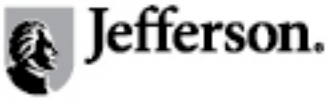

4-1-2015

Post-transcriptional modifications to tRNA--a response to the genetic code degeneracy.

Ya-Ming Hou
Department of Biochemistry and Molecular Biology, Thomas Jefferson University, ya-ming.hou@jefferson.edu

Department of Biochemistry and Molecular Biology, Thomas Jefferson University, howard.gamper@jefferson.edu

Wei Yang
Department of Chemistry and Biochemistry, Florida State University

Post-transcriptional modifications to tRNA—a response to the genetic code degeneracy

YA-MING HOU,[1] HOWARD GAMPER,[1] and WEI YANG[2]

[1]Department of Biochemistry and Molecular Biology, Thomas Jefferson University, Philadelphia, Pennsylvania 19107, USA
[2]Department of Chemistry and Biochemistry, Florida State University, Tallahassee, Florida 32306-4390, USA

We used to think of the genetic code degeneracy as an obligatory process resulting from assigning 61 sense codons to the 20 amino acids in protein. In this degeneracy, 18 of the 20 amino acids are endowed with multiple codons, and the conventional wisdom was that each of the 18 was translated equally from each codon. However, due to the degeneracy mostly manifest at the wobble position for each amino acid, we were led to believe that the degeneracy is simply to provide a deck of codons for safety control, in case one or two codons are damaged by nucleotide substitutions to the wobble position. However, work over a long path has now shown that the degeneracy has a much deeper meaning, with the ability to confer codon-by-codon translational regulation.

At its core, translation of a codon requires pairing interaction with the anticodon of its cognate tRNA. Due to the degeneracy, multiple codons are paired with one anticodon in different efficiencies and stabilities. This is a form of translational regulation executed at the level of individual codons. Nature has developed this translational regulation by two means: one by employing biased codon usage and the other by modifying the chemical structure of the tRNA anticodon or adjacent nucleotides via post-transcriptional modifications. While biased codon usage is genetically encoded, post-transcriptional modifications are epigenetic and have more flexibility to modulate the fidelity and efficiency of codon-anticodon pairing interactions. Indeed, among codons for the same amino acid, some require tRNA modifications that are stress-induced, conferring a capacity to regulate cellular adaptation to stress at individual codons. Thus, in response to the genetic code degeneracy, tRNA has evolved to carry post-transcriptional modifications to fine-tune translation one codon at a time. Enzymes catalyzing tRNA modifications are often linked to human diseases, ranging from metabolic defects, mitochondrial dysfunctions, and neurological disorders, to cancer.

The greatest diversity of tRNA modifications is at positions 34 (the wobble position) and 37 (on the 3′ side of the anticodon), two positions critical for translational accuracy and reading-frame maintenance. Three modifications at these positions have strong effects on codon-anticodon pairing interactions. First, U34 is modified to 5-methoxycarbonyl-methyl-2-thiouridine (mcm^5s^2U34) in $tRNA^{Lys/UUU}$ (anticodon 5′-UUU-3′), $tRNA^{Glu/UUC}$, $tRNA^{Gln/UUG}$, and $tRNA^{Arg/UCU}$. In the case of $tRNA^{Lys/UUU}$ for example, while it can read both Lys codons (5′-AAA and 5′-AAG), the codon-anticodon pairing in either case is weak, consisting of A-U and G-U base pairs. The addition of mcm^5s^2 to U34 can stabilize the A-U and G-U pairs at the wobble position by providing an extended chemical group to enhance the stacking interaction within each pair. The importance of this mcm^5s^2 modification is emphasized by its response to cellular stress. Specifically, mcm^5s^2U34 is synthesized in two sequential steps, first via the addition of mcm^5 to U34 by the *ELP* complex and second via the addition of s^2 to the mcm^5U34 intermediate by the *URM1* pathway. While the *ELP* complex is formally associated with the RNA pol II transcription complex in yeast, it catalyzes the mcm^5 addition to U34 upon cellular oxidative stress. Similarly, the *URM1* complex is activated by cellular stress in yeast, such as oxidative stress and temperature up-shift stress. Recent work shows that mcm^5s^2 addition to U34 facilitates translation of codon-biased mRNAs required for stress response. For example, upon stress with the alkylating agent methyl methane sulfonate, the methyl transferase Trm9 associated with the *ELP* complex is activated to help to translate DNA damage response genes each with a pattern of codon usage dependent on mcm^5s^2U34-tRNA for translation. Similarly, a human Trm9-like protein in cancer cells helps to translate the LIN9 protein, which has a codon bias requiring mcm^5s^2U34-tRNA for translation. LIN9 interacts with the tumor suppressor protein Rb to suppress cancer progression, thus implicating the Trm9-like protein and the mcm^5s^2 modification of U34-tRNA with a role in inhibition of tumor growth.

Corresponding author: ya-ming.hou@jefferson.edu

Article and publication date are at http://www.rnajournal.org/cgi/doi/10.1261/rna.049825.115. Freely available online through the *RNA* Open Access option.

