The effects of codon bias and optimality on mRNA and protein regulation

Fabian Hia[1] and Osamu Takeuchi[1,*]

[1]Department of Medical Chemistry, Graduate School of Medicine, Kyoto University, Kyoto, Japan

***Corresponding author. Tel: +81 75 753 9500; Fax: +81 75 753 9502; E-mail: otake@mfour.med.kyoto-u.ac.jp**

Abstract

The central dogma of molecular biology entails that genetic information is transferred from nucleic acid to proteins. Notwithstanding retro-transcribing genetic elements, DNA is transcribed to RNA which in turn is translated into proteins. Recent advancements have shown that each stage is regulated to control protein abundances for a variety of essential physiological processes. In this regard, mRNA regulation is essential in fine-tuning or calibrating protein abundances. In this review, we would like to discuss one of several mRNA-intrinsic features of mRNA regulation that has been gaining traction of recent – codon bias and optimality. Specifically, we address the effects of codon bias with regard to codon optimality in several biological processes centred on translation, such as mRNA stability and protein folding among others. Finally, we examine how different organisms or cell types, through this system, are able to coordinate physiological pathways to respond to a variety of stress or growth conditions.

Introduction

The degeneracy of the genetic code entails that 61 codons encode 20 different amino acids. With the exception of methionine and tryptophan, all amino acids are encoded by synonymous codons. One of the pioneering studies of synonymous codons was published in 1972, in a paper exhibiting a method to calculate codon frequencies in yeast and seven bacteria [1]. From the computations performed, Goel and colleagues arrived at a conclusion that synonymous codons were not fully equivalent and alluded that these differences in codon frequency was likely due to codons conferring different rates of translation and therefore, a "selection pressure to maintain certain ratios among the synonymous codons" [1]. These findings were further supplemented by analyses of part of the *Escherichia coli* (*E. coli*) chromosome which showed that the frequencies of synonymous codons were non-random in coding sequences [2]. This systematic bias in codon frequencies in organisms would come to be known as 'codon (usage) bias'. In 1980, Grantham and colleagues proposed the *Genome Hypothesis* which states that synonymous codons are used at different frequencies by different genomes, and that the usage remains constant for all genes within each genome [3]. In other words, every organism utilizes its own system of synonymous codons.

Because synonymous codons are decoded at different rates, codons can be briefly classified into two categories, optimal and non-optimal. Broadly speaking, optimal codons are decoded faster and more efficiently than their non-optimal counterparts, impacting the translation efficiency of transcripts [4-6]. As translation efficiency is modulated by the codon composition of the transcript, transcripts enriched with optimal codons are translated more efficiently than those enriched with non-optimal codons.

For this review, we would like to define codon optimality, in a general sense, as how efficiently a codon is translated by the ribosome. This sets codon usage bias apart from codon optimality in the sense that the former refers to the composition of codon frequencies in coding regions for the genome while the latter is a measure for how efficiently a codon is translated. As we will further elaborate in the proceeding sections, codon optimality has been shown to influence important biochemical processes such as translation initiation, elongation and inevitably, processes surrounding it such as protein folding, among others.

Because these processes consequently impact cellular fitness it has been proposed that codons may have been subjected to selective pressure during evolution [7]. In accordance to this, several theories have been pursued, demonstrating that codons have been selected for accurate translation at important protein residues and longer genes [8-10], as well as for optimizing translation speed for fast-growing organisms [11, 12]. It is thus important to note a corollary that codon bias is to a certain extent influenced by codon optimality. Finally, in accordance with the *Genome Hypothesis*, this also entails that codon optimality is not universal in that different organisms utilize different sets of optimized and non-optimized codons.

To date, several exceptional reviews have been published [7, 13-26]. In this review, we discuss in the light of recent findings and advancements, how the usage of synonymous codons influences mRNA and protein regulatory steps, their effects and the implications in biochemical and physiological processes. While discussed individually at the level of nucleotide sequences to how the resulting protein is folded, these processes should be thought of as intricately intertwined dynamic networks which contribute to a final protein output. We first explain the measures by which codon bias is measured and how they have evolved over time. We then discuss the evidence to show the effects of codon usage on translation followed by a discussion on an enabling tool used to probe transcriptome-wide translation, ribosome profiling. Subsequently, we assess evidence which indicates that synonymous codon usage can exert an impact during translation or processes prior to translation such as transcription and splicing. We then examine how RNA modification of codons can affect translation kinetics. Finally, we discuss the physiological relevance of codon bias/optimality and how cells have harnessed this system to coordinate physiological responses.

