In addition, I thank the members of my TRAC committee for their supports and constructive suggestions. A great portion of the results in this study was inspired from Allan and Andrei’s comments. Finally, I want to deeply thank my family in Taiwan and my husband Stanley for their endless love and support. Their faith on me gave me the strength to achieve something I thought was impossible.
Abstract

Covalent nucleotide modifications in RNAs affect numerous biological processes, and novel functions are continually being revealed even for well-known modifications. Among all RNA species, transfer RNAs (tRNAs) are highly enriched with diverse modifications, which are known to play roles in decoding and tRNA stability, charging, and cellular trafficking. However, studies of tRNA modifications have been limited in a small scale and performed by groups with different methodologies. To systematically compare the functions of a large set of noncoding RNA modifications in translational regulation, I carried out ribosome profiling in 57 budding yeast mutants lacking nonessential genes involved in tRNA modifications. Deletion mutants with enzymes known to modify the anticodon loop or non-tRNA substrates such as rRNA exhibited the most dramatic translational perturbations, including altered dwell time of ribosomes on relevant codons, and altered ribosome density in protein-coding regions or untranslated regions of specific genes. Several mutants that result in loss of tRNA modifications in locations away from the anticodon loop also exhibited altered dwell time of ribosomes on relevant codons. Translational upregulation of the nutrient-responsive transcription factor Gcn4 was observed in roughly half of the mutants, consistent with the previous studies of Gcn4 in response to numerous tRNA perturbations. This work also discovered unexpected roles for tRNA modifying enzymes in rRNA 2′-O-methylation, and in transcriptional
regulation of TY retroelements. Taken together, this work revealed the
importance and novel functions of tRNA modifications, and provides a rich
resource for discovery of additional links between tRNA modifications and gene
regulation.
Table of Contents

Title Page          i
Signature Page         ii
Dedication          iii
Acknowledgements         iv
Abstract          v
Table of Contents         vii
List of Figures         viii
List of Tables          x
List of Copyrighted Materials Produced by The Author xi
List of Third Part Copyrighted Material  xii
Preface            xiii

Chapter I: Introduction

    Part 1: tRNA Modifications        1
    Part 2: Ribosome Profiling        28

Chapter II: Results     36

Chapter III: Discussion  98

Bibliography                     114
List of Figures

Figure 1.1  Structures of modified nucleosides in tRNA
Figure 1.2  Location of modified nucleosides in tRNA
Figure 1.3  Synthesis of 5-aminomethyl uridine derivatives (xm⁵U)
Figure 1.4  Structure of stacking interactions of t⁶A37
Figure 1.5  Biosynthesis pathway of yW in S. cerevisiae
Figure 1.6  Ribosome profiling
Figure 2.1  Overview of dataset
Figure 2.2  Example of dataset reproducibility
Figure 2.3  Minor effects of cycloheximide on global ribosome occupancy
Figure 2.4  Overview of codon-level analysis of ribosome occupancy
Figure 2.5  Mutant effects on codon occupancy
Figure 2.6  Unanticipated mutant effects on codon occupancy
Table 2.6  Codon occupancy dataset
Figure 2.7  Dramatic effects of trm7 on ribosome occupancy profiles
Figure 2.8  RiboMeth-seq analysis of rRNA 2′-O-methylation in WT yeast
Figure 2.9  Trm7 effects on rRNA 2′-O-methylation
Figure 2.10  Comparison of RNA-Seq and RPF datasets
Figure 2.11  Effects of tRNA-Modifying Enzymes on RNA Abundance
Figure 2.12  Analysis of silencing-related phenotypes
Figure 2.13  Downregulation of TY1 expression in Elp-related mutants
Figure 2.14  Codon-level pauses have minimal impact on ORF-level ribosome occupancy data
Figure 2.15  Effects of tRNA-modifying enzymes on translational efficiency
Figure 3.1  Dual luciferase assay to determine elp-responsive elements
List of Tables

Table 1.1  
*S. cerevisiae* genes involved in tRNA modifications

Table 1.2  
tRNA species containing *xm^5U34* in yeast

Table 1.3  
m^5C locations in *S. cerevisiae* tRNAs

Table 2.1  
RNA-Seq dataset

Table 2.2  
Ribosome footprinting dataset

Table 2.3  
Translational efficiency

Table 2.4  
No cycloheximide ribosome footprinting dataset

Table 2.5  
Ribosome-protected fragment lengths

Table 2.6  
Codon occupancy dataset

Table 2.7  
RiboMeth-Seq dataset
List of Copyrighted Materials Produced by The Author

Chapter II is reprinted and adapted from the following co-authored work published by Elsevier:


Stated by the publisher, the author of this article retains the right to include it in a thesis or dissertation as long as it is not published commercially, and permission is not required.
# List of Third Part Copyrighted Material

<table>
<thead>
<tr>
<th>Figures</th>
<th>Publisher</th>
<th>License #</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.2</td>
<td>Annual Reviews</td>
<td>Not required</td>
</tr>
<tr>
<td>Figure 1.3</td>
<td>Annual Reviews</td>
<td>Not required</td>
</tr>
<tr>
<td>Figure 1.4</td>
<td>Springer Nature</td>
<td>4262910912964</td>
</tr>
<tr>
<td>Figure 1.5</td>
<td>John Wiley and Sons</td>
<td>4242910566753</td>
</tr>
<tr>
<td>Figure 1.6 A B</td>
<td>Elsevier</td>
<td>4250161258944</td>
</tr>
<tr>
<td>Figure 1.6 C D</td>
<td>Elsevier</td>
<td>Not required</td>
</tr>
</tbody>
</table>
Preface

The work in Chapter II was published as:


This work was contributed by Hsin-Jung Chou (H.-J.C.), Elisa Donnard (E.D.), H. Tobias Gustafsson (H.T.G.), Manuel Garber (M.G.), and Oliver J. Rando (O.J.R.). All experiments were conceived by H.-J.C. and O.J.R. and performed by H.-J.C. H.T.G. assisted with tye7Δ strain construction. Data analysis was carried out by H.-J.C., E.D., M.G., and O.J.R. Manuscript was written by H.-J.C. and O.J.R. and edited by all authors.
CHAPTER I: INTRODUCTION

Part 1: tRNA Modifications

General Introduction

All RNA species consist of not only the four common nucleotides A, C, G, and U, but also covalently modified nucleotides at a relatively lower abundance. Modified nucleotides have been implicated in a wide variety of biological roles, such as RNA stability, decoding, and regulation of gene expression (Frye et al., 2016). In eukaryotic messenger RNAs (mRNAs), the most common modification is the 7-methylguanylate cap at the 5’ end, which is vital for the integrity and function of mature mRNAs. Base or ribose methylations and pseudouridines are widespread modifications in ribosomal RNAs (rRNAs) in all branches of the tree of life and affect the structure and function of rRNAs.

Covalently modified nucleotides are particularly enriched in transfer RNAs (tRNAs). In fact, tRNAs have the highest portion of modified nucleotides among all RNA species. On average, about 15% of nucleotides in an individual tRNA are modified (Hopper, 2013; Machnicka et al., 2013; Phizicky and Hopper, 2010), although different cytosolic tRNA species have distinct frequencies of modified nucleotides, with a range of 7-17 modifications per nuclear encoded tRNA. For example, cytosolic tRNA-Lys-UUU has the most abundant (22%) modified
nucleotides among others, while tRNA-Glu-UUC and tRNA-Gly-UCC have the lowest frequency (9%) of modified nucleotides.

In addition, tRNAs have extremely varied types of modified nucleotides, with roughly 85 modifications identified across all kingdoms (El Yacoubi et al., 2012). Sequenced tRNAs revealed the chemical nature and position of almost all modifications in *Escherichia coli* and yeast *Saccharomyces cerevisiae* (Juhling et al., 2009), and the latter is the best characterized species for tRNA modifications among all eukaryotes. In *S. cerevisiae*, there are 25 known modified nucleotides (Figure 1.1 and 1.2 for their chemical structures and locations in tRNA, respectively). Almost half the modifications are present at positions 34 and 37, which are critical for decoding and stabilizing the codon-anticodon interaction (El Yacoubi et al., 2012; Phizicky and Hopper, 2010). The rest of the modifications are scattered elsewhere in the tRNA molecule and are responsible for stabilizing the tRNA structure (Motorin and Helm, 2010). Comprehensive reports for the tRNA sequences as well as the structures and positions of the modifications can be found in the tRNAdb database ([http://trna.bioinf.uni-leipzig.de/](http://trna.bioinf.uni-leipzig.de/); (Juhling et al., 2009)) and the Modomics database ([http://modomics.genesilico.pl/](http://modomics.genesilico.pl/); (Boccaletto et al., 2017; Czerwoniec et al., 2009; Dunin-Horkawicz et al., 2006; Machnicka et al., 2013).

tRNA modifications are thought to be crucial for decoding. As an essential component in protein synthesis, aminoacyl-tRNAs decode mRNAs via the pairing of tRNA anticodons and mRNA codons. The first two bases of the codon triplet
pair with the third and second bases of the anticodon (position 36 and 35 in tRNAs, respectively) obeying the Watson-Crick pairing rules. In contrast, the third base of the codon interacts with the first base of the anticodon (position 34) following the wobble hypothesis that allows less constrained and nonstandard pairings such as G:U (Crick, 1966). Thus, one tRNA molecule can theoretically decode several codons, and it is supported by the fact that the combinations of four nucleotides in anticodons of sequenced tRNAs are way less than the number of possible codons. Modifications in tRNAs, especially the ones in position 34, add chemical diversity to the four canonical nucleotides and thus enhance decoding accuracy. Some modifications such as inosine allow a relaxed pairing for a tRNA to decode multiple codons (Alseth et al., 2014), while other modifications such as 5-aminomethyl uridine derivatives improve the stringency of decoding to discriminate cognate codons from closely related codons (Yokoyama et al., 1985).

Although the majority of tRNA modifications were identified over 40 years ago, the roles of these modifications remained puzzling for decades. Powered by advanced techniques, recent studies using genetic, genomic, biochemical, and bioinformatic approaches uncovered a whole set of genes that are required for the synthesis of tRNA modifications (El Yacoubi et al., 2012; Phizicky and Hopper, 2010). These studies have shed light on the biological functions of tRNA modifications and increased the appreciation of the importance of these diverse cellular regulators. tRNA modifications have been implicated in a plethora of
cellular mechanisms including translation initiation and elongation (Liu et al., 2016), tRNA stability (Alexandrov et al., 2006; Motorin and Helm, 2010), charging, and cellular trafficking (Kramer and Hopper, 2013). Moreover, they have been implicated in a range of biological processes such as early development, neurodevelopment, and responses to environmental changes or stress (El Yacoubi et al., 2012; Phizicky and Hopper, 2010).

However, the roles of many modifications still are not fully understood. Do the modifications on the anticodon have functions other than accurate decoding? Do the modifications outside of the anticodon affect translation elongation? How do these modifications affect translation? How do they regulate gene expression at both translational and non-translational levels? These questions remain unexplored. In addition, each study to date has assayed the function of specific tRNA modifications differently and in a small number of mutants. The narrow scope of studying a handful of modifications combined with the experimental variance resulting from different techniques and different laboratories prevent researchers from comprehensively comparing the effects of functionally-related families of tRNA modifications.

To date, 72 yeast genes have been annotated with involvement in tRNA modifications (Table 1.1), of which 14 are essential and not included in this study. The remaining 58 non-essential genes are discussed in this work. The next section focuses on the synthesis pathway and roles of specific tRNA modifications, and the biological functions of these non-essential genes in S.
*cerevisiae*. Most of the topics discussed here are tRNAs encoded by the nuclear genome instead of the mitochondrial genome. Although the following discussions are studies mostly from budding yeast, critical discoveries from other model organisms as well as their similarities and differences to yeast studies also are described.
Figure 1.1. Structures of modified nucleosides in tRNA
25 known modified nucleosides in yeast tRNAs are shown with chemical structures, full names, and abbreviations in brackets. Modified parts are labeled in red, except for pseudouridine. Pictures adapted from Modomics database (Machnicka et al., 2013).
Figure 1.2. Location of modified nucleosides in tRNA

Known modifications in yeast tRNA are labeled on a tRNA with their position numbers. The majority of positions have one modification, while some positions may have different modifications presented in distinct tRNA species, such as position 34 and 37. Anticodons are in red circles. Picture adapted from (El Yacoubi et al., 2012).
Table 1.1. *S. cerevisiae* genes involved in tRNA modifications

<table>
<thead>
<tr>
<th>Yeast gene</th>
<th>Modification</th>
</tr>
</thead>
<tbody>
<tr>
<td>BUD32, CGI121, GON7, KAE1, PCC1, SUA5</td>
<td><em>t</em>⁶A³⁷</td>
</tr>
<tr>
<td>DUS1</td>
<td>D16, D17</td>
</tr>
<tr>
<td>DUS2</td>
<td>D20</td>
</tr>
<tr>
<td>DUS3</td>
<td>D47</td>
</tr>
<tr>
<td>DUS4</td>
<td>D20a, D20b</td>
</tr>
<tr>
<td>ELP1/IKI3, ELP2, ELP3, ELP4, ELP5, ELP6, KTI11/DPH3, KTI12, KTI13, KTI14, SIT4, SAP185, SAP195</td>
<td><em>ncm</em>⁵U³⁴, <em>ncm</em>⁵Um³⁴, <em>mcm</em>⁵U³⁴, <em>mcm</em>⁵s²U³⁴</td>
</tr>
<tr>
<td>MOD5</td>
<td><em>i</em>⁶A³⁷</td>
</tr>
<tr>
<td>NTS1, ISU1, ISU2, CFD1, NBP35, CIA1, URM1, UBA4, NCS2, NCS6/TUC1, TUM1</td>
<td><em>mcm</em>⁵s²U³⁴</td>
</tr>
<tr>
<td>PUS1</td>
<td>Ψ⁶26, Ψ⁶27, Ψ⁶28, Ψ⁶34, Ψ⁶35, Ψ⁶36, Ψ⁶65, Ψ⁶67</td>
</tr>
<tr>
<td>PUS2</td>
<td>Mitochondrial Ψ⁶27 and Ψ⁶28</td>
</tr>
<tr>
<td>PUS3/DEG1</td>
<td>Ψ⁶38, Ψ⁶3⁹</td>
</tr>
<tr>
<td>PUS4</td>
<td>Ψ⁶55</td>
</tr>
<tr>
<td>PUS6</td>
<td>Ψ⁶3¹</td>
</tr>
<tr>
<td>PUS7</td>
<td>Ψ⁶13, Ψ⁶3⁵</td>
</tr>
<tr>
<td>PUS8/RIB2</td>
<td>Ψ⁶3²</td>
</tr>
<tr>
<td>PUS9</td>
<td>Mitochondrial Ψ⁶3²</td>
</tr>
<tr>
<td>RIT1</td>
<td>Arp64</td>
</tr>
<tr>
<td>TAD1</td>
<td>Ψ³⁷</td>
</tr>
<tr>
<td>TAD2, TAD3</td>
<td>Ψ³⁴</td>
</tr>
<tr>
<td>TAN1</td>
<td>ac⁴C¹²</td>
</tr>
<tr>
<td>TRM1</td>
<td>m²⁶G²⁶</td>
</tr>
<tr>
<td>TRM2</td>
<td>m⁴U⁵⁴</td>
</tr>
<tr>
<td>TRM3</td>
<td>Gm¹⁸</td>
</tr>
<tr>
<td>Enzyme/Gene</td>
<td>Modifications</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------------</td>
</tr>
<tr>
<td>TRM4/NCL1</td>
<td>$m^5\text{C}27$, $m^5\text{C}34$, $m^5\text{C}48$, $m^5\text{C}49$, $m^6\text{C}50$, $m^6\text{C}56$</td>
</tr>
<tr>
<td>TRM5</td>
<td>$m^1\text{G}37$, $m^1\text{I}37$, yW37</td>
</tr>
<tr>
<td>TRM6/GCD10, TRM61/GCD14</td>
<td>$m^1\text{A}58$</td>
</tr>
<tr>
<td>TRM7, TRM732</td>
<td>Cm32</td>
</tr>
<tr>
<td>TRM7, TRM734/RTT10</td>
<td>Cm34, Gm34, ncm$^b$U$m34$</td>
</tr>
<tr>
<td>TRM8, TRM82</td>
<td>$m'\text{G}46$</td>
</tr>
<tr>
<td>TRM9, TRM112</td>
<td>ncm$^b$U$34$, ncm$^b$s$^2$U$34$</td>
</tr>
<tr>
<td>TRM10</td>
<td>$m^1\text{G}9$</td>
</tr>
<tr>
<td>TRM11, TRM112</td>
<td>$m^2\text{G}10$</td>
</tr>
<tr>
<td>TRM13</td>
<td>Am4, Gm4, Cm4</td>
</tr>
<tr>
<td>TRM44</td>
<td>Um44</td>
</tr>
<tr>
<td>TYW1, TYW2/TRM12, TYW3, TYW4</td>
<td>yW37</td>
</tr>
</tbody>
</table>

The table lists various RNA modifications catalyzed by different enzymes or gene products.
**Synthesis and Roles of Specific tRNA Modifications**

tRNA modifications are generated by specific tRNA modifying enzymes. Simple modifications such as 5-methylation or pseudouridine are catalyzed by single modifying enzymes. In contrast, some modifications require additional proteins such as scaffold proteins or kinases to initiate modification activity or to serve for site selection on tRNAs. For instance, complex and unusual modifications like 5-methoxycarbonylmethyl-2-thiouridine (mcm⁵s²U) or wybutosine (yW) require catalysis by multiple enzymes in several steps (El Yacoubi et al., 2012). Almost all complex modifications that require more than two enzymes in the synthesis pathway are located at positions 34 and 37.