A second example involves 5-methylation to cytosine (m^5C) of $tRNA^{Leu/CAA}$. Although m^5C is present in at least 34 yeast tRNA species, it is at the wobble position only in $tRNA^{Leu/CAA}$ as well as at position 48 in the V loop of the same tRNA. Upon cellular oxidative stress, however, the distribution of m^5C in the tRNA is dynamically changed, with a significant increase at the wobble position and a concomitant decrease at position 48, indicating a specific response to the stress. In yeast, the methyl transferase Trm4 is responsible for synthesis of m^5C34 to help decode the UUG codon. Without the modification, the codon-anticodon pairing interaction is weak and is dominated by A-U base pairs, whereas the presence of the modification provides a methyl group to enhance the stacking interaction at the wobble position. In yeast, the importance of m^5C34 for translation of the UUG codon is correlated with the highly biased distribution of UUG in genes involved in oxidative stress response. One of these genes is for the ribosomal protein Rpl22a, where all Leu residues are coded by UUG. Indeed, yeast cells exposed to oxidative stress preferentially synthesize Rpl22a, rather than its paralog Rpl22b, where only 34% of Leu residues are coded by UUG. The preferential synthesis of Rpl22a may confer adaptation to oxidative stress by several mechanisms, such as activating translation via the internal ribosomal entry site, regulating T-cell development, or engaging in the activity of the non-ribosomal complex of telomerase.

A third example is based on N^1-methylation to G37 (m^1G37) of $tRNA^{Pro/GGG}$. The cognate codon for this tRNA is CCC and the codon-anticodon pairing interaction is stable, consisting of three G-C base pairs. Nonetheless, without the modification, $tRNA^{Pro/GGG}$ is highly prone to +1-frameshifting, particularly on the slippery CC[C/U]-[C/U] sequences in mRNA. Taking the CCC-C mRNA sequence for example; the high propensity of the tRNA for making a +1-frameshift (+1FS) is in part due to its identical codon-anticodon pairing interactions in the un-shifted 0-frame and in the shifted +1-frame. Given that CC[C/U]-[C/U] sequences occur ~2500 times among protein-coding genes in *E. coli*, this poses a serious challenge for reading-frame maintenance. Unlike missense errors, +1FS errors are unforgiving; causing premature termination of protein synthesis and generating truncated products likely toxic to cells. The m^1G37 modification is known to suppress +1FS errors, although the mechanism has remained unresolved for a long time. Recent data from x-ray crystal structural work and advanced biochemical analysis have begun to shed light on the mechanism of suppression, showing that m^1G37 induces structural re-arrangement of the tRNA anticodon loop to restrict both the dynamics and kinetics of +1FS errors. Unexpectedly, while m^1G37 is conserved among all three isoacceptors of $tRNA^{Pro}$ (anticodons GGG, CGG, and UGG), its strength in suppressing +1FS errors varies; it is dominant and sufficient by itself in the UGG isoacceptor but is relatively weak in the GGG isoacceptor and requires the multi-faceted elongation factor EF-P for assistance. In the UGG isoacceptor, the wobble U34 is naturally modified to cmo^5U34 (uridine-5-oxyacetic acid), which enables strong pairing to CCA and CCG but weak pairing to CCC and CCU of Pro codons. The dominance of m^1G37 in the UGG isoacceptor may have arisen to prevent weak codon-anticodon pairs from +1-frameshifting.

The examples of mcm^5s^2U34, m^5C34, and m^1G37 illustrate that each post-transcriptional modification to tRNA is a response to the genetic code degeneracy and consequently the multiplicity of tRNAs for specific amino acid. In each case, the key question of how the modification stabilizes an otherwise unstable codon-anticodon pair remains to be addressed. The study of m^1G37 can provide a framework for moving forward; specifically, the role of m^1G37 was elucidated from analysis of each $tRNA^{Pro}$ in three states: (i) the G37-state made by in vitro transcription, lacking any post-transcriptional modification, (ii) the m^1G37-state made by in vitro transcription, followed by an enzyme-catalyzed synthesis of the modified base, and (iii) the native-state isolated from cells, containing all natural post-transcriptional modifications. These three states were compared for individual reaction steps of protein synthesis using purified ribosomes and translation factors. In parallel, crystal structures of m^1G37-tRNA on the ribosome were solved and genomic occurrences of the cognate CC[C/U]-[C/U] mRNAs were determined from bioinformatics. It is by combining data of these diverse approaches that an integrated view of how m^1G37 works begins to emerge. This example demonstrates the power of using complementary tools and methodologies to elucidate the mechanism and function of each tRNA modification.

Future work should focus more on tRNA modifications in multi-cellular organisms. While several modifications are conserved in all kingdoms of life, most show few phenotypes in single-cell organisms. For example, deficiency of mcm^5s^2U34 and m^5C34 in single-cell organisms has no phenotype in normal growth conditions, but manifests growth defects in cellular stress. The exception is m^1G37, whose deficiency is sufficient to arrest cell growth of both *E. coli* and yeast even without external stress. In contrast, deficiency of mcm^5s^2U34 and m^5C34 elicits diverse human diseases, indicating a stronger role of these modifications in human cells. For example, lack of the *ELP* complex for the first step synthesis of mcm^5s^2U34 results in severe human neurological defects, including the familial dysautonomia syndrome (causing patients to lose sensory neurons), intellectual disabilities, and non-familial amyotrophic lateral scelerosis. Similarly, deficiencies of m^5C34 in $tRNA^{Leu/CAA}$, due to pathogenic mutations in NSUN2 (the homolog of yeast Trm4), are associated with intellectual disabilities and a Dubowitz-like syndrome with mild microcephaly and congenital heart defects. Notably, in addition to these neurological disorders, deficiencies in tRNA modifications are linked to human diseases with metabolic defects, mitochondrial dysfunctions, and cancer. To date, very few human tRNA modifications are characterized; we anticipate that efforts directed

at elucidating human tRNA modifications at the enzymatic and mechanistic level will pave the way to understand how modifications are associated with human diseases and to gain insights into therapeutic interventions.

Acknowledgments

This work is supported by NIH R01GM108972 (to Y.-M.H.) and NIH R01GM111886 (to W.Y.).