Measures of codon bias and optimality

To scrutinize the effect of codon bias, several metrics have been proposed thus far. A classical metric, the codon adaptation index (cAI), proposed in 1987 by Sharp and Li, calculates the codon usage of a codon against a reference set of highly expressed genes [27]. Inspired by this index, another metric termed the tRNA adaptation index (tAI) was proposed by dos Reis and colleagues [28]. The tAI assumes that tRNA gene copy number in certain genomes such as *Escherichia coli*, *Saccharomyces cerevisiae* and *Caenorhabditis elegans* is highly correlated with tRNA abundance in the cell [29-31].

While the tAI is a useful indicator of codon bias, showing that different tRNA species have different affinities with its cognate codon, its predication that tRNA copy number reflects intracellular tRNA concentration may be limited to certain genomes [28]. Sabi and colleagues in 2014, proposed an approach to bridge this gap by introducing organism-specific adjusted tAI weights [32, 33]. Physiologically, tRNA abundances are dynamic, with translation rates depending on the balance between its "supply and demand", where tRNAs may be depleted depending on the frequency or demand of their cognate codons [34]. In order to accurately account for the cellular dynamics of tRNAs, Pechmann and Frydman in 2013, devised the normalized translational efficiency (nTE) scale that factored in the competition between cellular tRNA abundance and codon optimality [5]. Under this metric, codons could be designated optimal if the availability of cognate tRNAs is more abundant than their usage [5]. Additionally, several other estimators of codon translation speed, starvation indices as well as algorithms catered to calculating species-specific tAI have been conceived [33, 35, 36].

In 2015, Presnyak and colleagues laid out a metric which utilizes the Pearson correlation between codon frequencies and half-lives of each mRNA in yeast to derive an R-value known as the Codon occurrence to mRNA stability Correlation coefficient (CSC) [4]. Interestingly, comparison of the optimal and non-optimal assignments of codons based on the CSC with that of the tAI showed a high degree of similarity, with a good correlation between the respective values in yeast [4].

In 2016, Bazzini and colleagues in a paper investigating how codon identity and translation affected mRNA stability in zebrafish, *Xenopus*, mouse, and *Drosophila*, raised the possibility that strong amino acid bias in transcriptomes may have been a result of synonymous codons impacting mRNA stability [6]. As a measure of amino acid optimality, the authors utilized the Pearson correlation coefficient between mRNA half-life and the amino acids encoded in individual transcripts to derive an amino acid stabilization coefficient (ASC) [6]. Both CSC and ASC would then be used in a separate publication in 2019 to show that both codon and amino acid

identity were crucial in determining mRNA stability in human cells [37]. In several recent publications, a metric comparable to the ASC, the AASC, was defined by Narula and colleagues as well as Forrest and colleagues, to further investigate the effect of amino acids on mRNA stability in several half-life datasets [38, 39]. In concordance with the previous findings, the groups concluded that codon and amino acid content was associated with mRNA stability in human and mammalian cells respectively [38, 39].

Additionally, in our recent study in humans cells, codons with either a G or C at the third base position (GC3) was shown to be enriched in mRNA with longer half-lives while codons with either an A or U at the third base position (AU3) were shown to be associated with mRNA with shorter half-lives suggesting that GC3 and AU3 codons were optimal and non-optimal codons in humans respectively [40]. In this study, we showed that increased GC3-content entailed proportionately higher GC-content, suggesting that GC3- and GC-content could also be used as a tractable estimate of mRNA stability in human cells [40].

As the field of codon bias gains prominence, the development of metrics to reliably and easily quantify the variables associated with mRNA stability over a diversity of species is of vital importance. Given the inter-connectedness of translation as a complex process, further experiments are necessary to tease apart the intricacies of the system of codon bias and optimality.

Regulation of translation initiation and elongation

Translation efficiency can be determined by several parameters, such as the availability of cognate tRNA as well as the rate of initiation and elongation among others. The initiation step is regarded as the rate-limiting step of protein production [41-43]. In prokaryotes, this process is facilitated by the presence of the Shine Dalgarno sequence upstream of the start codon [44], while in eukaryotes, the process is guided by the Kozak sequence which encompasses the start codon [45]. In addition to this, the codon composition of the 5' end of the open reading frame (ORF) has been shown to be distinctly different from the rest of the ORF [21]. Interestingly, in bacteria, studies have shown that there are biases in codons usage at the translation start site [46-48].