Recent studies deleting specific genes responsible for distinct modifications revealed the consequence of loss of certain modifications and the functions of these modifying proteins. In addition to roles in protein synthesis, some modifying enzymes are implicated in multiple biological aspects such as DNA integrity, telomere homeostasis, chromatin remodeling, and transcription (El Yacoubi et al., 2012). The following sections describe the biosynthesis pathway of specific modifications, the modifying enzymes involved, and the functions of these enzymes.

**5-aminomethyl uridine derivatives (xm⁵U) at position 34**

There are four xm⁵U at position 34 in the budding yeast: 5-carbamoylmethyl-uridine (ncm⁵U), 5-carbamoylmethyl-2'-O-methyluridine (ncm⁵Um), 5-methoxy-
carbonylmethyluridine (mcm⁵U), and 5-methoxycarbonylmethyl-2-thiouridine (mcm⁵s²U). They are present in distinct tRNA species (Table 1.2), and their formation requires at least 25 proteins in a sequential synthesis pathway (Figure 1.3).

**Figure 1.3. Synthesis of 5-aminomethyl uridine derivatives (xm⁶U)**
Current model of the synthesis of xm⁶U at position 34 in yeast. R indicates the tRNA molecules in the structures of the modified base. Picture and legend adapted from (El Yacoubi et al., 2012).
Table 1.2. tRNA species containing \( \text{xm}^5\text{U34} \) in yeast

<table>
<thead>
<tr>
<th>Modifications</th>
<th>tRNA species</th>
</tr>
</thead>
<tbody>
<tr>
<td>ncm(^5\text{U})</td>
<td>Ser-UGA, Val-UAC, Pro-UGG, Thr-UGU, Ala-UGC</td>
</tr>
<tr>
<td>ncm(^5\text{Um})</td>
<td>Leu-UAA</td>
</tr>
<tr>
<td>mcm(^5\text{U})</td>
<td>Arg-UCU, Gly-UCC</td>
</tr>
<tr>
<td>mcm(^5\text{s2U})</td>
<td>Lys-UUU, Gln-UUG, Glu-UUC</td>
</tr>
</tbody>
</table>

The synthesis pathway starts from the formation of ncm\(^5\text{U}\), which is catalyzed by the Elongator Protein (Elp) complex consisting of six subunits Elp1/Iki3, Elp2, Elp3, Elp4, Elp5/Iki1, and Elp6 (Huang et al., 2005; Johansson et al., 2008). The modification activity of the Elp complex is regulated by phosphorylation at specific residues by several kinases (Kti11-Kti13/Ats1 heterodimer, Kti12, and Kti14/Hrr25), and by dephosphorylation by the protein phosphatase Sit4 and two associated proteins Sap185 and Sap190 (Abdel-Fattah et al., 2015; Jablonowski et al., 2004; Zabel et al., 2008). ncm\(^5\text{U}\) can be either 2'-O-methylated by Trm7-Trm734/Rtt10 heterodimer to form ncm\(^5\text{Um}\) (Guy et al., 2012; Pintard et al., 2002), or be converted into mcm\(^5\text{U}\) by a complex consisting of methyltransferase Trm9 and an activating scaffold protein Trm112 (Chen et al., 2011a; Kalhor and Clarke, 2003; Letoquart et al., 2015; Mazauric et al., 2010).

mcm\(^5\text{U}\) can be further 2-thiolated by tRNA targeting enzymes Ncs2 and Nsc6/Tuc1, which receive sulfur atoms that go through a sulfur transfer cascade
The sulfur atoms originate from cysteine and are carried by the cysteine desulfurase Nfs1 as an enzyme-bound persulfide in the mitochondria (Lill and Muhlenhoff, 2008; Nakai et al., 2004). The persulfide sulfurs are transferred from Nfs1 to Tum1, which acts as a sulfur carrier and an activator of Nfs1, and may serve as a sulfur shuttle protein between the mitochondria and cytosol (Noma et al., 2009). Subsequently, the sulfur atoms are passed to an E1-like protein Uba4 and then transferred to a ubiquitin-like protein Urm1. Eventually Urm1 transfers the sulfur atoms to tRNAs with the help of the Ncs2-Ncs6 complex. In addition, several proteins involved in the iron-sulfur cluster assembly machinery are required in 2-thiolation of cytosolic tRNAs, including three essential cytosolic proteins (Cfd1, Cia1, and Nbp35) and two mitochondrial scaffold proteins (Isu1 and Isu2) (Nakai et al., 2007). However, the role of these proteins in 2-thiolation synthesis pathway is not yet clear.

Loss of almost any one of these genes in the synthesis pathway results in an absence or decreased levels of distinct xmn5U modifications (Huang et al., 2008). As the E1p complex, Kti proteins, and Sit4 function upstream in the synthesis pathway, deletion of these genes leads to a reduction (kti13Δ) or an absence (other mutants) of nmn5U, nmn5U, and mcm5s2U. Although Sap185 and Sap190 are Sit4-associated proteins, only a double deletion mutant exhibits the same defect (Huang et al., 2008; Jablonowski et al., 2004), indicating that they may be functionally redundant. Lack of either TRM7 or RTT10 causes an absence of ncm6Um (Guy et al., 2012). Removal of TRM9, but not TRM112
which codes for a dispensable scaffold protein, results in a dramatic decrease in mcm$^5$U and mcm$^5$s$^2$U (Huang et al., 2008). Mutations in any one of the genes involved in 2-thiolation (except for ISU1 and ISU2, which are functionally redundant) leads to a reduction of mcm$^5$s$^2$U and an accumulation of the precursor mcm$^5$U.

xm$^5$U at position 34 plays an essential role in accurate decoding of optimal codons and translation fidelity. A tRNA containing an unmodified U34 can pair with codons ending in A and G, and possibly U and C as well (Crick, 1966; Lim, 1994). It can be an issue for codons in split codon boxes, in which the four codons have the same first two nucleotides, but the purine-ending codons encode for one amino acid and the pyrimidine-ending codons encode for another. mcm$^6$U and mcm$^5$s$^2$U are present in tRNAs that decode A-ending codons in two-split codon boxes (with an exception of tRNA-Gly-mcm$^5$UCC that encodes codons in an unsplit box), and enhance translation fidelity by promoting the codon discrimination by tRNAs. For example, mcm$^5$s$^2$U is stable in a C3'-endo form, which is a rigid confirmation that allows the tRNA pairs with A but not U (Yokoyama et al., 1985). This modification also enhances ribosomal A-site binding of aminoacyl-tRNA on a cognate codon and dipeptide formation in vitro (Rezgui et al., 2013). Likewise, mcm$^5$ modification enhances tRNA pairing with codons ending in both A and G, preventing misreading of pyrimidine-ending codons (Johansson et al., 2008). Moreover, a study using reporter assays with
slippery codons, which tend to result in ribosomal frameshifting, revealed that xm5U34 is important for reading frame maintenance (Tukenmez et al., 2015).

Lack of these modifications results in impaired proteome integrity and proteotoxic stress. Using ribosome profiling, a deep-sequencing method to monitor translation in vivo by mapping ribosome-occupied mRNAs, two groups uncovered that AAA, CAA, and/or GAA codons in elp3Δ, elp6Δ, ncs2Δ, ncs6Δ, and uba4Δ are more frequently occupied by ribosomes compared to wildtype. This suggests that ribosomes slow or stall on the codons that normally are decoded by tRNAs that contain mcm5s2 (Nedialkova and Leidel, 2015; Zinshteyn and Gilbert, 2013). The increased dwell time of ribosomes on these codons subsequently triggers protein misfolding and aggregation, and gene expression profile of proteotoxic stress (Nedialkova and Leidel, 2015). Overexpression of three corresponding tRNAs (tRNA-Lys-UUU, tRNA-Glu-UUC, tRNA-Gln-UUG) rescued these phenotypes, indicating that these abnormalities result from defective tRNAs. Moreover, translational upregulation of the nutrient-responsive transcription factor Gcn4 was observed in the mutants in these two studies, indicating de-repression of Gcn4 in response to the defect in anticodon modifications.

In addition to a general defect in protein integrity, loss of xm5U modifications leads to alteration of specific protein expression. Seven codons (AAA, CAA, GAA, AGA, GGA, GGG, and AGG) are decoded by tRNAs that contain Trm9-dependent modifications mcm5U or mcm5s2U. trm9Δ cells exhibit a
decrease of $\beta$-galactosidase activity from AGA-enriched lacZ reporter but an increase of the activity from the reporter enriched in near-cognate codon AGG (Begley et al., 2007). The protein level of several AGA-enriched genes ($YEF3$, $RNR1$, and $RNR3$) is also reduced in vivo. Another group reported the expression level of thousands of proteins in $trm9\Delta$ using quantitative proteomics (SILAC coupled to LC-MS/MS; Deng et al., 2015). 36.5% and 23% of proteins that are overrepresented with AGA and GAA codons, respectively, are significantly down-regulated in $trm9\Delta$. These down-regulated proteins are highly enriched in translation machinery, including aminoacyl-tRNA synthetases, proteins in both 40S and 60S ribosomal subunits, and proteins involved in translation initiation, elongation, and termination. That may explain the phenotypes observed in $trm9\Delta$, including protein errors and activation of protein stress response pathways (Begley et al., 2007; Patil et al., 2012). In another example, genes that are highly enriched in AAA, CAA, and GAA codons are preferentially down-regulated in $urm1\Delta$, which lacks 2-thiolation in tRNAs decoding these three codons (Rezgui et al., 2013).

Studies utilizing mutants that lack $xm^{5}U34$ revealed other functions of these tRNA modifying enzymes, including involvement in transcription elongation, polarized exocytosis, telomeric gene silencing, and DNA damage response. The Elp complex was originally identified as a component of hyperphosphorylated RNA polymerase II holoenzyme in transcription elongation (Otero et al., 1999), as Elp3 is a histone acetyltransferase (HAT) (Winkler et al.,
Deletion of ELP1 or ELP3 results in slow growth, delayed transcription of specific genes, and an acetylation defect of lys14 in histone H3 in vitro. In addition, Elp1 is implicated in the regulation of exocytosis independent from its function in transcription elongation (Rahl et al., 2005). Moreover, the absence of any one subunit of the Elp complex leads to a partial loss of silencing of reporter genes located in subtelomeric regions and the HMR mating locus, as well as increased sensitivity to DNA damage agents (Li et al., 2009b). Surprisingly, overexpression of hypomodified tRNA-Lys-UUU, tRNA-Glu-UUC, and tRNA-Gln-UUG can rescue all the above-mentioned phenotypes in the mutants, arguing that the defects from a loss of the Elp complex is caused by inefficient translation due to lack of tRNA modifications on the wobble uridine (Chen et al., 2011b; Esberg et al., 2006).

Interestingly, studies have linked uridine thiolation at position 34 to the response to environmental changes and stress. The level of mcm$^5$s$^2$U in wildtype yeast responds to the exposure to several chemicals such as H$_2$O$_2$, NaAsO$_2$, and methylmethane sulfonate (MMS) (Chan et al., 2010). Thiolation of specific tRNAs in S. cerevisiae reflects the levels of sulfur-containing amino acids cysteine and methionine in the growth media, and lack of tRNA thiolation in mutants affects the translation of relevant biosynthetic proteins (Laxman et al., 2013). In summary, at the molecular level, xm$^5$ modifications at U34 play essential roles in translation and multiple biological processes, and have a potential role in cellular response to environmental stimuli. At the organismal level, the Elp complex is
implicated in neurodevelopment in *C. elegans* (Chen et al., 2009) and several human neurological diseases including familial dysautonomia, intellectual disability, amyotrophic lateral sclerosis, and rolandic epilepsy (Bednarova et al., 2017). Although evidence links neurological dysfunctions to defects in tRNA modification or other functions of the Elp complex in the nervous system, the underlying pathogenesis remains elusive.

**2’-O-methylation (Xm) at the anticodon loop**

2’-O-methylations are found at positions 32 and 34 in tRNAs from bacteria, archaea, and eukaryotes (Machnicka et al., 2013). In *S. cerevisiae*, 2’-O-methylations include Cm32 in tLeu-UAA, tPhe-GAA, and tTrp-CCA, Gm34 in tPhe-GAA, Cm34 in tTrp-CCA, and the above-mentioned ncm⁵Um34 in tLeu-UAA. These methylations are generated by Trm7 methyltransferase and distinct partners Trm732 (required for 2’-O-methylation of C32) or Trm734/Rtt10 (required for 2’-O-methylation at position 34; Guy et al., 2012). The requirement of Trm7, Trm732, and Trm734/Rtt10 for tRNA 2’-O-methylation is conserved throughout eukaryotes, and the homologs of these enzymes are found in at least 14 organisms (Guy and Phizicky, 2015).

*TRM7* was identified by searching for *S. cerevisiae* homologs of *E. coli* 2’-O-methyltransferase FtsJ/RrmJ, which catalyzes the formation of 2’-O-methylation in both tRNA and rRNA (Pintard et al., 2002). Three homologs were identified for 2’-O-methylation at distinct substrates: Spb1 is involved in
methylations of 27S pre-rRNA, Mrm2 methylates mitochondrial 21S rRNA, and Trm7 generates Cm32 and Nm34 in tRNAs. In addition to loss of 2'-O-methylations, \textit{trm7}\Delta mutant exhibits slow growth and decreased level of polysomes compared to wildtype (Pintard et al., 2002). The slow growth defect can be rescued by overexpression of tPhe-UAA but not the other two affected tRNAs (Guy et al., 2012), arguing that the sickness of \textit{trm7}\Delta is due to the dysfunction of tRNA-Phe. Since the formation of Cm32 and Gm34 in tRNA-Phe drives the formation of wybutosine (yW) at position 37 from its precursor 1-methylguanosine (m\textsuperscript{1}G37), and that \textit{trm7}\Delta exhibits decreased level of yW and dramatically increased level of m\textsuperscript{1}G, it is unclear that the defect of \textit{trm7}\Delta is due to lack of 2'-O-methylation or yW.

\textit{trm7}\Delta is sensitive to H\textsubscript{2}O\textsubscript{2} and MMS treatments, and this sensitivity is likely caused by loss of 2'-O-methylation rather than loss of yW, as trm5\Delta, which is responsible for the formation of m\textsuperscript{1}G and thus yW, is resistant to the same treatments (Chan et al., 2010). Mutations in human methyltransferase \textit{FTSJ1}, the likely \textit{TRM7} homolog, is associated with nonsyndromic X-linked mental retardation (Freude et al., 2004; Ramser et al., 2004). The mutations in \textit{FTSJ1} are also linked to reduced level of 2'-O-methylation, with that Gm34 of tRNA-Phe may be a critical modification (Guy et al., 2015). Despite the fact that Trm7 is indispensable in tRNA function and translation, its detailed functions remain unknown. It is unclear how 2'-O-methylations at the anticodon loop of tRNAs contribute to translational fidelity and efficiency, and how the defects of specific
tRNA species affect gene expression and protein levels in cells. It is also unknown if Trm7 plays roles in other cellular processes apart from modifications of specific tRNA substrates.

While cytosolic Trm734/Rtt10 functions in tRNA modification, membrane-bound Trm734/Rtt10 has been implicated in endoplasmic recycling (Shi et al., 2011). TRM734/RTT10 was also identified in a screen for genes that alter yeast Ty1 retrotransposition (Nyswaner et al., 2008). Deletion of TRM734/RTT10 results in a dramatic increase in Ty1 mobility, and this alteration might be linked to the defect of 2'-O-methylation in tRNAs since TRM7 was also identified in the screen.

5-methylcytosine (m\textsuperscript{5}C)

5-methylcytosine (m\textsuperscript{5}C) occurs on both DNA and RNA. Eukaryotic m\textsuperscript{5}C of DNA has been studied intensively over the past years, and this DNA modification is traditionally considered as a marker for inactivate transcription in gene regulatory regions and as an epigenetic regulator of gene expression ranging from X-chromosome inactivation and genomic imprinting to suppression of repetitive elements (Breiling and Lyko, 2015). However, the distribution and function of m\textsuperscript{5}C in RNA is relatively unclear. Fueled by recent advances in sequencing techniques, m\textsuperscript{5}C was mapped transcriptome-wide on tRNA, rRNA, mRNA, and non-coding RNA in bacteria, archaea, S. cerevisiae and human cell lines.
(Edelheit et al., 2013; Hussain et al., 2013; Khoddami and Cairns, 2013; Squires et al., 2012).

In *S. cerevisiae*, m$^5$C is present in various locations at the anticodon loop, the variable region, and the T loop in many cytosolic tRNAs (Table 1.3), where it functions to stabilize the secondary structure of tRNA and may respond to environmental changes (El Yacoubi et al., 2012). For example, the presence of m$^5$C40 in tRNA-Phe-GAA enhances Mg$^{2+}$ binding to tRNA and promotes stabilization of tRNA structure (Chen et al., 1993). The m$^5$C level in tRNA-His increases in response to several growth arrest conditions, such as growth at nonpermissive temperature, starvation of required amino acids, glucose, or uracil, and rapamycin treatment (Preston et al., 2013).