Of interest, is the hypothesis that a 'ramp' sequence is present immediately after the start codon. It entails that translation is slow at the start of translation, increasing thereafter [49, 50] (**Fig. 1a**). Several purposes of the ramp have been proposed. Studies predominantly in bacteria propose that these reduce elongation speed, prevent 'ribosomal traffic jams' [50, 51] while others asserting that the ramps are an effect of selection for reduced secondary structure [46-48, 52]. Within the latter studies, results pertaining to the use of codon bias has been varied, with studies proposing that rare codons are selected for efficient translation [47], in contrast to studies which argue that codon bias does not have significant effects on mRNA or protein levels [48].

However, the 5' end of the coding region is home to several regulatory signals which can influence translation initiation and elongation. Tuller and colleagues in a review aimed to demystify these findings, describe in detail the positions and purposes of several known features and regulatory signals at the first 70 codons of the ORF [21]. Namely, the review in addition to others, describes sequences at the first ~10 codons conferring weak mRNA folding followed by a region of strong mRNA folding, suggesting that codon usage bias is in part directly selected for, as opposed to weak indirect selection [21]. Accordingly, Verma and colleagues show that the nucleotide positions 7-15 of the coding sequence (3-5 on the peptide) strongly affect the efficiency of translation [49]. Research by Bentele and colleagues show that the first 5-10 codons of the protein coding sequence in bacteria contain largely rare codons to reduce mRNA folding at the translation start [46]. Furthermore, Tuller and colleagues in their work demonstrate that further downstream from the initial 10 codons, at positions ~30-50 exists a region of low translation efficiency which has been selected for to reduce ribosomal collisions at the expense of slower translation initiation [50]. Additionally, the 5' end of the ORF was described to have a relatively weak adaptation to the tRNA pool, with the length of the ramp agreeing well with the length of translated polypeptide needed to fill the exit tunnel of the ribosome suggesting a role of the ramp in transiting from a slow initiation to a fast elongation phase [21]. Other signals such as the environment surrounding the start codon as well as amino acid exit tunnel interactions in addition to others can also be found within these 70 codons [21]. The authors caution that while certain of these signals may be universal, a portion of them may be

specific to or have yet to be investigated in certain organisms [21]. Indeed, with an abundance of regulatory features which are linked to translation initiation and even elongation, future studies need to consider the position-specific features and their consequences working in tandem with respect to translation initiation and elongation.

Early studies in the MS2 bacteriophage and *E. coli* genome showed a strong, non-random system of codon bias in highly expressed coding regions of their mRNA [53, 54]. Supplemented with information of studies about tRNA anti-codon sequences, Grosjean and colleagues showed that less abundant tRNAs were less utilized in highly expressed genes [53, 55]. The study went on to conclude that for efficient translation, certain degenerate codons were selected for based on their optimal codon-anticodon binding energies [53]. It was not until a study in 1984 by Pedersen that the direct effect of rare codons in *E. coli* under several growth conditions was measured [56]. The study concluded that protein synthesis rates were adversely affected due to the lack of cognate rare tRNAs when reporters harbouring rare codons were utilized [56] (**Fig. 1b**). This was one of several other early studies which would go on to show that codon bias usage and its associated tRNA pool could influence translation elongation [57, 58]. Advancements in genome engineering technologies enabled one particular study to recode genes in *E.coli* in order to investigate how codon usage bias influences translation [59]. Results from this investigation revealed that genome-scale perturbation of codons from frequent to rare codons resulted in an adverse effect in translation efficiency and therefore, cellular fitness [59]. In line accordance with previous hypotheses, increasing the supply for these codons resulted in a recovery of fitness, proving that translation is largely dependent on the tRNA pool and that a codon-to-tRNA balance is crucial in maintaining cellular viability [59]. Indeed, with advancements in deep sequencing and proteomics, it has also been shown that tRNA levels better correlate with codon usage in mice under fasting conditions [60]. Several models of investigating codon optimality or bias have been utilized to date. Besides the replacement of non-optimal codons with their optimal counterparts, heterologous protein expression models have also been employed in bacteria, yeast, zebrafish and human cells among others [4, 37, 38, 40, 61, 62]. Interestingly, in filamentous fungus *Neurospora crassa*, codon bias of genes has been postulated to evolve to efficiently transcribe and translate coding sequences [63, 64]. Yu and colleagues, in their work with a *Neurospora* cell-free system, presented several lines of evidence to demonstrate the effect of codon bias on translation elongation and consequently, protein structure [64]. Additionally, with ribosome profiling, the authors demonstrated that optimal and non-optimal codons influenced ribosome occupancy *in vitro* and *in vivo*, increasing and decreasing translation elongation respectively [64].

Other well established protocols such as polysome profiling have been used to measure ribosome density for individual mRNAs [65]. In studying the effect of codon bias in two proto-oncogenes, *KRAS* and *HRAS* with similar protein identity but differing levels of optimality, it was demonstrated that optimization of mRNA increases its levels in polysome fractions in human cells [66, 67]. These, together with other studies hint that codon optimization of coding sequences increases translation efficiency and consequently higher protein production [40, 66, 67] (**Fig. 1c**).

Indeed, the complexities and intricacies of translation are influenced by a multitude of factors. Riba and colleagues showed that in several yeast stains, translation elongation rates were influenced by amino acid and codon content as well as tRNA abundance to a similar extent [43]. Looking forward, the increased sensitivity of instruments coupled with increasing availability of transcriptomic and proteomic data will further allow us to tease apart the individual aspects of translation regarding codons usage and their downstream constituents such as amino acid content. In the next section, we would like to discuss one of the enabling technologies vital to understanding translation – ribosome profiling and its implications.

Ribosome profiling studies

The advancement of functional genomic tools has enabled an unprecedented and deeper understanding of the dynamics of translation. In particular, ribosome profiling, the capturing and sequencing of mRNA fragments protected by the translating ribosome, has enabled transcriptome-wide analysis of translation with nucleotide resolution [68, 69] (**Fig. 2a**). The premise was simple. Ribosomal velocity should be inversely related to

ribosomal density. If a system of codon optimality was present, translating ribosomes with their A-sites at non-optimal codons would possess a relatively longer dwell time compared to that of their counterparts at optimal codons (**Fig. 2b**). Surprisingly, initial ribosome profiling experiments in yeast, mouse embryonic stem cells and bacteria saw no correlations between codon usage and ribosome density or tRNA abundance when cells were treated with translation inhibitors cycloheximide (CHX) or chloramphenicol [70-73]. Several studies also proposed that positively charged amino acids were the major determinant of ribosomal pausing [70, 74]. Moreover, conflicting results were numerous, with several studies reporting significant correlations instead [75]. Additionally, it has been shown that ribosomal profiling studies can be biased by a variety of factors such as sequencing bias, coverage, experimental and analysis methodology as well as culture conditions [41, 74, 76-81]. Compounding to this, the use of different translational inhibitors may yield different ribosome footprint sizes with their distributions among codons uncorrelated between different inhibitors [82].

The use of CHX in particular, in ribosome profiling studies would prove to be a challenging one. To account for the discrepancy of the findings in ribosome profiling studies, several studies in yeast analysed unexpected patterns in the ribosome density of individual downstream codons and concluded that the inhibition of ribosome translocation by CHX was not immediate – elongation was able to proceed and ribosomes were inhibited in a gradual manner [79, 83]. Additionally, CHX has been shown to affect transcription of ribosome biogenesis genes in budding yeast under nutrient starvation, thereby affecting measures of translation efficiency [84]. To address the issue of ribosome arrest, alternative experimental protocols have been developed. 5PSeq, a method to identify 5' capped and phosphorylated RNA that are products of exo- or endonucleolytically cleaved RNAs has been proposed to be a drug-free alternative. In this approach, these degradation intermediates which follow a three-nucleotide periodicity are sequenced to give a transcriptome-wide view of ribosome dynamics [85, 86]. Flash-freezing of cells to allow the omission of CHX pre-treatment has also been widely adopted as another alternative [87-89]. With this protocol, studies in yeast have shown that non-optimal codons are indeed translated slower than their optimal counterparts [75, 76, 82]. Furthermore, to investigate the effect of codons on the before-mentioned translational ramp at the 5' end of the ORF, ribosome profiling has also been used to supplement experimental findings. However, conclusions from ribosome profiling experiments have been mixed, with studies claiming that no reduction in elongation speed could be observed and that any evidence of rare codon-mediated slowing was due to experimental artefacts or analysis issues [71, 90]. In this regard, Tuller and Zur in a 2015 publication, citing differences in analysis and normalization protocols argue that elongation speed was still slower in at the start of the ORF [21].