The formation of all m$^5$C in *S. cerevisiae* requires Trm4 methyltransferase (Motorin and Grosjean, 1999), but studies of Trm4 or m$^5$C in *S. cerevisiae* are limited. It is unclear if m$^5$C at position 34 contributes to decoding efficiency or if m$^5$C at other positions plays different roles. In contrast, relatively extensive evidence in higher eukaryotes showed that m$^5$C methyltransferases such as Nsun2 and Dnmt2 are implicated in several activities, including stress response, protection of tRNAs from nuclease cleavage, processing ncRNA into regulatory small RNAs, and neurological abnormalities (intellectual disability, microcephaly, and behavioral deficits among them; Blanco et al., 2014; Chan et al., 2012; Hussain et al., 2013; Schaefer et al., 2010).
Table 1.3. \textit{m}^5\textit{C} locations in \textit{S. cerevisiae} tRNAs

<table>
<thead>
<tr>
<th>Position</th>
<th>tRNA species</th>
</tr>
</thead>
<tbody>
<tr>
<td>27</td>
<td>Pro-UGG</td>
</tr>
<tr>
<td>34</td>
<td>Leu-CAA</td>
</tr>
<tr>
<td>40</td>
<td>Phe-GAA</td>
</tr>
<tr>
<td>48</td>
<td>Asn-GUU, Cys-GCA, Ile-AAU/UAU, Ini-CAU, Leu-UAA/CAA, Lys-UUU, Ser-UGA/IGA/CGA/GCU, Thr-UGU, and Tyr-GUA</td>
</tr>
<tr>
<td>49</td>
<td>Arg-ICG, Asp-GUC, Glu-UUC, Gly-GCC, His-GUG, Ini-CAU, Pro-UGG, Val-AAC/UAC, and Phe-GAA</td>
</tr>
<tr>
<td>50</td>
<td>Gly-GCC</td>
</tr>
<tr>
<td>56</td>
<td>Leu-UAA</td>
</tr>
</tbody>
</table>

**N6-threonylcarbamoyladenosine (t$^6$A) at position 37**

N6-threonylcarbamoyladenosine (t$^6$A) is present at position 37 in tRNAs that decodes ANN codons. A structural study demonstrated that t$^6$A37 strengthens the A1-U36 codon-anticodon interaction in the ribosome (Murphy et al., 2004). This modification forms a planar structure that is coplanar with the adenine ring, and expands the area of t$^6$A37 available to stack with A38 of the tRNA and A1 of the codon (Figure 1.4). The bulky and freely rotatable threonyl moiety of t$^6$A37 forbids it from being incorporated into the helical stack (prevent the formation of U33-A37 base pair of the anticodon loop), and thus restricts the position of t$^6$A37 and stabilize the structure of the anticodon loop. Moreover, t$^6$A37 plays roles in frameshifting maintenance in vitro and initial codon selection in vivo (El Yacoubi et al., 2011).
There are six conserved proteins involved in the biosynthesis of t^6A: a universal protein (Sua5), a kinase-associated endopeptidase (Kae1), a protein kinase (Bud32), and three small polypeptides (Cgi121, Pcc1, and Gon7) (Daugeron et al., 2011; El Yacoubi et al., 2011; El Yacoubi et al., 2009; Srinivasan et al., 2011). Deletion of SUA5, KAE1, BUD32, or GON7 results in a complete absence of t^6A in S. cerevisiae, while deletion of CGI121 or PCC1 leads to a ~20% reduction in the level of t^6A compared to wildtype (El Yacoubi et al., 2011; Srinivasan et al., 2011; Thiaville et al., 2016). While Sua5 alone was identified as a t^6A37 modifying enzyme (El Yacoubi et al., 2009), the other five proteins were first identified as the KEOPS complex (Kinase, Endopeptidase and Other Proteins of Small size) responsible for telomere maintenance (Downey et al., 2006), and as the EKC complex (Endopeptidase-like and Kinase associated to transcribed Chromatin) involved in transcription regulation (Kisseleva-Romanova et al., 2006). It remains elusive if these functions of the KEOPS/EKC complex are separate from t^6A biosynthesis, or if the misregulation of telomere and transcription in these mutants is dependent on t^6A deficiency.

Loss of t^6A37 shares similar phenotypes with xm^5^U34 deficiency. Disruption of the KEOPS/EKC complex causes a severe growth phenotype, as well as translational upregulation of transcription factor Gcn4, which is normally expressed in a low level and is regulated at the translational level by four upstream open reading frames (Daugeron et al., 2011). Consequently, these mutants exhibit upregulation of Gcn4 target genes at the RNA level. Mutations in
the KEOPS complex are associated with neurological disorders such as early-onset nephrotic syndrome, microcephaly, brain anomalies, and developmental delay (Braun et al., 2017; Edvardson et al., 2017). The above phenotypes were also observed in xmsU34 deficient mutants (Nedialkova and Leidel, 2015; Zinshteyn and Gilbert, 2013).

**Figure 1.4. Structure of stacking interactions of t^6A37**

Structure of the decoding site for anticodon loop of tRNA-Lys-UUU-t^6A37 and codon AAA. 16S rRNA is grey, with interacting bases indicated. S12 is orange, codon is purple, and anticodon loop is cream. The base of t^6A37 is orange, and the modification is pink. The t^6A37 base forms stacks with tRNA A38 and the first base of the codon (A1). Picture and legend adapted from (Murphy et al., 2004)
N6-isopentenyladenosine (\textit{i}^6\textit{A}) at position 37

N6-isopentenyladenosine (\textit{i}^6\textit{A}) is another modification at position 37 and present in tRNAs harboring A36-A37-A38 (tRNA-Cys-GCA, Ser-NGA, and Tyr-GUA in \textit{S. cerevisiae}). This modification is linked with ribosome binding affinity, altered gene expression, and translational efficiency and fidelity (Schweizer et al., 2017). The formation of both cytosolic and mitochondrial \textit{i}^6\textit{A} requires Mod5, a delta 2-isopentenyl pyrophosphate:tRNA isopentenyl transferase in budding yeast (Dihanich et al., 1987; Najarian et al., 1987). In addition to its function in tRNA modification, Mod5 has been known to contribute to the tRNA gene-mediated silencing of genes that are near tRNA genes by binding to tRNA genes and pre-tRNAs (Pratt-Hyatt et al., 2013), and was identified as a prion protein that promotes drug resistance and cell survival upon environmental stress via prion conversion (Suzuki et al., 2012).

In humans, the isopentenyl transferase homolog TRIT1 has been implicated in diseases. For example, a mutation in TRIT1 is associated with severe combined mitochondrial respiratory chain defects (Yarham et al., 2014). TRIT1 is also a possible tumor suppressor, as its RNA level was downregulated in lung adenocarcinomas, and as overexpression of a functional TRIT1 gene in lung cancer cells alleviated their colony size and reduced the probability of tumor incidence in nude mice injected with the transfected cells (Spinola et al., 2005).
**Wybutosine (yW) at position 37**

Wybutosine (yW) is an extensively modified nucleotide at position 37 in tRNA-Phe. The biosynthesis pathway of yW is a multistep process involving five enzymes: Trm5, a S-adenosylmethionine (SAM) dependent tRNA methylase, and Tyw1-4, tRNA-yW synthesizing enzymes 1-4 (refer to Figure 1.5 for detailed synthetic process; Kalhor et al., 2005; Noma et al., 2006). Loss of any one of these enzymes leads to a complete absence of yW. Similar to the function of t⁶A37, yW37 has been known to stabilize codon-anticodon interaction. The stacking interactions of the tricyclic yW with adjacent bases A36 and A38 decrease the entropic penalty for ribosomal A-site binding (Konevega, 2004).

The human homolog TYW2 has been implicated in breast cancer, as it was amplified and overexpressed in several breast cancer cell lines. However, the level of yW modification in tRNA-Phe was not affected in a mouse tumor model with overexpressed TYW2 (Rodriguez et al., 2012), suggesting human TYW2 may have a role in tumorigenesis unrelated to yW biogenesis.
Figure 1.5. Biosynthesis pathway of yW in S. cerevisiae
Trm5 uses Ado-Met as a methyl donor to methylate G at position 37 of tRNA-Phe in order to generate 1-methylguanosine (m\(^1\)G). Tyw1 catalyzes tricyclic formation on m\(^1\)G to produce 4-demethylwyosine (imG-14), using Ado-Met and FMN as cofactors. Tyw2 transfers \(\alpha\)-amino-\(\alpha\)-carboxypropyl group from Ado-Met to the side chain of imG-14 to yield 7-aminocarboxypropyl-demethylwyosine (yW-86). Tyw3 methylates N4 position of yW-86 to generate 7-aminocarboxypropyl-wyosine (yW-72). Tyw4 catalyzes methylation and methoxycarbonylation in the final step of yW biosynthesis to complete yW formation. Picture and legend adapted from (Noma et al., 2006).

**N4-acetylcytidine (ac\(^4\)C) at position 12**
N4-acetylcytidine (ac\(^4\)C) is present in tRNAs from all three domains. In bacteria and archaea, ac\(^4\)C is found at position 34 and responsible for preventing codon misreading during translation. In eukaryotes, this modification occurs only at position 12 in a subset of tRNAs, e.g., tRNA-Ser and tRNA-Leu in S. cerevisiae (Juhling et al., 2009). ac\(^4\)C12 is also found in intron-containing precursor tRNA in yeast, implicating involvement in an early step of the tRNA maturation (Etcheverry et al., 1979). The biosynthesis of ac\(^4\)C12 requires Tan1 (tRNA acetylation), and ac\(^4\)C12 deficiency in tan1\(\Delta\) was linked with a decreased level of mature tRNA-Ser-CGA, indicating the role of this modification in tRNA stability (Johansson and Bystrom, 2004).
Part 2: Ribosome Profiling

Translation is an essential cellular process that decodes genetic information of mRNA into functional proteins. This energy-consuming process is under tight regulation. Any subtle defect during translation may result in severe dysfunction of proteins and impact on human health. Although the mechanism of translation has been extensively studied, and much is known about the function and structure of the major players in translation—the ribosome and tRNA—our understanding of translational regulation has largely been limited.

Measuring mRNA abundance using techniques to analyze nucleic acids, from the beginning of microarrays to the present high-throughput sequencing, has been the major focus to monitor global gene expression, despite the fact that translation is a fundamental step in determining levels of protein expression. Although the steady-state protein level can be quantitatively measured by proteomic mass spectrometry (Aebersold and Mann, 2003), this technique has not been implemented to a user-friendly workflow and the readout only provides partial information of the translation process. Recently, the emergence of ribosome profiling (Ingolia et al., 2009), a sequencing-based approach to monitor \textit{in vivo} translation, fills the technical gap between RNA sequencing and proteomic mass spectrometry, offering more details on how a given mRNA is translated by the ribosomes.
Ribosome profiling measures the mRNA fragments (~30 nucleotides) that are protected by the ribosomes from nuclease digestion. The sequencing reads infer precise positions of the ribosomes on a given mRNA at the moment when the translation was halted, providing a high-resolution and quantitative profile of translation across the transcriptome. The ribosome density on a given transcript thus can be used to estimate translation rate and efficiency. This technique provides many aspects of insights into the mechanism of translational regulation, including identifying regulatory ribosome pausing sites and translated upstream open reading frames (uORFs) (Ingolia et al., 2009; Ingolia et al., 2011; Nedialkova and Leidel, 2015; Smith et al., 2014; Zinshteyn and Gilbert, 2013). It also unveiled the translational control in certain organelles such as mitochondria (Rooijers et al., 2013; Williams et al., 2014). Recently, a number of studies have employed ribosome profiling to illuminate the roles for specific tRNA modifying enzymes (e.g., Elp-related proteins and KEOPS complex) in translational regulation (Laguesse et al., 2015; Nedialkova and Leidel, 2015; Thiaville et al., 2016; Zinshteyn and Gilbert, 2013). Nonetheless, the majority of modifying enzymes are still omitted from the genome-wide analysis on translation and their roles in translational control are await to be investigated using this powerful technique.

There are seven major steps in ribosome profiling: cell lysis, footprinting using nuclease digestion, ribosome purification, footprint RNA purification, library generation, deep sequencing, and data analysis (Figure 1.6A; Ingolia, 2016).
Manipulation of each step may affect remarkably the final interpretation of the translation profile. Translation and ribosome occupancy can change in seconds upon stress or environmental changes. As a result, the first step of cell lysis plays a critical role in capturing the snapshot of translation in cells. In yeast, this is tricky because yeast has a cell wall and cannot be directly lysed in detergent like mammalian cells. Yeast cell lysis is commonly achieved by two approaches: 1) fast centrifugation followed by detergent lysis with bead beating (Heyer and Moore, 2016; Zinshteyn and Gilbert, 2013); or 2) rapid filtration followed by cryogenic grinding (Ingolia et al., 2009; Nedialkova and Leidel, 2015). Although the second method appears to better capture the physiological state of translation, both methods have been used to generate reliable and reproducible datasets.

The cell culture is treated with elongation inhibitors such as cycloheximide (CHX) shortly prior to lysis in order to stabilize the ribosomes on mRNA and preserve translation at the moment of cell lysis. This treatment may, however, result in several artifacts. Translation initiation and termination continuously occur after CHX treatment since it only inhibits elongation. It causes an extensive accumulation of ribosomes at the start codon and a depletion at the stop codon, hindering the studies that require the information of ribosome profile around the 5' and 3' of the transcripts (Ingolia et al., 2011). In addition, a reversible inhibitor like CHX appears to allow slow and concentration-dependent elongation until cell lysis (Gerashchenko and Gladyshev, 2014; Hussmann et al., 2015). CHX
treatment seems to lock the ribosomes in a certain conformation resulting in ~30-nucleotide footprints, while the sequencing reads without CHX consist of an additional size of footprints at ~22 nucleotides (Lareau et al., 2014). Studies using or not using CHX revealed opposite conclusions on the correlation of translation rate and codon optimality (Hussmann et al., 2015). Fortunately, the drug effects seem to exclusively affect the codon-level analysis but not the gene-level analysis (the average ribosome density on a given transcript)—CHX treatment simply distorts the distribution of ribosomes on mRNA, but does not add or remove ribosomes on the ORFs. Nevertheless, tRNA modification studies that employed ribosome profiling with CHX treatment uncovered the roles of tRNA modifying enzymes in translation rate and in the regulation of gene expression (Laguesse et al., 2015; Nedialkova and Leidel, 2015; Thiaville et al., 2016; Zinshteyn and Gilbert, 2013). Nedialkova et al. also showed that using or omitting CHX resulted in consistent findings. Taken together, CHX treatment might affect the outcomes of certain experiments, while it displays a minor effect on other studies.

The methods to construct a deep-sequencing library may somewhat influence the outcome of sequencing reads. Library generation for small RNA sequencing is commonly used in ribosome profiling because of the small size (~30 nt) of the footprints. Typically, it includes ligation with a preadenylylated oligonucleotide at the 3’ of the footprint fragment, reverse transcription, and cDNA circularization prior to amplification across the footprints (Heyer et al.,
Recent optimized protocols have added unique molecular identifiers (UMIs) at the 5’ of the preadenylylated oligonucleotide and at the beginning of the primer for reverse transcription to ameliorate the sequence biases in ligation and circularization, as well as to allow the computational removal of duplicate reads during PCR amplification (Lecanda et al., 2016; McGlincy and Ingolia, 2017).

Data analysis of ribosome profiling includes mainly two parts: gene-level and codon-level analyses. For gene-level analysis, the relative translation efficiency of a given transcript can be inferred by the quantification of the ribosome footprints from that transcript. The quantity of ribosome footprints observed on a given transcript is assumed to positively correlate with the outcome of protein synthesis, based on the premise that all the ribosomes complete their translation. Simply calculating the ribosome footprints, however, would not truly reflect translation efficiency, because an increase in ribosome footprints could imply two scenarios: increased translation or increased mRNA abundance (Figure 1.6B). Hence, an mRNA sequencing library that is generated side-by-side with the ribosome footprinting is required for the data interpretation. The ribosome density normalized to the mRNA abundance, termed translational efficiency (TE), is generally used to compare the translation efficiency between samples.

Codon-level analysis calculates specific ribosomal A, P, or E site occupancy on each codon transcriptome-wide (Ingolia et al., 2009; Nedialkova
This analysis reveals ribosome pausing or stalling on specific codons across all transcripts. With a visualized genome browser track, it is also able to uncover the position in which the ribosome pauses (Brar and Weissman, 2015; Ingolia, 2016). Codon-level analysis is carried out as: 1) assignment of ribosomal A, P, E sites in sequencing reads, and 2) quantification of A, P, or E site occupancy on each codon. The assignment of ribosomal P site is based on the distance from the 5' end to the AUG codon (named p-site offset) of the accumulated reads aligned at the start codon (Figure 1.6C). Since the p-site offset varies with the size of ribosome footprints, each length of footprints (commonly 28-32 nt) must be calculated separately and may have distinct values of the offset. Approaches for quality control and normalization can remarkably improve the signals of codon-level analysis (Nedialkova and Leidel, 2015). The removal of the first 15 and the last 5 codons from the reference ORFs reduces the noise caused by artefactual ribosome accumulation at and downstream of the start codon and ribosome runoff at the stop codon. Normalization of the codon frequency of A, P, or E sites by the basal codon frequency of 3 downstream sites of the A site (Figure 1.6D) also improves the accuracy and consistency of ribosome occupancy data between samples.
Figure 1.6. Ribosome profiling

(A) Workflow of ribosome footprint profiling.

(B) Ribosome footprint density encompasses mRNA abundance and translation. More ribosome footprints can result from higher mRNA abundance or increased translation.

(C) P-site offset assignment. Cumulative coverage of 5’ nucleotides from ribosome footprint reads mapping near start codons across all transcripts is shown. The length of mapped reads is between 29 and 31 nucleotides (nt). The peak located 12-13 nt upstream of the start codon is inferred to represent ribosomes for translation initiation that contain the AUG codon in their P-site. The P-site offsets for the mapped reads are 12 or 13 based on the length of the reads.

(D) Approach for determining ribosomal A, P, E sites and the downstream 3 sites within ribosome footprints inferred from (C).

(A and B) Pictures and legend adapted from (Ingolia, 2016); (C and D) pictures and legend adapted from (Nedialkova and Leidel, 2015).
CHAPTER II: RESULTS

Abstract

Covalent nucleotide modifications in noncoding RNAs affect a plethora of biological processes, and new functions continue to be discovered even for well-known modifying enzymes. To systematically compare the functions of a large set of noncoding RNA modifications in gene regulation, we carried out ribosome profiling in budding yeast to characterize 57 nonessential genes involved in tRNA modification. Deletion mutants exhibited a range of translational phenotypes, with enzymes known to modify anticodons, or non-tRNA substrates such as rRNA, exhibiting the most dramatic translational perturbations. Our data build on prior reports documenting translational upregulation of the nutrient-responsive transcription factor Gcn4 in response to numerous tRNA perturbations, and identify many additional translationally regulated mRNAs throughout the yeast transcriptome. Our data also uncover unexpected roles for tRNA-modifying enzymes in regulation of TY retroelements, and in rRNA 2′-O-methylation. This dataset should provide a rich resource for discovery of additional links between tRNA modifications and gene regulation.
Introduction

In addition to the four common nucleotides present at roughly equal abundance in RNA—A, G, C, and U—it has long been known that covalently modified nucleotides are present at lower abundance. A classic example of such a modified nucleotide is the 7-methylguanylate cap found at the 5′ end of eukaryotic mRNAs. Although decades of investigation have identified scores of modified nucleotides in a variety of coding and noncoding RNAs, new nucleotide modifications continue to be discovered. Moreover, the functions of many nucleotide modifications remain obscure: while the chemical and structural properties of specific modified nucleotides are often well understood, the detailed functional and regulatory consequences remain unknown for many RNA modification events in vivo.

Nucleotide modifications are particularly common in tRNAs, and it is estimated that ~15% of all tRNA nucleotides are covalently modified (Czerwoniec et al., 2009; El Yacoubi et al., 2012; Phizicky and Hopper, 2015). Modified tRNA nucleotides include relatively common species such as 5-methylcytosine (m5C) and pseudouridine (Ψ), as well as unusual and complex nucleotides such as the wobble modification 5-methoxycarbonylmethyl-2-thiouridine (mcm5s2U), which is generated in a multi-step process by the Elongator complex along with a number of additional factors (Esberg et al., 2006; Huang et al., 2005; Huang et al., 2008). At the molecular level, tRNA modifications have been implicated in a wide variety of processes, including stabilization of tRNA secondary structure (Motorin and
Helm, 2010), translation initiation (Liu et al., 2016), decoding (Li et al., 1997; Nedialkova and Leidel, 2015; Zinshteyn and Gilbert, 2013), reading frame maintenance (Lecointe et al., 2002), protection of tRNAs from nuclease cleavage/degradation (Alexandrov et al., 2006; Schaefer et al., 2010), and subcellular trafficking of tRNAs (Kramer and Hopper, 2013). At the organismal level, tRNA-modifying enzymes have been implicated in processes ranging from neurodevelopment to meiotic chromosome pairing to abscisic acid signaling to early development (Phizicky and Hopper, 2010). Interestingly, some tRNA modifications can be regulated in response to environmental conditions, as, for example, thiolation of specific tRNAs in *S. cerevisiae* is responsive to the levels of sulfur-containing amino acids cysteine and methionine in the growth media, with altered tRNA thiolation affecting the translation of relevant biosynthetic proteins (Laxman et al., 2013). Although these and many other examples of tRNA modification biology have been uncovered over decades of study, the functions of many tRNA modifications remain mysterious.

A handful of recent studies have carried out transcriptome-wide analysis of translation, using ribosome profiling (Ingolia et al., 2009) to illuminate the roles for specific tRNA-modifying enzymes in translation and proteostasis (Laguesse et al., 2015; Nedialkova and Leidel, 2015; Thiaville et al., 2016; Tuorto et al., 2015; Zinshteyn and Gilbert, 2013). Here, we adopt this approach to systematically study the roles for tRNA modifications in translational regulation in *S. cerevisiae*, using ribosome profiling to generate ribosome occupancy maps for 57 yeast
deletion strains. As expected, deletion of genes encoding enzymes that modify nucleotides in the tRNA anticodon caused the most dramatic translational phenotypes, while loss of enzymes responsible for more distant tRNA modifications often resulted in few discernible translational phenotypes.

Codon-level analysis in many cases recovered the expected stalling or slowing of ribosomes at codons corresponding to relevant modified anticodons, as well as identifying codon-level translational perturbations in a number of previously unstudied mutants. Scrutiny of the dramatic translational phenotypes observed in *trm7Δ* cells revealed a potential role for this gene in methylation of rRNA as well as tRNAs. At the level of individual transcripts, mutant effects on translation of specific genes resulted in a variety of downstream outcomes, both expected (altered transcription of amino acid metabolism genes secondary to altered translation of *GCN4* mRNA) and surprising (impaired heterochromatin-mediated gene silencing). Most surprisingly, we find a role for the Elongator complex (and other factors involved in generation of mcm5*s*2U) in maintaining expression of transcripts associated with *TY1* retrotransposon long terminal repeats (LTRs).

Overall, our data illuminate unanticipated aspects of Trm7 and Elongator function, reveal many additional examples of translational regulation by upstream open reading frames (uORFs), and provide a rich source of hypotheses for future study.
Results

**Ribosome Footprinting in Mutant Yeast Strains**

Budding yeast encode 73 genes currently annotated to play a role in tRNA modification, of which 14 are essential. Here, we set out to characterize translational phenotypes for the remaining 59 nonessential genes. Haploid deletion mutants for 57 genes (two mutants—\emph{pcc1Δ} and \emph{pus6Δ}—failed quality control several times) were freshly obtained by sporulation of heterozygous mutant diploids to confirm the viability of the deletion in question and to minimize the potential for suppressor mutations.

Initial studies revealed aneuploidies in a subset of the mutants, the majority of which we subsequently re-derived and confirmed to be euploid. However, for two mutants—\emph{bud32Δ} and \emph{gon7Δ}, both members of the conserved KEOPS complex (Daugeron et al., 2011; Downey et al., 2006; Kisseleva-Romanova et al., 2006; Srinivasan et al., 2011)—we repeatedly obtained haploid strains bearing an additional copy of ChrIX, suggesting that the KEOPS complex may be essential in our strain background. We included these mutants in the final set of 57 mutants despite this aneuploidy, as many of the dramatic ribosome footprinting phenotypes observed in these two mutants were also observed to a lesser extent in the euploid \emph{cgi121Δ} mutant (which exhibits partial but incomplete loss of the t⁶A modification; (Thiaville et al., 2016)), suggesting that many of the observed phenotypes accurately reflect KEOPS function. Nonetheless, we urge
caution in interpreting results obtained from bud32Δ and gon7Δ cells, as observed phenotypes may be secondary to second-site mutations.

Figure 2.1A shows these 57 genes, grouped according to their modified nucleotide product. This list includes not only the catalytic subunits of tRNA-modifying enzymes, but also other factors that affect a given tRNA modification, such as the phosphatase Sit4, which is required for Elongator function in vivo. In addition, it is important to note that many of the encoded proteins are also known to affect nucleotide modifications on other RNA species such as mRNAs or rRNA, or play more pleiotropic roles in cell biology: the multimethylase-activating scaffold protein Trm112, for example, is required for methylation of tRNAs, rRNA, and elongation factors (Liger et al., 2011). When appropriate, we will discuss the potential for non-tRNA targets as the relevant mechanistic basis for translational changes observed below.

For these 57 deletion strains, we assayed the consequences of loss of tRNA modification on translational control proteome-wide, using ribosome profiling (Ingolia et al., 2009) to provide codon-resolution insight into ribosome occupancy (Figure 2.1B). Matching mRNA abundance data were gathered for each strain as a reference for the ribosome footprinting dataset to enable calculation of translational efficiency per transcript, and to identify any mutant effects on transcription or mRNA stability. Wild-type (WT) and mutant strains were grown to mid-log phase in rich media and processed for ribosome profiling, with biological duplicates for each strain. As the use of the translation inhibitor
cycloheximide can affect the distribution of ribosomes across ORFs (Gerashchenko and Gladyshev, 2014; Hussmann et al., 2015), we also generated replicate datasets for six strains collected without the use of cycloheximide. Tables 2.1, 2.2, 2.3, and 2.4 provide complete datasets for mRNA, ribosome occupancy, and translational efficiency, as well as the six no-cycloheximide replicates.

Overall, biological replicates were well correlated with one another (Figures 2.1B and 2.2), with pairwise correlations ranging from 0.97 to >0.99. Moreover, the six datasets generated from cultures not subject to cycloheximide treatment exhibited similar pairwise correlations between experiments run with or without cycloheximide (Figure 2.2) as observed for replicate pairs run in the presence of cycloheximide. Ribosome-protected footprints (RPFs) were predominantly 28–32 nt (Table 2.5), as expected, and exhibited known features such as 3 nt periodicity over coding regions and absence of ribosomes over introns. In addition, our data recapitulated prior ribosome profiling analysis of Elongator and other mutants involved in the formation of mcm5’s2U and related modifications (Nedialkova and Leidel, 2015; Zinshteyn and Gilbert, 2013); see below). This dataset thus provides a high-quality resource for analysis of the roles for tRNA-modifying enzymes in translation.

Below, we analyze the dataset at three levels of granularity: averaged across codons, averaged over metagenes, and averaged across individual genes.
Figure 2.1
**Figure 2.1. Overview of dataset**

(A) Nonessential genes involved in tRNA modifications in budding yeast. Encoded proteins are grouped roughly according to function, as, for example, the Elongator complex is grouped along with other enzymes required for formation of the mcm$^5$s$^2$U wobble modification. For each group of enzymes, the known product is shown (R indicates ribose in the tRNA backbone for modified bases), along with a tRNA cartoon showing the best-characterized modification locations. For some sets of mutants, the modification shown represents only a subset of products, as, for example, Elongator and associated factors also catalyze the formation of mcm$^5$U, ncm$^5$U, and ncm$^5$Um, in addition to mcm$^5$s$^2$U as shown. Throughout the manuscript, modifying enzymes are generally color-coded as indicated here, except in cases where subsets of related factors must be distinguished.

(B) Example of RNA-seq and ribosome footprinting data for chr3:57,000–107,000, showing strong correlations between biological replicate experiments, and also, for the majority of the transcriptome, between mutant strains.
**Figure 2.2. Example of dataset reproducibility**

(A) Data for the *GCN4* locus is shown for the indicated strains, with biological replicates, as well as replicates of distinct but functionally-related mutants, exhibiting highly similar RNA and RPF profiles.

(B) Correlation matrix for all ribosome occupancy replicates. Correlations were computed for log10(rpkm+1) to mitigate the correlation driven by expression levels. In general, replicates clustered together, with similarly high correlations observed for datasets for functionally-related mutants (such as *elp1Δ* and *uba4Δ*).

(C) Biological replicates were significantly better-correlated than other pairs of datasets. Histograms show percent of pairwise comparisons exhibiting the indicated correlation, for all biological replicate pairs and for all remaining pairs, as indicated.
Figure 2.3. Minor effects of cycloheximide on global ribosome occupancy

(A) Scatterplots showing ribosome occupancy for all annotated ORFs, comparing replicates generated using cycloheximide to arrest translation before cell harvest and during cell lysis (X axis, rpkm+1, log10 scale), and replicates generated without cycloheximide treatment prior to cell lysis (Y axis, rpkm+1, log10 scale). With the exception of a few dubious ORFs (indicated in the tan1Δ scatterplot, but which appear in all samples), cycloheximide had negligible effects on overall ribosome profiling data at the level of gene-specific ribosome occupancy –
correlations between replicate pairs run in the presence of cycloheximide for these six strains were all 0.97-0.99 (indicated on the scatterplots), nearly identical to the correlations between no cycloheximide and cycloheximide replicates.

(B) Length of ribosome-protected fragments for all datasets. Left two panels show the highly reproducible footprint size distributions observed in two major batches of samples – Batch 1, comprising 28 samples, was collected throughout 2015, while Batch 2, comprising 98 samples, was collected throughout 2016. We do not understand the reason for the subtle difference between these batches (30-32 nt for Batch 1 and 29-31 nt for Batch 2) – we note that in the 2015 batch, yeast was incubated on ice for 20 sec prior to collection by centrifugation, while this step was omitted in the 2016 batch, but otherwise samples were processed identically. Whatever the reason for this effect, we note that all related mutants that were split between batches (e.g., Batch 1 included $\text{elp1} \Delta$, $\text{elp3} \Delta$, $\text{kti11} \Delta$, and $\text{urm1} \Delta$ from the Elongator-related mutants) nonetheless exhibited extraordinary reproducibility at every scale of analysis from codon-level (where fragment length is explicitly accounted for) to gene-level.

In contrast to the primary dataset collected in the presence of cycloheximide, the length distribution of ribosome-protected fragments was highly variable in the no cycloheximide samples, ranging from 28-30 nt for $\text{trm7} \Delta$ to 30-33 nt for $\text{tan1} \Delta$ and $\text{tyw4} \Delta$, in contrast to the nearly-uniform footprint size distributions observed for cycloheximide-treated samples. Previously reported 20-22 nt footprints in no-CHX samples (Lareau et al., 2014) were excluded from this study and thus were not shown here. This variability in fragment size precluded us from providing a detailed analysis of codon-level ribosome occupancy, which is the likeliest characteristic of ribosome occupancy profiles to be affected by cycloheximide treatment. We therefore conclude that codon-level analyses may be influenced by the use of cycloheximide in the primary dataset in this manuscript, but that all analyses at coarser resolution – gene-level ribosome
occupancy measures, etc. – are robust to the use of cycloheximide in the ribosome profiling protocol.

Table 2.1. RNA-Seq dataset
Sequencing depth-normalized RNA-Seq data for mutant and wild-type (columns), with rows showing the 4884 genes with at least 10 or more reads in all wild-type replicate RPF datasets. Link to dataset:
https://www.ncbi.nlm.nih.gov/geo/download/?acc=GSE100626&format=file&file=GSE100626%5FChou%5FTable%5FS1%2Etxt%2Egz

Table 2.2. Ribosome footprinting dataset
As in Table 2.1, but for ribosome footprinting data. Link to dataset:
https://www.ncbi.nlm.nih.gov/geo/download/?acc=GSE100626&format=file&file=GSE100626%5FChou%5FTable%5FS2%2Etxt%2Egz

Table 2.3. Translational efficiency
Mutant effects on translational efficiency, relative to wild-type, expressed as log2 fold change. Mutants effects on translational efficiency are calculated as RPF rRPKM/RNA rRPKM, with rRPKM calculated by normalizing RPKM in each mutant by the average of RPKM in WTs grown in the same batch. Link to dataset:
https://www.ncbi.nlm.nih.gov/geo/download/?acc=GSE100626&format=file&file=GSE100626%5FChou%5FTable%5FS3%2Etxt%2Egz

Table 2.4. No cycloheximide ribosome footprinting dataset
As in Table 2.2, for six replicate ribosome profiling experiments carried out without cycloheximide treatment. Link to dataset:
https://www.ncbi.nlm.nih.gov/geo/download/?acc=GSE100626&format=file&file=GSE100626%5FChou%5FTable%5FS6%2Ettx%2Egz

Table 2.5. Ribosome-protected fragment lengths
Fraction of each fragment size in 27-34 nt for all ribosome footprinting datasets after sequence demultiplexing, 5' CC and 3' adaptor trimming, and rRNA removal. Link to dataset: http://www.cell.com/molecular-cell/fulltext/S1097-2765(17)30837-7. Refer to Table S5 in the link.
Codon-Level Analysis of Ribosome Occupancy Changes

Changes in tRNA levels or modifications can affect the dwell time of ribosomes on the relevant codon occupying the A site, and this is readily observed as changes in codon-averaged ribosome occupancy. We therefore analyzed global codon occupancy for all 57 mutants, as previously described (Nedialkova and Leidel, 2015) (Figures 2.4 and Table 2.6). Our data recapitulate recent studies of mutations affecting mcm$^5$s$^2$U formation (henceforth collectively referred to as Elongator-related mutants) that documented increased A site ribosome occupancy over AAA, CAA, and GAA codons, which are decoded using mcm$^5$s$^2$U34-containing tRNAs (Nedialkova and Leidel, 2015; Zinshteyn and Gilbert, 2013), providing further confidence in our dataset (Figures 2.4 and 2.5A).

For many modifiers with known tRNA substrates, we document altered ribosomal A site occupancy at the relevant codons (Figures 2.5 and 2.6), consistent with the tRNA modification in question affecting tRNA stability, charging, or codon recognition. For example, loss of Mod5, which generates N6-isopentenyladenosine (i$^6$A) at position 37 in a number of tRNAs (Cys-GCA, Ser-NGA, and Tyr-GUA) (Laten et al., 1985), results in dramatically decreased A site ribosome occupancy over all relevant codons (Figure 2.5B). In the case of Tan1, responsible for generation of N4-acetylcytidine (ac$^4$C) at position 12 of leucine and serine tRNAs (Johansson and Bystrom, 2004), deletion mutants exhibit dramatically decreased A site ribosome occupancy, but a corresponding increase in P site occupancy over the relevant codons (Figure 2.5C). The observation
suggests that this modification could potentially play some role(s) in peptidyl-tRNA positioning, enhancing translocation of codons from the P to the E site, or delaying translocation from the A site to the P site. Curiously, loss of the dehydroyuridine synthase Dus2, which is responsible for dU20 formation in the majority of tRNAs (Xing et al., 2004), caused a similar (albeit less dramatic) reduction in A site occupancy of serine and leucine codons (Figure 2.4), although it did not cause the same compensatory increase in P site occupancy.