To date, several other methods developed from the ribosome profiling approach have been developed. Translation complex profile sequencing (TCP-seq) has been used to provide valuable *in vivo* evidence to support the scanning model of translation initiation [91, 92]. TCP-seq involves the snap-chilling of cells followed by the crosslinking of the translation complexes with their bound mRNA. RNA digestion of these complexes ensues to retrieve RNA footprints which are then separated by sedimentation velocity [91]. Finally, recovered RNA is subjected to high-throughput sequencing [91]. Importantly TCP-seq was able to capture, in addition to the 80S initiation complex, 40S ribosomal small subunit (SSU) footprints along the 5'UTRs and stop codons, providing valuable insights in the special and temporal sequence of events at these locations [91]. Another method, selective ribosome profiling (SeRP), uses an immunoprecipitation protocol to isolate specific factor-bound ribosome-nascent-chain complexes in via ribosome profiling to study the co-translational activity of specific protein maturation factors [93-96]. To date, SeRP has been successfully applied in both *E.coli* and yeast to investigate protein factor engagement to nascent polypeptide chains [93-96]. Given the engagement of chaperones to nascent polypeptide chains, investigating the maturation dynamics of proteins derived from codon optimized/deoptimized mRNA would surely provide insights to co-translational folding as discussed in the proceeding section.

As elongation speed along a transcript is never uniform, the pausing of ribosomes on certain sequences may result in ribosome arrest and collisions on the transcript. Collided ribosomes are collectively known as disomes and are subjected to mRNA surveillance processes known as ribosome-associated quality control (RQC) and No-Go-Decay (NGD) whereby the nascent peptide and mRNA are degraded respectively [97-99] (**Fig. 2c**). Where ribosome profiling has enabled the capturing of transcriptome-wide footprints, new analysis methods

have been developed to capitalize on ribosome profiling to analyze phenomena associated with ribosome arrest. One such method developed by Diament and colleagues to analyze ribosome queuing showed that at least one in five translating ribosomes in yeast is stalled [100]. Analysis to date shows that in yeast, CGA–CCG and CGA–CGA codons as well as poly(A) tracts induce stalling [98, 99, 101, 102] while a recent study in humans and zebrafish by Han and colleagues, showed that disomes occur at Pro-Pro/Gly/Asp, Arg-X-Lys E-P-A-sites as well as stop codons and 3'UTRs [103]. Interestingly, another recent study by Meydan and colleagues showed that in yeast, recognition of collided ribosomes do not always result in the nascent peptide and transcript undergoing RQC and NGD [104], suggesting another level of discernment that can further calibrate protein output.

At this point, it remains to be conclusively seen if codon optimality-mediated deceleration of ribosomes can be addressed through ribosome profiling. One might ask about the potential of a significantly decelerated ribosome resulting in a ribosome collision or, in such cases, how a cell is able to accurately discern the difference between both. With the increasing use of ribosome profiling, the continuous development and modification of ribosome profiling techniques will be essential in illuminating and bridging the gap between the transcriptome and proteome for both studies from basic biology to complex diseases [105-107]. However, at present, while ribosome profiling is a revolutionary and useful gauge of translation, careful consideration must be given in interpreting its conclusions in relation to its experimental parameters, analyses and limitations.

Regulation of mRNA stability

mRNA degradation is a complex process facilitated by many protein complexes (**Fig. 3a**). A multi-faceted participant in translation, codon optimality has also been implicated in regulating mRNA stability. For years, translation was known to be intricately coupled to mRNA degradation [108, 109] via several different decay pathways [110-113]. In particular, several experiments have shed light with regard to how translation is coupled to mRNA repression and degradation in yeast via proteins Dhh1p and Pat1p [114-117]. Specifically, during translation, Dhh1p is bound to slowly translocating ribosomes, engaging mRNA decapping and deadenylating factors such as Dcp1p, Lsm1p, Pat1p and Pop2p to induce mRNA decay [116, 117]. In addition, the 5' end of mRNA decay intermediates which follow a triple nucleotide periodicity have been identified suggesting that the RNA can be degraded from the 5' to 3' direction concurrent to translation [85, 86]. The effect of codon bias on mRNA stability has also been uncovered in several other model organisms such as *E. coli*, *Schizosaccharomyces pombe*, zebrafish, *xenopus*, trypanosomes and human cells [6, 37, 40, 62, 118-120]. It is also interesting to note that in a recent study in yeast, the Ccr4-Not complex was shown to be recruited to the ribosome via Not5 at regions of non-optimal codons when the ribosome A-site was vacant suggesting that Not5 acts as a guide of optimality-based degradation [121]. Additionally, a cryo-electron microscopy structure study by Tesina and colleagues show that in yeast, the mRNA exit site of the ribosome can be bound by exoribonuclease Xrn1, providing evidence that the Xrn1 can co-translationally interact with the translation machinery to degrade mRNAs [122].