In addition to relatively specific changes occurring over known target codons, we observed more widespread changes in A site occupancy in mutants lacking Trm1 or Trm112 (Figure 2.4B), consistent with the broad substrate range for these methylases: Trm1 generates N2,N2-dimethylguanosine (m2G26) in the majority of cytoplasmic tRNAs (Ellis et al., 1986), while the Trm112 methylase scaffold is required for appropriate methylation of tRNAs, rRNAs, and translation factors (Liger et al., 2011). Interestingly, in both of these mutants, A site occupancy tends to increase over codons beginning with purines and decrease over codons beginning with pyrimidines (Figure 2.4A).

In addition to these and other cases that confirm and extend expected aspects of tRNA modification, we uncovered a number of surprising changes in ribosome occupancy that suggest additional roles for tRNA modifications in translation (Figures 2.6). For example, in several analyses (here and below) we observed discrepant phenotypes for the four tyw mutants, despite the fact that loss of any of the four encoded enzymes completely eliminates wybutosine (yW
modification at position 37 in tRNA-Phe) synthesis in vivo (Noma et al., 2006). Here, we noted that all four deletion mutants exhibited increased ribosomal P site occupancy over relevant codons, but that A site occupancy at these codons was increased only in mutants lacking Tyw1–3 (Figure 2.6B). As the yW precursor “yW-72” (lacking a methyl group on the α-carboxyl group and a methoxycarbonyl group on the α-amino group of yW) accumulates in target tRNAs in tyw4Δ mutants (Noma et al., 2006), we speculate that yW-72 may be sufficient for appropriate decoding of tRNA-Phe, but that the full yW modification is required for efficient peptidyl transfer and/or translocation. As another example, although we observe significantly increased P site occupancy over valine codons GUC/GUG/GUU in mutants affecting the Trm8/Trm82 heterodimer (Alexandrov et al., 2006), these mutants also exhibit unexpected decreases in P site occupancy of UGC and UGU, which are decoded by tRNA-Cys-GCA (Figure 2.6C).

Taken together, these analyses recapitulate previously reported translational deficits and thereby validate the quality of our dataset, as well as illuminating additional functions or targets of various tRNA modifications, providing hypotheses for mechanistic follow-up.
Figure 2.4. Over view of codon-level analysis of ribosome occupancy
(A) Effects of all 57 mutations on average ribosome occupancy at A, P, and E sites over all 61 codons (excluding stop codons). Columns depict mutations, with key mutations identified above clusters. Heatmaps show log2 fold changes
relative to the WT average (red, increased codon occupancy; green, decreased codon occupancy). Data for A, P, and E sites are all sorted identically, based on A site dataset clustering. Note that P and E site panels are scaled to 50% of the width of the A site panel.

(B) Expanded view of A site occupancy with all 57 mutants annotated.

Figure 2.5. Mutant effects on codon occupancy

(A) A site ribosome occupancy for all 61 codons for Elongator-related mutants, relative to WT average. Gray diamonds show average occupancy (zero, by definition) and SD for nine replicates of BY4741; red diamonds show average and SD for 30 Elongator-related datasets (two biological replicates for 15 deletion
mutants). The expected increase in ribosome occupancy is confirmed over AAA, CAA, and GAA, as indicated.

(B) A site occupancy for two biological replicates of mod5Δ, with 8 known target codons at left columns (indicated) and all remaining codons at right columns sorted alphabetically.

(C) A and P site occupancy for two biological replicates of tan1Δ, with 12 known target codons for serine and leucine at left columns (indicated) and all remaining codons at right columns sorted alphabetically.

Figure 2.6
**Figure 2.6. Unanticipated mutant effects on codon occupancy**

(A) A site ribosome occupancy for all 61 codons sorted alphabetically for KEOPS mutants, with two biological replicates for each of bud32Δ, gon7Δ, and cgi121Δ. We noted that although t6A37 is found in tRNAs decoding ANN codons, KEOPS mutants affect mostly codons that start with C, which are labeled in red.

(B) Discrepant behavior between mutants affecting yW synthesis. A, P, and E site occupancy data shown for Phe codons for the four tyw mutants, as indicated: Tyw1–3 are indistinguishable by design, while the discrepant behavior of Tyw4 is visually emphasized.

(C) A and P site occupancy for trm8Δ and trm82Δ, with 16 relevant codons to the left (indicated) and all remaining codons sorted alphabetically.

**Table 2.6. Codon occupancy dataset**

Ribosome footprints of 28-31 nt were analyzed as described in Materials and Methods. Columns show individual replicates of wild type and various mutants, while rows show A-, P-, or E-site occupancy on 61 codons (no stop codons are included in the analysis since the first 15 and the last 5 codons of each transcript were removed from the reference). Link to dataset:

https://www.ncbi.nlm.nih.gov/geo/download/?acc=GSE100626&format=file&file=GSE100626%5FChou%5FTable%5FS4%2Etxt%2Egz
**Trm7 Methylates Both tRNAs and rRNAs**

Turning from the relatively subtle codon-level ribosome occupancy phenotypes described above to gene-level analysis of ribosome occupancy, we noted a particularly dramatic phenotype in initial surveys of the ribosome footprint landscape. Specifically, yeast lacking the tRNA methylase Trm7, which is required for 2′-O-methylation at positions 32 and 34 of the anticodon loop of several tRNAs (Pintard et al., 2002), exhibit dramatic and widespread changes in ribosome occupancy at a large number of 5′ UTRs and 5′ coding regions (Figure 2.7A). This was unique to \textit{trm7Δ} among all 57 mutations in this study (Figure 2.7B) and appeared to be an artifact of globally reduced translation in this mutant: the 5′ enrichment of ribosomes was only observed in libraries prepared in the presence of cycloheximide (Figure 2.7C), suggesting that an excess of free ribosomal subunits is present in this mutant that can assemble onto mRNA upstream of cycloheximide-arrested ribosomes during cell lysis. Consistent with this hypothesis, we confirmed a significantly decreased abundance of polysomes in \textit{trm7Δ} mutants (Figure 2.7D), as previously reported (Pintard et al., 2002).

This dramatic translational phenotype suggested that Trm7 may have additional substrates, particularly since neither of the two known partners of Trm7—Trm732 and Rtt10—had similar effects on ribosome occupancy at 5′ UTRs (Figure 2.7B). As the closest Trm7 homolog in bacteria is an rRNA methylase (Pintard et al., 2002), we tested the hypothesis that Trm7 might also methylate rRNA in budding yeast. We assessed 2′-O-methylation of rRNA using
RiboMeth-seq (Marchand et al., 2016), in which limited alkaline hydrolysis of RNA is used to cleave RNAs at all positions with an unmodified 2′ hydroxyl, allowing sequencing-based identification of 2′-O-methylation sites based on a reduction in sequencing reads starting immediately downstream of the methylated ribose.

Our RiboMeth-seq data obtained from WT yeast recovered all known 2′-O-methylation sites in yeast 18S and 25S rRNA, with high concordance between eight replicate libraries (Figures 2.8A and Table 2.7). Comparing rRNA methylation in WT, trm7Δ, and the unrelated trm3Δ mutant (n = 8 each), we identified five significantly hypomethylated sites, all of which were specific to the trm7Δ mutant (Figures 2.9A-D). Extending this analysis to several additional mutants, including other 2′-O-methylases as well as Trm7’s known dimerization partners, revealed that four of the five candidate Trm7 target sites were affected exclusively in trm7Δ, but also that the strongest candidate, C663, was in addition hypomethylated in mutants lacking Rtt10, one of the two known heterodimerization partners for Trm7 (Figure 2.9E). Given that the observed depletion of 5′ end reads starting downstream of C663 is not complete (Figure 2.9C), we infer that this nucleotide is likely to be partially methylated across the population of ribosomes in the cell, perhaps suggesting a role for Trm7/Rtt10 in context-dependent methylation of specific subpopulations of ribosomes. Thus, although the basis for the global effects of Trm7 on translation
remains unclear, our data do support the hypothesis that the tRNA methylase Trm7 also methylates a subset of rRNA molecules \textit{in vivo}.

Figure 2.7. Dramatic effects of \textit{trm7} on ribosome occupancy profiles
(A) Increased ribosome occupancy at 5’ UTRs in \textit{trm7Δ} mutants. RNA-seq and RPF data for WT and \textit{trm7Δ} mutant yeast at characteristic genomic loci. Red arrows show examples of increased ribosome occupancy in the mutant.
(B) 5’ ribosome accumulation is unique to \textit{trm7Δ} mutants in this dataset. Metagene shows ribosome occupancy data averaged across all genes, aligned by start codon. Main plot shows data for wild-type and \textit{trm7Δ} mutant, while zoom-in on 5’ UTRs shows all individual datasets.
(C) Increased ribosomal 5’ UTR and 5’ coding occupancy in *trm7Δ* mutants is only observed in ribosome footprinting libraries generated in the presence of cycloheximide. Averaged ribosome occupancy from -100 to +110 nt relative to the AUG is shown for the four indicated datasets: ribosome profiling for wild-type and *trm7Δ*, carried out in the presence or absence of cycloheximide. Excess ribosome occupancy of 5’ UTRs is only observed in the presence of cycloheximide. Cycloheximide-dependent 5’ UTR occupancy is a known consequence of global defects in translation. Briefly, in mutants that cause an excess of unassembled 40S and 60S ribosome subunits, artefactual assembly onto mRNAs can occur during cell lysis. In the presence of cycloheximide-stalled ribosomes, the artefactual ribosome assembly primarily occurs at 5’ ends of mRNAs, as observed for the *trm7Δ* mutant. Along with the overall loss of polysomes in this mutant (Fig 2.7D), we conclude that *trm7Δ* mutants are characterized by a global defect in translation.

(D) Polysome profiles of the indicated strains reveal a global deficit in translation in *trm7Δ*. WT and unrelated *trm4Δ* are shown for comparison.
Figure 2.8. RiboMeth-seq analysis of rRNA 2′-O-methylation in WT yeast

(A) Validation of 18S RiboMeth-Seq dataset. Top panel shows counts (normalized to reads per million rRNA-mapping reads) of sequencing reads starting across 54 nt of 18S rRNA. The three annotated locations are dramatically underrepresented and correspond to three well-known 2′-O-methylation sites on 18S rRNA. Bottom panel shows methylation “A scores” (Birkedal et al., 2015; Marchand et al., 2016) aggregated for eight WT datasets—individual replicates are nearly indistinguishable—with asterisk indicating previously validated methylation sites.

(B) Validation of 25S RiboMeth-Seq dataset. A scores for the WT aggregate dataset are plotted as in (A) bottom panel, with previously described 2′-O-methylation sites represented with asterisk.
Figure 2.9

A. 25S rRNA

B. A score, trm3Δ

C. 5' end counts

D. 5' read starts

E. A score, mutant - WT
Figure 2.9. Trm7 effects on rRNA 2’-O-methylation

(A-B) Five hypomethylated sites of 25S rRNA in trm7Δ were identified using RiboMeth-seq. Scatterplot compares methylation A scores for WT (x axis, n = 8) and trm7Δ (y axis, n = 8) strains (A). The five significantly differentially methylated nucleotides are indicated with large purple points and lose methylation in trm7Δ but are unaffected in the unrelated trm3Δ mutant (B).

(C) Normalized RiboMeth-seq 5’ end read starts (as in Figure 2.8A, top panel) for 14 nt surrounding 25S rRNA C663, as indicated.

(D) Statistically significant effects of Trm7 on highly-methylated rRNA nucleotides represent minor subpopulations of ribosomes. Normalized 5’ end counts shown for a short region encompassing Gm805 and Am807, with insets showing that significant effects of Trm7 here represent an increase in a very small subpopulation of unmethylated ribosomes at these sites.

(E) Comparison of mutant effects on the five candidate Trm7 target sites in 25S rRNA, shown as the change in A score for each mutant replicate relative to the average of eight WT replicates. For the five Trm7 target nucleotides, data are shown for WT, trm7, and trm3 mutants (n = 8 each), and for trm13, trm44, trm732, and rtt10 mutants (n = 2 each). Note that C663 methylation is lost in mutants affecting Trm7 as well as one of its heterodimerization partners, Rtt10, while the remaining four potential Trm7 target sites are not affected by either Trm732 or Rtt10.

Table 2.7. RiboMeth-Seq dataset

RiboMeth-Seq data for WT yeast and the indicated mutants. Sheets include raw counts of 5’ read starts. Link to dataset:

https://www.ncbi.nlm.nih.gov/geo/download/?acc=GSE100626&format=file&file=GSE100626%5FChou%5FTable%5FS5%2Etxt%2Egz
mRNA Abundance Changes Report on Diverse Cellular Functions Impacted by tRNA Modifications

We next sought to assess how the loss of specific tRNA modifications affects genomic output at the levels of mRNA abundance and ribosome occupancy (and, by extension, translation efficiency) of individual genes. As expected, changes in mRNA abundance were generally reflected in the ribosome occupancy dataset (Figure 2.10). We focus first on mRNA levels, as changes in translational efficiency of key regulators (such as transcription factors) can cause widespread physiological and transcriptional changes that are readily appreciated by RNA sequencing (RNA-seq), effectively amplifying the signal for biologically important translational regulation events.

Overall, we find robust mRNA abundance changes in roughly half (27/57) of the mutants analyzed, with minimal or no effects on mRNA abundance observed for the remaining 30 mutants (Figure 2.11A). The vast majority of mutants exhibiting substantial impacts on cell function, as measured by gene expression changes, are known to affect tRNA modifications in the anticodon itself or immediately adjacent to the anticodon, including (1) ncm$^5$U, mcm$^5$U, and mcm$^5$s$^2$U at the wobble position (Elongator and related factors); (2) threonylcarbamoyladenosine (KEOPS); (3) wybutosine (Tyw3); (4) isopentenyladenosine (Mod5); (5) pseudouridine (Pus3/Deg1 and Pus7); and (6) ribose 2'-O-methylation (Trm7). These findings are consistent with the expectation that anticodon modifications should have greater effects on the
primary tRNA function—decoding mRNAs—than modifications that occur elsewhere in the tRNA molecule. In addition to the factors involved in anticodon modification, a small number of proteins involved in tRNA modifications distant from the anticodon, such as Rit1, which is required for the 2′-O-ribosylphosphate modification at A64 of initiator tRNA that prevents initiator tRNA from participating in translational elongation (Astrom and Bystrom, 1994), also exhibited robust gene expression changes. Interestingly, several mutants that robustly affected codon-specific translation, such as trm82Δ and tan1Δ, had only modest effects on gene expression.

Consistent with prior studies on Elongator (Deng et al., 2015; Nedialkova and Leidel, 2015; Zinshteyn and Gilbert, 2013) and KEOPS (Daugeron et al., 2011), we observe upregulation of a large group of genes, primarily involved in amino acid biosynthesis and related metabolic pathways, in these mutants. This upregulation can be attributed to translational upregulation of the nutrient- and tRNA-responsive transcription factor Gcn4 (Hinnebusch, 2005) (Figures 2.11B-D). Our data recapitulate these prior findings, further validating the dataset. Moreover, we find that expression of GCN4 is translationally upregulated in several additional mutants, including pus3Δ, pus7Δ, rit1Δ, trm1Δ, trm7Δ, mod5Δ, and tyw3Δ. Thus, a wide variety of aberrations in tRNA function convergently result in increased synthesis of Gcn4, presumably as a consequence of impaired translation of regulatory uORFs in the GCN4 5′ UTR. We also find that the upregulation of a proteostasis stress response previously described in
Elongator mutants (Nedialkova and Leidel, 2015) is exhibited in additional mutants, with \( \textit{PRE3} \) upregulation occurring in mutants affecting Elongator, KEOPS, and Trm112, and \( \textit{MSN4} \) upregulation occurring more broadly across the set of mutants that affect the Gcn4 response.

While the loss of multiple distinct tRNA-modifying complexes induced a common transcriptional response through \( \textit{GCN4} \) upregulation, other gene expression changes were confined to a more limited set of mutants (Figure 2.11A) and thus were clearly not secondary to increased cellular levels of Gcn4. Among upregulated genes, a large group of genes related to mitochondrial function and carbohydrate metabolism were upregulated in KEOPS mutants and in \( \textit{tyw3} \Delta \) and \( \textit{sit4} \Delta \) mutants. Importantly, this was not a result of these strains having lost their mitochondrial DNA, as we detected abundant transcripts for mitochondrially encoded genes in all of these strains. Other potential explanations for the physiology underlying this gene expression program include altered mitochondrial function resulting from loss of mitochondrial tRNA modifications, or altered expression of the respiration-regulating Hap4 transcription factor. However, in several mutants it is unlikely that this gene expression signature results from loss of the relevant tRNA modifications. Most notably, although the cell-cycle phosphatase Sit4 is required for formation of \( \textit{mcm}^5\text{s}_2\text{U} \) (Huang et al., 2008), \( \textit{sit4} \Delta \) is the only Elongator-related mutant exhibiting the carbohydrate/mitochondria transcriptional phenotype. Similarly, although all four Tyw proteins are required for wybutosine formation, only \( \textit{tyw3} \Delta \)
mutants (which accumulate tRNAs modified with the “yW-86” intermediate; (Noma et al., 2006)) upregulate HAP4 and related genes. On the other hand, this phenotype might potentially reflect a bona fide consequence of loss of t6A in KEOPS-related mutants—although the aneuploid bud32Δ and gon7Δ strains exhibit much stronger upregulation of carbohydrate metabolism genes than docgi121Δ mutants, cgi121Δ mutants, which maintain ~80% of WT levels of t6A, do exhibit a modest effect on these genes (Figure 2.11A). Given the unusual pattern of mutants exhibiting upregulation of carbohydrate metabolism genes, we did not further pursue this connection, although it may prove (at least in the case of KEOPS) an interesting area for future study.
Figure 2.10

mRNA abundance

Ribosome footprints
Figure 2.10. Comparison of RNA-Seq and RPF datasets
Overview of all RNA-seq (left panel) and RPF (right panel) changes across the 57 mutants in this study. Data are shown for all genes changing at least 2-fold in at least two mutants (filtered for average mRNA abundance >10 RPKM; reads per kilobase of transcript per million mapped reads). RPF data were sorted identically to the RNA-Seq dataset to enable side-by-side comparison.
Figure 2.11. Effects of tRNA-Modifying Enzymes on RNA Abundance

(A) Overview of all RNA-seq changes across the 57 mutants in this study. Data are shown for all genes changing at least 2-fold in at least two mutants (filtered for average mRNA abundance >10 RPKM; reads per kilobase of transcript per million mapped reads). Boxes show five relatively coherent gene expression clusters, with prominent functional annotations enriched in each geneset indicated.

(B) Translational upregulation of GCN4 is a common occurrence in tRNA modification mutants. Top panels show RNA-seq and RPF data for the GCN4 ORF and its 5' UTR, which carries four well-studied regulatory upstream ORFs (uORFs). Zoom-ins focusing on the GCN4 coding region show RNA-seq and RPF data for the indicated mutants.

(C) Mutant effects on GCN4 RNA and RPF levels are shown for all mutants, sorted from high to low Gcn4 translational upregulation.

(D) RNA-seq correlates of GCN4 translational upregulation. Rows show Gcn4 targets (genes that exhibit >2-fold increase in RNA Pol2 occupancy in (Qiu et al., 2016) for all mutants, sorted as in (C).
Silencing-Related Phenotypes in tRNA Modification Mutants

We next addressed silencing-related phenotypes in the dataset, as tRNA-modifying enzymes have previously been implicated in silencing of subtelomeric and mating-type reporters (Chen et al., 2011b; Li et al., 2009b), and in telomere capping and recombination (Downey et al., 2006; Peng et al., 2015). Defects in various aspects of heterochromatin silencing in yeast result in separable transcriptional responses, which in turn provide robust proxies for function of the relevant silencing pathways. For example, the dramatic downregulation of haploid-specific genes (HSGs) (such as those encoding the pheromone response pathway) observed in \textit{bud32Δ} and \textit{gon7Δ} mutants (Figure 2.11A) is typical of the “pseudo-diploid” state of budding yeast mutants that fail to repress the silent mating loci (Rusche et al., 2003).

To systematically explore mutant effects on mating locus and subtelomeric silencing (Figure 2.12A), we plotted expression of HSGs and subtelomeric genes across all mutants in this study (Figure 2.12B, top and middle panels). We also included \textit{PHO} genes as a separate class (bottom panel)—although several \textit{PHO} genes are located near chromosome ends, we noted that both telomere-proximal and -distal \textit{PHO} genes were downregulated in essentially all mutants that affect \textit{GCN4} mRNA translation (Figure 2.11A). Mutants in Figure 2.12B are sorted according to their effects on HSG expression, as \textit{bud32Δ} and \textit{gon7Δ} exhibit the most dramatic downregulation of HSGs. A small number of additional mutants (\textit{sit4Δ}, \textit{kti11Δ}, \textit{trm7Δ}, \textit{mod5Δ}, and \textit{cgi121Δ})
showed moderate (~1.5-fold) downregulation of these genes, with the majority of
mutants, including most of the Elongator-related mutants, exhibiting extremely
subtle changes in HSG expression. Derepression of subtelomeric genes was
similarly restricted primarily to KEOPS mutants. Downregulation of
subtelomeric \textit{PHO} genes such as \textit{PHO89} (Figures 2.12A and B) is shown here
to emphasize the distinction between the small subset of mutants that specifically
affect silencing-related phenotypes, and the larger class of mutants exhibiting
relatively nonspecific phenotypes (such as Gcn4 upregulation).

As the silencing phenotype previously reported for \textit{elp} mutants was
ascribed to defective translation of the \textit{SIR4} mRNA (Chen et al., 2011b), we next
examined the translational efficiency of the \textit{SIR} mRNAs in our dataset
(Figure 2.12C). Consistent with the dramatic silencing defects observed
in \textit{bud32Δ} and \textit{gon7Δ}, we found that the translational efficiency of \textit{SIR2} mRNA
was significantly decreased in these two mutants (Figure 2.12C). However,
outside of this connection, silencing phenotypes were otherwise poorly correlated
with \textit{SIR} mRNA translational efficiency as assayed by ribosome footprinting. For
example, although \textit{mod5Δ} mutants did exhibit modest changes in \textit{SIR2} mRNA
translation accompanied by moderate downregulation of HSGs, nearly identical
changes in \textit{SIR2} translation in \textit{pus3Δ} and \textit{ncs2Δ} mutants did not result in
appreciable mating locus derepression (Figures 2.12B and C).

Taken together, our data indicate that defects in silencing of endogenous
loci are relatively rare in mutants that affect tRNA modifications, with the most
dramatic silencing defects being confined to the unusual case of the two aneuploid KEOPS mutant strains.

Figure 2.12. Analysis of silencing-related phenotypes

(A) RNA-seq data for an ~15 kb locus adjacent to TEL2R. Two notable phenotypes are indicated with arrows: repression of PHO genes, observed in a wide range of mutants in this study (Figure 2.11A), and derepression of a subset of subtelomeric genes, which is confined primarily to mutants in the KEOPS complex. (B) Mutant effects on expression of HSGs (a robust reporter for silent mating locus derepression), subtelomeric genes, and PHO genes, as indicated. Mutants are sorted by their average effect on HSGs. (C) Mutant effects on translational efficiency of Sir proteins.
**Regulation of TY1 Expression by Elongator**

Although robust silencing defects were largely confined to bud32Δ and gon7Δ, we uncovered a surprising silencing-related phenotype in Elongator-related mutants. Specifically, we observed substantial downregulation of TY1 retrotransposon expression occurring almost exclusively in Elongator-related mutants (Figure 2.13A). Interestingly, a small number of endogenous protein-coding genes were also downregulated in the same subset of mutants, and inspection of these genes revealed that all such genes are located in genomic neighborhoods in proximity to intact TY elements, solo TY1 LTRs, and tRNA genes (Figure 2.13B). This link was of great interest to us given the central role for tRNAs in the biology of LTR retrotransposons (Marquet et al., 1995; Weiner and Maizels, 1987), and raises the question of how the Elongator complex affects TY-linked gene expression.

Although diminished TY1 expression occurs, counterintuitively, in sir mutants (Lenstra et al., 2011), this is unlikely to explain the downregulation we document for Elongator mutants: KEOPS and other mutants that affect other aspects of Sir-dependent silencing in this dataset do not cause TY1 repression, and conversely, the various Elongator-related mutants exhibit subtle or no effects on other Sir-dependent phenotypes (Figures 2.12B and 2.13A). In addition, the highly Elongator-specific effects on TY1 regulation cannot be a result of GCN4 upregulation, which occurs in a much broader group of mutants (Figure 2.11).
We next considered the possibility that TY-adjacent genes are affected in Elongator mutants as a secondary consequence of altered levels of the target gene TYE7 (Figures 2.13A and B), which encodes a known transcriptional activator of TY LTRs (Lohning and Ciriacy, 1994). However, qRT-PCR of several target genes in a tye7Δ background revealed further decreases in mRNA abundance in tye7Δelp3Δ and tye7Δuba4Δ double mutants (Figure 2.13C), demonstrating that altered regulation of TY1 elements does not result from Elongator’s effects on the endogenous TYE7 locus.

Finally, although most phenotypes of Elongator mutants are suppressed by overexpression of a subset of its target tRNAs (Chen et al., 2011b; Esberg et al., 2006; Nedialkova and Leidel, 2015; Zinshteyn and Gilbert, 2013), we found that Elongator’s effect on expression of TY-adjacent genes was unaltered by overexpression of two such tRNAs (Figure 2.13D). These data suggest that Elongator’s effects on TY element expression could result from the modification of other target tRNAs, or, more intriguingly, that its control of TY expression may not require wobble nucleotide modification. Given the many links between tRNAs and LTR element replication (Marquet et al., 1995), it will be of great interest in future studies to determine how Elongator functions to support expression of genes located near TY LTRs.
Figure 2.13

A

B

C

D

Fold change relative to WT
Figure 2.13. Downregulation of TY1 expression in Elp-related mutants

(A) Cluster shows TY1 and TY-adjacent genes, same as “TY1 elements” cluster in Figure 2.11A. Structural genes encoded by the TY1 retroelement are indicated with orange boxes.

(B) ORFs downregulated in Elongator-related mutants are associated with TY1 long terminal repeats (LTRs). Top panels show RNA-seq data for TYE7 for WT and a representative Elongator-related mutant. Bottom panels show genomic loci associated with the ORFs shown in (A).

(C) Elongator effects on target genes are not mediated via changes in TYE7 expression. qRT-PCR for two ORFs and for a TY1 element, as well as two normalization controls (TEF1 and TDH3), was performed in one of six strain backgrounds (WT, uba4D, elp3D, tye7D, tye7Duba4D, and tye7Delp3D, as indicated) in four replicates. All data are normalized to the WT expression levels. Left panel validates our RNA-seq observations, while right panel shows Elongator effects on these genes in the absence of Tye7. TY1 mRNA levels are decreased in tye7D, as expected, but, importantly, deletion of Elongator leads to a further decrease in TY1 expression.

(D) q-RT-PCR data for the three indicated genes is shown for a panel of 12 mutant yeast strains, as indicated. For each gene, data are normalized to the wild-type value. tRNA overexpression plasmids are described in (Leidel et al., 2009).
Gene-Specific Changes in Translational Efficiency Reveal Regulatory uORFs

We finally turn to analysis of translational efficiency in our dataset. Although mutant effects on the translational control of key regulatory genes, such as GCN4, result in an amplified response at the level of mRNA abundance, altered synthesis of many proteins is not expected to cause dramatic transcriptional phenotypes and thus mutant effects on translational efficiency must be addressed directly.

We note that increased ribosome occupancy on a coding region can result from increased translation, as observed, for example, for GCN4, or from slowed or stalled translation. However, analysis of mutants with significant effects on codon-level occupancy (Figures 2.5 and 2.6) revealed that although in some cases genes with high levels of the affected codons exhibited increased ribosome occupancy (Figure 2.14), suggesting slowed or stalled translation, this effect was quantitatively extremely modest overall, with codon frequency in ORFs typically explaining no more than ~1%-2% of the variance in translational efficiency. Thus, slowed translation contributes modestly to overall ribosome occupancy, which instead primarily reports on translational efficiency.

Figure 2.15A shows clustered translational efficiency for all 57 mutants, relative to WT. As with mutant effects on mRNA abundance, we noted that mutants that affect nucleotide modifications at or adjacent to anticodons exhibited altered translation of many more transcripts than did mutants affecting
distal nucleotides. Interestingly, although we identified a handful of relatively specific translational changes in subsets of these mutants, overall, we find that the majority of mutants that affect anticodon modifications tend to affect translation of a common group of mRNAs, suggesting that many of the affected genes respond to some aspect of overall translational efficiency (e.g., efficiency of uORF translation), rather than to levels or functionality of individual tRNAs.

Focusing first on translationally upregulated genes, we identified a small group of target genes that exhibited a similar mutant profile to that of *GCN4*. Most notably, we found that translation of *SER3* mRNA was highly correlated with that of *GCN4* across our dataset. Closer inspection of the *SER3* locus revealed clear evidence for ribosome occupancy upstream of the *SER3* start (Figure 2.15B), falling within a previously described regulatory transcript known as *SRG1* (Martens et al., 2004). Although *SRG1* was originally described as a sense-strand cryptic transcript that is terminated near the *SER3* AUG, Martens et al. also noted the presence of long readthrough *SRG1* transcripts extending to the *SER3* 3′ end, and we find multiple sequencing reads spanning the *SRG1/SER3* junction, indicating that a subset of *SER3* transcripts include the *SRG1* sequence as their 5′ UTR. These results are most consistent with a model in which translation of *SER3* mRNA is regulated by a uORF in a manner analogous to the intensively studied mechanism of *GCN4* regulation (Hinnebusch, 2005), and imply that *SRG1* plays separable roles in regulation of *SER3* at both transcriptional and translational levels.
Turning next to translationally downregulated genes, we noted that CPA1, which is known to be translationally regulated by an upstream “attenuator peptide” (Gaba et al., 2005), was downregulated in essentially the same broad set of mutants that affect GCN4 translation (Figures 2.15A and C). A number of other transcripts were translationally repressed in the same set of mutants that affected CPA1, and in many cases we found evidence for uORFs that likely confer translational regulation on the downstream ORFs (Figures 2.15A, D, and E). Together, these data provide an expanded survey of presumptive regulatory uORFs, with potential implications for understanding the distinctions between uORFs with stimulatory, versus repressive, effects on downstream ORF translation.
Figure 2.14. Codon-level pauses have minimal impact on ORF-level ribosome occupancy data

Scatterplots for six mutants showing mutant effects on translational efficiency (y axis, log2 fold change relative to wild-type) for all genes vs. the fraction of "relevant codons" in the open reading frame (x axis). Relevant codons here are defined as those exhibiting increased A site (or P site, in the case of trm82) occupancy in Figures 2.5 and 2.6, as follows: kti11: CAA, AAA, GAA; tan1: CUA, CUC, CUG, CUU, UUA, UUG, AGC, AGU, UCA, UCC, UCG, UCU; trm82: GUC, GUG, GUU; elp5: CAA, AAA, GAA; tyw3: UUC, UUU; trm7: UUU. Although there is a modest positive correlation for a handful of these scatterplots, the largest R² here is 0.015 for elp5, meaning that ORF codon composition explains at most 1-2% of ribosome occupancy changes in mutants.
Figure 2.15. Effects of tRNA-modifying enzymes on translational efficiency

(A) Overview of translational efficiency dataset. Heatmap shows log2 fold changes, relative to WT, of all genes with TE changes of at least 2-fold in two or more mutants.

(B) Translational regulation of SER3 by uORFs. RNA and RPF (ribosome-protected footprint) data are shown for WT, pus7D (where SER3 is unaffected), and elp1D, in which SER3 translational efficiency is increased. Notable here is a peak of ribosome occupancy over the upstream regulatory transcript SRG1, which is lost (despite no change in SRG1 RNA abundance) in mutants that translationally derepress SER3.

(C–E) Examples of genes translationally repressed in various tRNA-modifying enzyme mutants. Data shown as in (B), with green arrows highlighting diminished ribosome occupancy of ORFs, and red arrows highlighting likely regulatory uORFs. Here, known (CPA1; C) or putative (CMR3 and YGP1; D and E) upstream regulatory ORFs are highlighted in red in the genomic annotation.
Discussion

This dataset provides a unique resource for understanding the roles for tRNA-modifying enzymes and their various cofactors in translational regulation. A key feature of this study is the comparison of multiple disparate mutants within the same dataset, which provides a valuable opportunity to constrain hypotheses for the mechanisms underlying translational phenotypes of interest (see, for example, Figures 2.12 and 2.13).

Overall, we find that those tRNA modifications that occur at or adjacent to the anticodon have the greatest effects on ribosome occupancy, as expected. That said, some modifications that are distant from the anticodon (e.g., tRNA position 12) also have strong effects on ribosome occupancy. The absence of transcriptional or ribosome occupancy phenotypes for many of the remaining mutants involved in tRNA modification could reflect a variety of factors: regulatory feedback could maintain high levels of tRNAs that are destabilized in the absence of a given modification, or certain tRNA modifications could play important roles under alternative growth conditions that stress the proteostasis machinery.

More granular analyses at varying levels of resolution from gross transcriptional phenotypes to gene-centric and codon-centric ribosome occupancy reveal both expected behaviors of various mutants as well as unanticipated observations that inform mechanistic hypotheses for future study. We highlight several striking examples of such findings, such as the distinction
between the *tyw* mutants in A and P site occupancy at relevant codons, and many additional related examples can be found in the various Tables.

Most surprisingly, we discover a role for the Elongator complex and other factors required for wobble U modifications (*mcm^5^s^2^U and related) in control of expression of both intact *TY1* elements as well as multiple endogenous genes associated with solo LTRs. This finding is of great interest given the ancient links between tRNAs and LTR retroelements—tRNAs or tRNA-like RNA structures almost universally serve as primers for reverse transcriptase (Marquet et al., 1995; Weiner and Maizels, 1987), and recent studies implicate cleaved tRNA fragments in control of LTR-associated genes (Martinez et al., 2017; Schorn et al., 2017; Sharma et al., 2016). We consider a number of hypotheses for the mechanistic basis for Elongator control of *TY1*, ruling out roles for the Sir complex, Gcn4, or Tye7 as mediators of this effect. These findings reveal a surprising connection between a tRNA-modifying complex and control of LTR elements, and mechanistic dissection of the role for Elongator in *TY* transcription or mRNA stability will be of great interest.
Materials and Methods

**Yeast strains and culture conditions**

All strains were generated in the BY4741 background (MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0). Haploid deletions were generated by sporulation and tetrad dissection of heterozygous deletions in the diploid BY4743 background, which were obtained either from the Yeast Knockout Heterozygous Collection (Dharmacon), or generated *de novo* via replacement of genes of interest with KanMX6 in BY4743. Haploid deletions were selected on YPD+G418, Lys minus (SD +His +Leu +Met +Ura with glucose), and Met minus (SD +His +Leu +Lys +Ura with glucose) media. MATα and deletion genotypes were verified by PCR.