Presnyak and colleagues with the formulation of the CSC, showed that mRNA half-lives correlated with optimal codon content in yeast [4]. The study further showed that mRNA degradation was linked to codon optimality via the regulation of ribosome elongation [4]. In a follow-up study, it was established that this effect was brought about by ribosome-bound Dhh1p acting as a sensor of ribosome velocity, associating with mRNAs containing predominantly non-optimal codons to facilitate their degradation [123] (**Fig. 3b**). However, a study by He and colleagues that studied decapping activators Pat1, Lsm1 and Dhh1, showed that the average codon optimality score of individual transcripts targeted by Dhh1 in their study did not agree with the observation in the previous study [124]. Furthermore, given that transcripts with low codon optimality scores were also targeted by Lsm1 and Pat1, independently of Dhh1, the authors suggested that other decay factors as well as the identity and distribution of non-optimal codons along a transcript may be factors that influence degradation[124]. Interestingly, in a separate study, the loss of the mammalian homolog of Dhh1, DDX6, did not result in a change in the correlation between mRNA stability and translational levels but instead, led to increased translation of microRNA targets [125]. Further analysis revealed that the stability of transcripts stabilized upon DDX6 loss were independent of codon optimality suggesting DDX6 does not link mRNA stability to codon optimality in

mammalian cells [125]. Separately, Courel and colleagues in their work in mammalian cells, showed that DDX6 targets were GC-rich mRNAs that were instead enriched in optimal codons rather than non-optimal ones [126]. Furthermore, the group reported that the GC-content of genes influences mRNA storage and decay in human cells with AU-rich and GC-rich mRNA degraded via different pathways; GC-rich mRNAs are degraded from the 5' end via XRN1 while AU-rich mRNAs are localized to P-bodies where they are targeted for degradation by PAT1B in a 3' to 5' manner [126]. The effects of GC- as well as GC3-content on mRNA expression and stability have also been investigated. In our work with human cells, increased GC3-content, mirrored by increased GC-content was shown to increase both mRNA stability and translation efficiency in contrast to AU-rich mRNA which exerted a converse effect [40] (**Fig. 3c**). Additionally it was shown that AU-rich discerning RBPs such as ILF2 and ILF3 can bind to the coding sequences of mRNAs to exert a destabilizing effect independent of translation [40].

In the same vein as codon bias, amino acid content has also been shown to be a determinant of mRNA stability [38, 39]. Investigations in yeast demonstrate that during translation elongation, decoding at the ribosome A-site determines mRNA degradation [127]. On the other hand, in higher eukaryotes such as zebrafish and *Xenopus*, the effect on mRNA stability is comprised of both codon and amino acid usage [6]. This discrepancy between the importance of codon- or amino acid-level usage on mRNA stability has been suggested to be a reflection of differences in the translation and/or degradation machinery between organisms [127]. Indeed, uncoupling the effect of amino acid usage bias from codon bias on mRNA stability is complicated even within studies in human cells [37-39]. While these recent findings show that amino acids could be defined as optimal or non-optimal, factors such as the hydropathy of amino acids have also been proposed [128]. Further investigations would be required to completely uncouple the effects of codon and amino acid optimality on mRNA stability.

While is it the prevailing consensus that codon bias affects translation, several models of how ribosome speed and therefore density, affects mRNA stability. Some models posit that the competition between the decay and translation initiation machinery in addition to increased ribosome density, excludes mRNA decay factors, consequently increasing mRNA stability [129-133] (**Fig. 3d**). On the other hand, other studies propose that the effect of codon optimality on mRNA stability is the major determinant during the translation elongation phase, utilizing ribosomes as sensors to recruit mRNA decay factors [4, 114, 117, 123, 127]. While not entirely mutually exclusive, it would be interesting to pry apart and assess the contributions of each of the above-mentioned factors in different model organisms.

Co-translational protein folding

Alongside discoveries of how synonymous codons affected translation efficiency, it was proposed that gene sequences had, by extension, evolved to control the kinetics of translation at defined parts of proteins to allow efficient folding [134, 135]. Early studies in *E. coli* demonstrated that in several multi-domain proteins, codons which slowed translation were enriched in highly ordered regions in the polypeptide chain suggesting a form of co-translational control of protein folding [135]. Follow-up investigations revealed that the specific topology of proteins where correlated to codon usage; α-helices by optimal codons, β-strands and coils by non-optimal codons [136, 137]. In one study, analysis of several genomes including *E. coli*, yeast, *Drosophila*, and mice revealed that translational fidelity is critical at structurally sensitive sites; optimal codons are enriched in buried residues, in addition to regions whereby mutations would have resulted in significant changes in folding energy [9] (**Fig. 4a**). Additionally, conserved rare codon clusters, demonstrated to improve protein folding, have been found to be predominantly enriched in membrane-related proteins involved in targeting, insertion or, interaction with other proteins [138, 139]. Using the nTE, Pechmann and Frydman revealed that in yeast, non-optimal codon clusters are associated with α-helices while β-sheets are associated only with optimal codons [5]. Additionally, they showed that hydrophobic regions, critical for protein folding and aggregation, contain mainly optimal codons suggesting a need for accurate error-free translation in these regions [5]. Kinetic modelling of the translation of *E. coli* proteins by Bitran and colleagues revealed that the usage of rare codons in the translation of intermediate folding regions of large proteins improves co-translational folding efficiency by allowing nascent chains more time to fold into their proper conformations [140]. Investigations in animal

systems such as *Drosophila* cells have also shown that structure based-codon manipulation affects protein structure and function [141].

Interestingly, it was noted the type of synonymous codons used in controlling protein folding was different for various species, reflecting differences in translational mechanisms [142, 143], in accordance to Grantham's hypothesis [3]. Indeed, work with heterologous expression systems has been shown to yield inactive proteins [144, 145]. Since then, such correlations have been demonstrated in various experimental systems with various genes [146-148]. Critically, the fact that organisms had evolved their own codon usage systems to control translation hints at the importance of an optimized system in producing functionally intact proteins. Early studies which involve replacement of rare codons with frequent ones demonstrate a negative impact on the synthesized protein's structure and thus activity [147] (**Fig. 4b**). Additionally, perturbations to the tRNA pool by increasing the abundance of rare codons in *E. coli* has been shown to generate misfolded proteins [149]. These findings have also been reflected in the *Neurospora* cell-free system as well as *in vivo* via measurements of optimized and non-optimized luciferase reporters; optimization of non-optimized regions resulted in a decrease in luciferase activity [64, 150]. Additionally, these optimized variants were more resistant to partial trypsin digestion than the wildtype ones [64]. Furthermore, investigations into mammalian gamma-B crystalline proteins have shown that proteins translated from corresponding synonymous codon variants also possessed different conformations when investigated by 2D nuclear magnetic resonance as well as protease resistance assays [151].

Conversely, an alternative explanation accounting for protein folding besides codon bias or optimality has been proposed. In a study in yeast and mouse embryonic stem cells, Yang and colleagues argue that increased protein production relies on a trade-off of translational accuracy for elongation speed and that the only way to improve translational fidelity is to decrease the speed of translation [152]. Importantly, the authors demonstrate in yeast that strong mRNA structures are used to slow elongation in order to improve translational fidelity and consequently protein structure of highly expressed mRNAs [152]. However, a study in multiple genomes explains that synonymous substitutions appear to be selected for in maintaining stable RNA structures which in turn can influence translation [153] suggesting that these findings can to a certain extent be reconciled with the use of synonymous codons.

While studies on co-translational folding are pervasive, it should be noted that the complexity behind protein folding is manifold. While in general, perturbations or substitutions to synonymous codons have yielded fascinating insights to how proteins are co-translationally folded, these results may vary depending on the algorithms, organism, reporter system that was employed as well as the sensitivity of detection of the folding signals. Even within bacteria, the strength of codon usage bias was found to be varied among species [154]. As mentioned before, results from heterologous expression systems should be interpreted in the light of differing tRNA pool requirements. Finally, given the involvement of protein chaperones on folding [155], it would be appealing to investigate how chaperones are involved in the translation of optimized and non-optimized sequences.