Cells were cultured in YPD+G418 at 30°C for small amounts and were amplified in YPD at 30°C for experiments. After initial analysis of RNA-Seq from all deletion mutants, we identified clear evidence of aneuploidy (consistently elevated expression across entire chromosomes) for several strains. These strains were freshly re-made, and, with the exception of *gon7Δ* and *bud32Δ* for which all isolates obtained carried an addition ChrIX copy, all remade strains were confirmed to be euploid.

Strains *tey7Δ, tey7Δ elp3Δ*, and *tey7Δ uba4Δ* were generated by replacing *TYE7* with *URA3* in BY4741, *elp3Δ*, and *uba4Δ*, respectively, and selected on SC-Ura media. The cells were cultured in SC-Ura media for small amounts and were amplified in YPD at 30°C for experiments. Yeast strains with
tRNA overexpression plasmids were generated using lithium acetate-based transformation and selected on SC-Leu media. The plasmids were generous gifts from Sebastian Leidel and Wendy Gilbert.

**Ribosome profiling and RNA-seq**

Ribosome profiling was carried out as described in (Heyer and Moore, 2016), with minor modifications. Yeast strains were grown overnight to mid-log phase (OD$_{600} = 0.5$-$0.6$) in YPD at 30°C, treated with cycloheximide (CHX) in a final concentration of 100 μg/ml for 30 s with vigorous shaking, and harvested by centrifugation at 4°C for 2 minutes. The cell pellets were flash-frozen in liquid nitrogen immediately. Total time from adding CHX to snap-freezing was about 4 minutes.

For ribosome footprinting, cells were lysed in ice-cold lysis buffer (20 mM Tris-HCl, 140 mM KCl, 1.5 mM MgCl$_2$, 1% Triton X-100, 0.5 mM DTT, 100 μg/ml CHX) by glass-bead beating in an ice-cold block, and the clarified ribosome extract was obtained by centrifugation for 5 minutes at 9,500 rpm, 4°C. The ribosome extract was diluted with ice-cold lysis buffer, aliquoted to 5 A$_{260}$ units per tube, and followed by incubating with 500U of RNase I at 23°C for 1 hour. Monosomes were separated on 12 mL of 10%–50% sucrose gradient (20 mM Tris-HCl, 140 mM KCl, 5 mM MgCl$_2$, 0.5 mM DTT, 20 U/ml SUPERase-In, 100 μg/ml CHX) in Sw40 ultracentrifuge tubes by centrifugation for 160 minutes at 35,000 rpm, 4°C. Gradients were fractioned using Brandel Density Gradient
Fractionation System, and 80S monosome fractions were collected and flash-frozen immediately. RNA was extracted using TRIzol and 2-propanol precipitation. 27-34 nt ribosome footprints were isolated using denaturing 15% polyacrylamide-TBE-urea gels and then purified with Zymo ZR small-RNA PAGE Recovery Kit.

For CHX-free ribosome profiling, yeast cells were harvested by rapid filtration onto 0.8-um-pore-size mixed cellulose membrane. The cells were scraped and snap-frozen in liquid nitrogen immediately. Frozen cells and frozen pellets of lysis buffer (the same as above) were ground together into fine powder using a Retsch CryoMill with sample chambers pre-chilled in liquid nitrogen. After thawing the frozen powder on ice, the ribosome extract was obtained by centrifugation for 5 minutes at 9,500 rpm, 4°C, and the same procedure was followed.

For RNA-Seq, total RNA was extracted from the same yeast culture for ribosome footprinting using TRIzol, followed by DNase digestion and purification using Zymo RNA Clean & Concentration Kit. DNase-free RNA was depleted of rRNA using Illumina Ribo-Zero Gold Kit, followed by zinc-based fragmentation for 10 min at 70°C. RNA fragments and ribosome footprints were constructed into deep-sequencing libraries as described in (Heyer et al., 2015). Briefly, RNA 3’ ends were dephosphorylated using T4 PNK, ligated with a 5’-preadenylated adaptor using T4 RNA Ligase 2 truncated K227Q, and reverse transcribed with barcode primers. cDNA was precipitated using 2-propanol and size selected
using denaturing 6% polyacrylamide-TBE-urea gels. cDNA was purified using traditional “crush and soak” method, followed by circularization with epicenter CircLigase and PCR amplification prior to sequencing.

2′-O-methylation sequencing (RiboMeth-seq)

RiboMeth-seq was carried out essentially as described in (Marchand et al., 2016), except for 3′ end dephosphorylation. 500 ng of total RNA extracted for RNA-seq was subjected to fragmentation using alkaline solution (100 mM Na₂CO₃-NaHCO₃, pH 9.2) for 10 minutes at 95°C, followed by RNA purification using Zymo RNA Clean & Concentration Kit. Purified RNA fragments were 3′ dephosphorylated using various enzymes, 5′ phosphorylated using T4 PNK in the presence of ATP, and finally constructed into deep-sequencing library using NEBNext Small RNA Library Prep. For WT, trm7Δ, and trm3Δ, we initially generated 2 replicate datasets each using T4 PNK, Antarctic phosphatase, or Shrimp alkaline phosphatase for 3′ end dephosphorylation, with no significant effects of any of these variant protocols on methylation. A second round of libraries was built using T4 PNK for 3′ end dephosphorylation, with two additional replicates for WT, trm7Δ, and trm3Δ (final n = 4 for T4 PNK protocol, n = 8 across all 3 protocols for these three strains), as well as 2 replicates each for trm13Δ, trm44Δ, trm732Δ, and rtt10Δ.
**Polysome profiling**

The preparation of ribosome extract and sucrose gradient for polysome profiling was the same as mentioned above, but with incubating 5 A260 units of ribosome extract with SUPERase-In instead of RNase I. Fractionation and detection of ribosomes were performed using the same system and its data capture software.

**Quantitative RT-PCR**

Total RNA was isolated using TRIzol with glass-bead beating, followed by the purification using Zymo Direct-zol RNA MiniPrep with in-column DNase I treatment. cDNA was generated using SuperScript IV Reverse Transcriptase with random hexamers for priming. Quantitative PCR was performed using KAPA SYBR Fast qPCR Master Mix 2X Universal with ~10 ng of cDNA and ROX Low in QuantStudio 3 Real-Time PCR System. Relative fold changes in mutants were calculated as $2^{\Delta\Delta Ct}$, and the gene expression was normalized to the expression of TEF1 and TDH3.

**Analysis for ribosome profiling and RNA-seq**

**Sequencing read mapping and gene-level analysis**

Barcoded libraries were pooled and sequenced on an Illumina NextSeq500. Raw fastq reads were de-multiplexed and removed of adaptor sequence using HOMER package (Heinz et al., 2010). RPF reads were mapped to S. cerevisiae rDNA and the mapping reads were discarded. The remaining
RPF reads and RNA-seq reads were mapped to sacCer3 genome using TopHat v2.0.12 (Trapnell et al., 2009) with parameters `-p 4 -l 5000 --no-coverage-search`. Unique mapping reads were saved using SAMtools (Li et al., 2009a) `view` function and parameter `-q 10`. Reads in length of 27-34 nt (RPF) or ≥ 27 nt (RNA-seq) were used for gene-level analysis, in which reads were quantified as raw counts or reads per kilobase of transcript per million mapped reads (RPKM) using HOMER `analyzeRepeats.pl` function with open reading frame annotations downloaded from Saccharomyces Genome Database (SGD). The read coverage in UCSC Genome Browser tracks were generated using HOMER `makeUCSCfile` function with parameters `-fragLength 30` for RPF or `-fragLength 38` for RNA.

Translational efficiency (TE) was calculated as relative RPKM of ribosome footprint density divided by relative RPKM of RNA abundance, with relative RPKM calculated as RPKM in each mutant divided by the average of RPKM in WTs grown in the same batch. Hierarchical clustering was performed using standard methods. Metagene analyses in Figures 2.7B and C show normalized ribosome occupancy averaged across all genes, aligned in a defined region from 100-nt upstream to 110-nt downstream of the first nucleotide of start codon. It was analyzed using a Python-based package Plastid (Dunn and Weissman, 2016), in which it quantified the counts of the nucleotide in a read corresponding to the first nucleotide of ribosomal p-site in the defined region of each gene. Within each gene, the counts of each position in the defined region were
normalized by the total number of counts in a normalization region, which is 70-100 nt downstream of the first nucleotide of start codon. Finally, the mean of normalized counts in each position across all genes was calculated.

**Codon occupancy analysis**

Global codon occupancy analysis was calculated as described in (Nedialkova and Leidel, 2015), with minor modifications. The P-site offsets were calculated by examining the cumulative distribution of 28-31 nt reads aligning at start codons using online package Plastid. After applying the respective offset to reads of each size, only in-frame reads were used. The first 15 and the last 5 codons of each transcript were removed from the reference. The frequency of each codon in ribosomal A, P, and E sites was calculated, and divided by the average frequency of the same codon in the three downstream codons from the A site for normalization.

**Analysis for RiboMeth-seq**

Indexed libraries were pooled and sequenced on an Illumina NextSeq500. De-multiplexed reads were mapped to *S. cerevisiae* rDNA using Bowtie2 v2.1.0 (Langmead and Salzberg, 2012) with parameters `-p 4–no-unal`. 5’ ends of uniquely mapping reads were quantified using bedtools (Quinlan and Hall, 2010) `genomecov` function with parameters `-d –5`. Methylation A scores were calculated as described in (Birkedal et al., 2015; Marchand et al., 2016).
A score of position $i = \max\{0, 1 - \frac{2n_i}{0.5|\mu_l - \sigma_l| + n_i + 0.5|\mu_r - \sigma_r|}\}$

where $n_i$ is the read counts at position $i$, $\mu_l$ and $\sigma_l$ are the mean and standard deviation of $\sum_{j=i-6}^{i-1} n_j$, respectively. $\mu_r$ and $\sigma_r$ are the mean and standard deviation of $\sum_{j=i+1}^{i+6} n_j$, respectively.

**Data Availability**

The accession number for all fastq files of deep-sequencing data is GEO: GSE100626. UCSC Genome Browser Tracks for all RPF and RNA data can be accessed at the following links:

http://genome.ucsc.edu/cgi-bin/hgTracks?hgS_doOtherUser=submit&hgS_otherUserName=chouhj&hgS_otherUserSessionName=11Del_2015batch

The first link consists of tracks for ribosome profiling and RNA-seq from 11 mutants: $elp1\Delta$, $elp3\Delta$, $kti11\Delta$, $urm1\Delta$, $trm4\Delta$, $trm7\Delta$, $trm732\Delta$, $rtr10\Delta$, $rit1\Delta$, $pus1\Delta$, and $pus7\Delta$.

http://genome.ucsc.edu/cgi-bin/hgTracks?hgS_doOtherUser=submit&hgS_otherUserName=chouhj&hgS_otherUserSessionName=46Del_2016batch
The second link consists of tracks for ribosome profiling and RNA-seq of another 46 mutants.

http://genome.ucsc.edu/cgi-bin/hgTracks?hgS_doOtherUser=submit&hgS_otherUserName=chouhj&hgS_otherUserSessionName=noCHX_2017

The third link consists of tracks for cycloheximide-free ribosome profiling of 5 mutants: *elp3Δ, mod5Δ, tan1Δ, trm7Δ*, and *tyw4Δ*. 
Acknowledgments

We thank A. Korostelev, A. Jacobson, and N. Krietenstein and other members of the Rando lab for discussions and critical reading of the manuscript. tRNA overexpression plasmids were a generous gift from S. Leidel and W. Gilbert. This work was funded by NIH grants R01HD080224 and R01GM079205 to O.J.R.
CHAPTER III: DISCUSSION

Covalent nucleotide modifications in tRNAs have been studied for decades, and many of their roles have been discovered, including in translation efficiency and fidelity, tRNA stability, and response to environmental changes. However, each study utilized a small number of mutants and dissected their functions using different approaches, preventing the direct comparison between functionally related mutants. A key feature of this work is the systematical comparison of 57 tRNA-modifying mutants within the same dataset, with the same experimental and analytical methods. It provides a valuable opportunity to constrain hypotheses for the mechanism underlying translational phenotypes of interest. This work thus offers an unprecedented resource for understanding the roles of tRNA modifications in translational regulation, as well as the functions of these tRNA-modifying enzymes and their cofactors in gene expression.

Overall, we found that the modifications present at the anticodon loop have the greatest effects on both codon- and gene-centric ribosome occupancy, consistent with the fact that the fundamental role of tRNAs is decoding. The modifications located distant from the anticodon also affect the dwell time of ribosomes at relevant codons, e.g., ac\textsuperscript{4}C12 modified by Tan1. In addition, we discovered a distinction between the tyw mutants in A and P site occupancy at relevant codons, implicating that yW intermediate product “yW-72” is sufficient for
decoding but the full yW modification is required for efficient peptidyl transfer or translocation.

In addition to the aspect of translational regulation, this work revealed several novel roles for various modifying enzymes and cofactors. First, we showed that the tRNA methylase Trm7 also contributes to 2’-O-methylation in a subset of 25S rRNAs. Second, we uncovered a surprising finding that the Elp complex and others involved in the mcm^5s^2U34 modification is implicated in control of the expression of TY1 retrotransposon and TY-associated genes, as their expression was decreased in these mutants. Finally, we identified a number of upstream open reading frames (uORFs) that translationally regulate the expression of their downstream genes in a similar way of the regulation of Gcn4 expression.

Together, the systematical study of the proteins involved in tRNA modifications provides a number of mechanistic hypothesis for future studies. In the following sections, I will broadly discuss four topics in details: 1) codon occupancy analysis, 2) mRNA abundance changes, 3) the role for Elp-related proteins in TY1 expression, and 4) translational regulation via uORFs.

**Codon occupancy analysis**

Changes in tRNA level and modification can affect the binding rate and affinity of cognate tRNAs to the ribosomes, and thus alter the dwell time of ribosomal A-site on the relevant codons during translation. This effect has been reported as a
change in ribosome occupancy at the codon level (Nedialkova and Leidel, 2015; Zinshteyn and Gilbert, 2013). Prior studies reported that five Elp-related mutants (elp3Δ, elp6Δ, ncs2Δ, ncs6Δ, and uba4Δ) lacking wobble uridine modifications exhibited higher ribosomal A-site occupancy on relevant codons AAA, CAA, and GAA (Nedialkova and Leidel, 2015; Zinshteyn and Gilbert, 2013). This work further extends this finding to all viable Elp-related mutants (15 in total), confirming the importance of this modification in decoding efficiency and ribosome pausing time.

In addition to recapitulating previous studies, this work identified many modifications that affect ribosome dwell time on known target codons (Figure 2.4, 2.5, and 2.6 in Chapter II), including ac4C12 (observed in tan1Δ), dU20 (dus2Δ), m2-G26 (trm1Δ), 2′-O-methylation at position 34 (trm7Δ and rtt10Δ), i6A37 (mod5Δ), yW37 (tyw mutants), and m7G46 (trm8Δ and trm82Δ). Some of these modifications are present at the anticodon loop, consistent with the widely accepted idea that defects in this area may have a great impact on translation efficiency. For instance, modifications at position 37 are known to stabilize the codon-anticodon interaction and enhance the binding affinity of tRNAs to their cognate codons in ribosomes (Konevega, 2004; Schweizer et al., 2017). Loss of these modifications may alter the base stacking around tRNA position 37 and disturb the interactions of the anticodon, codon, and 40S ribosome in the decoding center, resulting in altered dwell time of ribosomal A-site on cognate codons.
Interestingly, some identified modifications that affect ribosome dwell time locate distant from the anticodon. These modifications were generally considered for the role in tRNA stability—loss of these modifications may result in a reduction of mature tRNAs or may alter the tRNA structure (Alexandrov et al., 2006; Johansson and Bystrom, 2004; Pallan et al., 2008; Rider et al., 2009). Although our finding appears to link tRNA stability to the ribosome dwell time at particular codons, it still remains elusive in which mechanism these modifications affect the ribosome dwell time. For instance, lack of ac4C12 causes a decrease in A site occupancy but an increase in P site occupancy at serine and leucine codons, suggesting that this modification could potentially play some roles in peptidyl-tRNA positioning—enhancing translocation from the P to E site, or delaying translocation from the A to P site. tan1Δ was reported to exhibit decreased level of one specific leucine tRNA species (Johansson and Bystrom, 2004), which might contribute to the mutant effect on codon occupancy. However, it is unlikely to explain the changes in ribosome occupancy on all serine and leucine codons. It is also counterintuitive to the fact that lower tRNA abundance typically increases the dwell time of ribosomal A site. To investigate how these away-from-anticodon modifications affect the ribosome pausing time, two approaches may be considered. First, using tRNA sequencing to quantify the relative abundance of affected tRNAs in mutants could provide a direct correlation between the changes in ribosome occupancy vs. the tRNA levels. Second, utilizing *in vitro* assay that measures the binding and dissociation
constants of acetyl-tRNA to the ribosomal A site (Konevega, 2004) could potentially determine if the modification defects affect the ribosome binding ability and thus change the ribosome dwell time.

The A-site ribosome occupancy on a particular codon is widely considered as an estimate of the translation rate on that codon—the more ribosomes observed on the codon, the slower the translation rate is. Many studies have exploited ribosome profiling to discover the factors that correlate with the translation rate, including codon usage, codon-tRNA balance, and protein folding and aggregation (Ingolia et al., 2011; Nedialkova and Leidel, 2015; Qian et al., 2012). In our dataset, both increased and decreased ribosome occupancy were observed. Do they truly reflect the translation rate? It could be validated using two different approaches. The \textit{in vivo} translation rate could be determined by measuring the protein level of a reporter gene that is highly enriched with affected codons after an incubation of radioactive labeling amino acids (Cannarozzi et al., 2010). Alternatively, one could track the real-time luminescence signal from an \textit{in vitro} translation assay by mixing the cell-free lysates prepared from WT/mutants and Firefly luciferase mRNAs with specific codon manipulation. The time of the first appearance of luminescence signal was shown to reflect the translation rate (Yu et al., 2015).

Ribosome stalling on affected codons may potentially diminish the protein production from the genes enriched with those codons. In \textit{urm1}\Delta, proteins with the highest frequency (1% of the genome) of affected codons were significantly
decreased (Rezgui et al., 2013). Likewise, a high percentage of proteins enriched with relevant codons was significantly downregulated in \textit{trm9A} (Deng et al., 2015). As our dataset discovered many mutants with altered ribosome dwell time, an essential question is raised: is the protein level of certain genes disturbed in these mutants? This question may not be fully answered using the ribosome profiling data, as the mutants displayed no significant change in translational efficiency (average ribosome density normalized to RNA abundance in a given ORF) from the ORFs enriched with affected codons. In addition, further analysis revealed an extremely minimal correlation between the translational efficiency and the frequency of affected codons in ORFs, with the codon frequency explaining less than 2% of the variance in ribosome occupancy at gene level (Figure 2.14 in Chapter II). These observations imply that the reduction of protein levels might result from cellular processes that cannot be monitored by ribosome profiling. A recent study revealed that ribosome stalling on affected codons during translation may cause the misfolding of actively translating protein and protein aggregation (Nedialkova and Leidel, 2015). Accordingly, the change in the protein level in mutants may likely be due to post-translational regulation, such as protein misfolding and degradation. Overall, it will be of a great interest to determine the ribosome-pausing effect on protein abundance using quantitative proteomics such as liquid chromatography tandem-mass spectrometry (LC-MS/MS).
Elongation inhibitors such as cycloheximide (CHX) have been widely used in ribosome profiling to stabilize polysomes on mRNAs (Ingolia et al., 2009). However, a number of studies reported possible artifacts of adding cycloheximide prior to cell lysis, including ribosome accumulation at the start codon and ribosome depletion at the stop codon (Ingolia et al., 2011), and allowing slow elongation before cell lysis (Gerashchenko and Gladyshev, 2014; Hussmann et al., 2015). CHX treatment may distort the distribution of ribosomes on mRNAs and affect the codon-level analysis particularly. On the other hand, Nedialkova et al. consistently observed increased A-site occupancy on affected codons in elp-related mutants in both +CHX and -CHX conditions, with a smaller effect on -CHX samples (Nedialkova and Leidel, 2015), arguing that the ribosome pausing on cognate codons is probably not an artifact. To address this concern, we generated ribosome footprinting datasets for 6 strains with a CHX-free protocol. Briefly, we found overall negligible effects on gene-level ribosome occupancy, but we were unable to carry out detailed codon-level analysis. These libraries exhibited high variability in ribosome footprint length and an absence of ribosome accumulation at the start codon, which precluded A/P/E site assignment and the analysis of codon-level ribosome occupancy. Taken together, we could not rule out the possibility that codon-level analysis may be influenced by the use of cycloheximide, although many quality control and normalization steps were employed to improve the consistency and accuracy of the dataset. It might be
worthwhile to examine the mutant-effect on codon occupancy using an optimized CHX-free protocol for ribosome profiling.

**mRNA abundance changes**

This work revealed remarkably altered mRNA abundance in roughly half of the mutants investigated. The vast majority of these mutants are known to affect tRNA modifications at the anticodon loop, including \( \text{xm}^5\text{U34} \) (Elp complex and related proteins), \( t^6\text{A37} \) (KEOPS), Wy37 (Tyw3), \( i^8\text{A37} \) (Mod5), pseudouridine (Pus3 and Pus7), and 2’-O-methylation (Trm7). A large group of genes involved in amino acid biosynthesis and related metabolic pathways are upregulated in these mutants, primarily due to translational upregulation of the nutrient- and tRNA-responsive transcription factor \textit{GCN4} (Figure 2.11 in Chapter II). \textit{GCN4} translation is regulated by a well-studied mechanism in which four upstream open reading frames (uORFs) are involved (Hinnebusch, 2005). In the normal condition, the eIF2-GTP-Met-tRNA-iMet ternary complex (TC) binds to 40S subunits scanning on the mRNA and initiates translation on uORF1. After termination on uORF1, nearly all 40S subunits continue scanning downstream and rebind the TC before reaching uORFs 2, 3, or 4. They reinitiate at these uORFs and dissociate from the mRNA after termination, leaving \textit{GCN4} untranslated. In the condition of amino acid starvation, uncharged tRNAs trigger the activation of protein kinase Gcn2, which phosphorylates eIF2 and prevent the formation of TC. As the concentration of TC is low, only a portion of the 40S
subunits scanning downstream from uORF1 rebind the TC before reaching uORFs 2-4 and reinitiate translation at one of these uORFs. The remaining 40S subunits lack the TC when they reach uORFs 2-4, by pass the start codons at these uORFs, and rebind the TC before reaching the GCN4 AUG. It leads to reinitiate translation at GCN4 and an increased level of Gcn4 protein.

Then, how does GCN4 translation respond to tRNA modification defects? There are some hints from previous studies on different tRNA modification mutants. The translational upregulation of GCN4 in KEOPS mutants requires the uORFs, as mutations in four uORFs block the upregulation of GCN4 translation (Daugeron et al., 2011). Unlike in the starvation condition, GCN4 induction in Elp and related mutants (elp3∆ and ncs6∆) is GCN2-independent (Zinshteyn and Gilbert, 2013), ruling out a possibility that hypomodified tRNAs interact with Gcn2. Moreover, defects in tRNA 5’-end processing and nuclear export that lead to an accumulation of premature tRNAs in the nucleus induce GCN4 translation independently of GCN2 and phosphorylation of eIF2 (Qiu et al., 2000). These evidences prompt us to speculate that hypomodified tRNAs might accumulate in the nucleus and induce GCN4 translation. It could be tested by measuring overall or specific tRNA levels in the nucleus using tRNA-seq. As all of the mutants that exhibit translationally upregulated GCN4 have modification defects in specific tRNA species, for instance, lack of 2-thiolaiton at position 34 in only 3 tRNAs in Elp related mutants, it is likely that the accumulated tRNAs in the nucleus are specific as well.
A substantial number of genes are downregulated in distinct mutants. These genes can be clustered into several cellular functions, including iron transport, *PHO* regulation, *TY1* elements, and haploid-specific genes. It is unclear how tRNA modification defects cause mRNA abundance changes in these mutants. Nonetheless, two mechanisms that regulate the mRNA abundance can be involved—transcription and mRNA stability. Compared to transcription, more evidences showed that translation defects can alter the mRNA stability via different degradation mechanisms. For instance, codon optimality is known to be a major factor of mRNA stability (Presnyak et al., 2015). Codon composition correlates with mRNA stability, as a transcript containing a higher portion of optimal codons has a longer half-life. Substituting non-optimal codons in a given transcript with synonymous optimal codons decreases the mRNA decay rate, and vice versa. As some mutants in this work affect ribosome dwell time and/or tRNA levels, and given that the tRNA amounts and translational rate are two major determinants for codon optimality, transcripts enriched with affected codons might turn into “less optimal transcripts” and have lower mRNA stability. This could be tested by measuring mRNA decay rate in mutants vs. wildtype, and could be further confirmed in strains with exonuclease deletion background. Nonsense-mediated mRNA decay (NMD) might be one of the degradation mechanisms for these “less optimal transcripts”, as a set of NMD-regulated transcripts has lower average codon optimality and lower translational efficiency (Celik et al., 2017). It will be of great interest to test if decreased mRNA
abundance in the mutants is resulted from NMD by performing RNA-seq in double deletion strains that lose both specific tRNA modifications and one of NMD regulatory factors (Upf1, Upf2, and Upf3).

The role for Elp-related proteins in TY1 and TY-associated gene expression

tRNAs have been functionally linked to the biology of retrotransposons. The cDNA synthesis of retrotransposons requires specific tRNAs as primers (Marquet et al., 1995; Weiner and Maizels, 1987). In S. cerevisiae, the 3’ end of tRNA-iMet is complementary to the sequence immediately downstream of 5’ LTR of Ty1 retrotransposon RNA, and serves as the primer for reverse transcription (Curcio et al., 2015; Marquet et al., 1995). Beyond this conventional function, several recent studies implicated tRNA-derived fragments (tRFs) in control of retrotransposons and LTR-associated genes (Martinez et al., 2017; Schorn et al., 2017; Sharma et al., 2016). tRFs in the plant Arabidopsis thaliana are processed by Dicer-like 1 and incorporate into Argonaute1 (AGO1). The tRF-AGO1 complex can specifically target and cleave the mRNAs of endogenous transposable element (Martinez et al., 2017). In mouse stem cells, the 3’ tRFs inhibit the expression and reverse transcription of retrotransposon RNA (Schorn et al., 2017). In addition, 5’ tRFs from tRNA-glycine-GCC repress genes associated with the endogenous retroelement MERVL in mouse embryonic stem cells and embryos (Sharma et al., 2016).

In this work, we surprisingly discovered a decreased expression of TY1 retrotransposon and the endogenous genes that are located near intact TY, solo
LTRs, and tRNA genes in Elp-related mutants that lack mcm$_5^s^2$U34 modification. We also showed that this downregulation of TY1 expression is unlikely a result of Sir-dependent silencing, a secondary consequence of TYE7 downregulation, or Gcn4 regulation. This finding is of great interest given many links between tRNAs and retrotransposons. To investigate the roles for enzymes and cofactors involved in mcm$_5^s^2$U34 in the regulation of TY-associated gene expression, three hypotheses are proposed. First, Elp-related proteins may play roles in control of the transcription of these genes, possibly via the TY1LTR, given that LTRs can serve as an alternative promoter for their downstream genes. Second, Elp-related proteins may be involved in post-transcriptional regulation, such as mRNA stability. Third, the mutant effects on altered expression of TY-associated genes may be mediated by tRFs, given the links between tRNA modifications and tRNA cleavage (Schaefer et al., 2010).

To test the first possibility, we attempted to determine the Elp-responsive elements around the affected genes using promoter mutations in combination with the dual luciferase reporter assay. The genomic sequences of the TYE7 5’ and 3’UTR were cloned respectively into upstream and downstream of the Firefly luciferase (Fluc) reporter gene (Figure 3.1A). The normalized expression of Fluc in the plasmid p14, which contains an intact upstream TY1LTR and downstream tRNA-Met gene, was decreased in Elp-related mutants, resembling the downregulation of endogenous TYE7 from the RNA-seq data (Figure 3.1B). The preliminary result showed that mutating the tRNA-Met gene slightly ameliorated
the reduction of Fluc in elp3Δ and uba4Δ, while deletion of TY1LTR or segments of TYE7 5’ UTR had no effect (Figure 3.1B and C). It will be worthy to perform a complete promoter analysis with deletion of more segments of the remaining 5’ and 3’ UTRs. This will allow us to determine the elp-responsive sequences in the context of the reporter assay, or to rule out the possibility of the first aim.

To investigate the mRNA stability of endogenous TY-associated genes in the second hypothesis, the RNA abundance of specific genes in mutants and WT could be measured using quantitative RT-PCR or northern blot over time after treatment with transcription inhibitors. Finally, the third hypothesis can be initially tested using small RNA sequencing to quantify the abundance of tRNA fragments upon chemical treatments, which is required to induce the tRNA cleavage in S. cerevisiae.

Overall, this work unveiled a surprising role for Elp-related modifying proteins in TY1 and TY-associated gene expression, and raised an intriguing possibility to investigate the link between tRNAs and retrotransposons.
Figure 3.1. Dual luciferase assay to determine elp-responsive elements

(A) The plasmids designed for the dual luciferase assay. The CEN/ARS, LEU2+ plasmid harbors Firefly luciferase (Fluc; green) and Renilla luciferase (Rluc; blue) reporter genes for test and for reference, respectively. The 5' and 3' UTRs of genomic TYE7 were cloned upstream and downstream of the Fluc coding sequence, respectively, with various sequence bashings in different plasmids as indicated. The Rluc coding sequence was fused with *S. cerevisiae* TEF1 promoter and CYC1 terminator, which are selected deliberately to ensure a consistent Rluc expression between Elp-related mutants and WT. (B and C) Luciferase assays using various plasmids in WT and elp-related mutants as
indicated. X axis is the Fluc luminescence normalized to the reference Rluc luminescence, with the mutants’ values normalized to the WT’s.

**Translational regulation via uORFs**

The analysis of translational efficiency in this study uncovered several genes whose expression responds to a general tRNA modification defect instead of losing of specific modifications. Specifically, the translation on these mRNA was up- \((GCN4, SER3, \text{etc.})\) or down-regulated \((CPA1, CMR3, YGP1, \text{etc.})\) in a subset of mutants lacking deferent modifications. Gcn4, which is translationally upregulated in almost half of the mutants, is a nutrient-responsive transcription factor and its translation is regulated by a well-known mechanism involving ribosome reading through its four upstream open reading frames (uORFs) \((\text{Hinnebusch}, 2005)\). The upregulation of Gcn4 in elp-related and KEOPS mutants indeed requires the uORFs \((\text{Daugeron et al.}, 2011; \text{Zinshteyn and Gilbert}, 2013)\). In addition, \(CPA1\), which encodes a protein catalyzing the synthesis of an arginine precursor, also has an uORF that is translated into the arginine attenuator peptide (AAP) \((\text{Gaba et al.}, 2005; \text{Wang et al.}, 1999)\). The uORF hinders the translation of \(CPA1\) mRNA by stalling ribosomes at its own termination codon in response to arginine. We observed increased ribosomes at the 3’ of this uORF and a decrease in ribosome occupancy on \(CPA1\) mRNA.

Closer inspection of the locus of other affected genes revealed evidence for ribosome occupancy upstream of these genes. For example, \(SER3\), which
encodes a 3PG dehydrogenase that catalyzes serine and glycine biosynthesis, has an upstream transcript named \textit{SRG1} that represses \textit{SER3} mRNA expression (Martens 2004). \textit{SER3} transcript includes the \textit{SRG1} sequence as its 5’ UTR, based on the presence of long readthrough \textit{SRG1} transcripts extending to the \textit{SER3} 3’ end (Martens 2004), and the evidence of sequencing reads across the \textit{SRG1}/\textit{SER3} junction in this study. Our dataset revealed a remarkable decrease in ribosome occupancy at the 5’ of \textit{SRG1} and increased ribosomes on \textit{SER3}. It is likely that \textit{SER3} translation is regulated by \textit{SRG1} in a manner similar to the regulation of \textit{GCN4} translation. Other translationally downregulated genes might also resemble the uORF regulation mechanism of \textit{CPA1}, as they showed a comparable pattern of ribosome occupancy as \textit{CPA1}.

It deserves a future study to investigate if these translationally affected genes are regulated by their uORFs. 5’ RACE could be employed to validate the readthrough transcripts of uORFs and the downstream transcript. To determine if uORFs are necessary for downstream gene regulation, one may utilize a reporter assay with the gene in interest (for example, \textit{SER3}) fused with the reporter gene prior to the stop codon, and with three variations for the 5’ UTR—a full length of uORF, an uORF without the start codon, or an absence of uORF.
Bibliography


Presnyak, V., Alhusaini, N., Chen, Y.H., Martin, S., Morris, N., Kline, N., Olson, S.,

levels increase in response to several growth arrest conditions in Saccharomyces
cerevisiae. RNA 19, 243-256.

Qian, W., Yang, J.R., Pearson, N.M., Maclean, C., and Zhang, J. (2012). Balanced codon

(2016). Genome-wide cooperation by HAT Gcn5, remodeler SWI/SNF, and
chaperone Ydj1 in promoter nucleosome eviction and transcriptional activation.
Genome Res 26, 211-225.

(2000). Defects in tRNA processing and nuclear export induce GCN4 translation
independently of phosphorylation of the alpha subunit of eukaryotic translation

genomic features. Bioinformatics 26, 841-842.

disease syndrome protein, negatively regulates exocytosis independently of
transcriptional elongation. Mol Cell 17, 841-853.

Ramser, J., Winnepenningckx, B., Lenski, C., Errijgers, V., Platzer, M., Schwartz, C.E.,
gene FTSJ1 in Xp11.23 is associated with non-syndromic mental retardation in a

Rezgui, V.A., Tyagi, K., Ranjan, N., Konevega, A.L., Mittelstaet, J., Rodnina, M.V., Peter,
M., and Pedrioli, P.G. (2013). tRNA tKUUU, tQUUG, and tEUUC wobble position
modifications fine-tune protein translation by promoting ribosome A-site binding.
Proc Natl Acad Sci U S A 110, 12289-12294.

dihydouridine synthase 2 from yeast and the importance of modifications for

Rodriguez, V., Vasudevan, S., Noma, A., Carlson, B.A., Green, J.E., Suzuki, T., and
required for the biosynthesis of a highly modified Wybutosine (yW) base in

reveals features of normal and disease-associated mitochondrial translation. Nat
Commun 4, 2886.