Regulation of transcription and splicing

Given its multiplicity of effects towards influencing protein expression, codon optimality has unsurprisingly been shown to exert its effects even at the level of transcription. By the 1990s, optimality modelling based on the simplifying assumption that natural selection favours increased transcription of genes bearing preferred codons had already been made [156]. Indeed, subsequent studies in this field have yielded concrete evidence implicating codon optimality in altering transcription rates. As part of a study aimed at analyzing the connection between sequence features of translation to transcription elongation, Cohen and colleagues showed that codon usage bias significantly affects both transcription and translation in highly expressed genes in yeast [157]. Zhou and colleagues in a study utilizing the *Neurospora* cell-free system, showed that the effects of codon optimization were also positively exerted at the level of mRNA transcription [63]. Conversely, it was shown that histone H3 lysine 9 (H3K9) trimethylation was responsible for the transcriptional silencing of genes containing predominantly non-optimal codons [63] (**Fig. 5a**). Kudla and colleagues, aiming to uncover the effects of

synonymous silent-site GC substitutions on gene expression in mammalian cells, demonstrated that GC-rich reporters had increased mRNA abundances compared to their GC-poor counterparts [158]. Interestingly, the study also reported that mRNA degradation was not responsible for this difference and attributed the increase to either increased transcription or mRNA processing [158]. Evidence from our work with human cells demonstrated that GC-rich reporters derived from REL and IL6 was more abundant compared to their AU-rich counterparts in HEK293 cells suggesting increased transcription of GC-rich transcripts [40].

Apart from reporters, closely related genes have been employed. Newman and colleagues in their experiments show that differences in codon bias in closely related nucleic acid-sensing receptors TLR7 and TLR9 which contain predominantly rare and common codons respectively, possess different levels of expression [159]. The authors then demonstrate that codon bias is correlated to GC-content and that suboptimal codon bias related to low GC-content limits the expression of TLR7 with respect to TLR9 [159]. Furthermore, proto-oncogenes *KRAS* and *HRAS*, genes with similar amino acid identity but differing levels of optimality were demonstrated to be differentially regulated at the level of translation and, transcription via histone modification and chromatin structure [67].

Corroborating these studies, a recent study by Mordstein and colleagues demonstrated that through genome-wide analysis and optimization of codons, high GC-content increased mRNA and protein abundance, as well as cytoplasmic localization [160]. Interestingly however, the authors also showed that splicing increases the expression of AT-rich genes via increased cytoplasmic localization; an effect not observed for GC-rich genes [160] (**Fig. 5b**). In another study, Fontrodona and colleagues demonstrate that in humans, exons coregulated by splicing factors possess similar nucleotide composition bias in that codons encode amino acids with similar physicochemical properties [161]. For example, G/C-rich motif-binding SRSF2 promotes the inclusion of GC-rich exons which code preferentially for small amino acids while C-rich motif-binding SRSF3 promotes the inclusion of GC-rich exons which code preferentially for uncharged amino acids [161]. These studies are extremely relevant in that they demonstrate that nucleotide sequence bias (and consequent codon bias) not only has an impact on the production of a functional gene, but is also implicated in the localization and splicing of transcripts while maintaining their physiochemical protein features.

Unsurprisingly, Stergachis and colleagues show that codon preferences in mammalian genomes can also be accounted for by transcription factor binding [162]. In their study, codons which specify both amino acids and exonic transcription factor recognition sites, termed ‘duons’, are evolutionarily constrained by the need to preserve transcription factor recognition sequences [162]. Importantly, with data obtained from genome-wide associated studies (GWAS), the authors show that a large proportion of disease- and trait-associated duons are associated with non-synonymous substitutions suggesting that these variants may have an impact on regulation and/or protein function [162]**.**

RNA modifications in tRNA and protein-coding RNA

Given the increasing body of evidence demonstrating how tRNA pools and codon usage regulate translation, one would call into question how organisms, which do not possess a canonical repertoire of all 61 tRNA species (possessing generally 23 to 45 tRNA species), are able to circumvent this limitation in resources. Francis Crick, in 1966, in explaining the nature of the genetic code’s degeneracy, proposed the *Wobble Hypothesis* [163]. The hypothesis states that while regular base-pairing could occur between the first and second positions of the codon with the corresponding third and second bases of the anticodon, the third base of the codon could form a non-Watson-Crick base pairing with the first base of the anticodon [163]. Where many organisms do not possess all 61 tRNA species, wobbling compensates this limitation by conferring broad specificity, allowing one tRNA molecule to be read by multiple codons [164]. Wobble bases have been shown to affect translation kinetics although the impact and extent to are still uncertain, with several studies reporting mixed results [76, 165, 166]. In this section we discuss how organisms utilize wobbling to adjust and optimize translation kinetics by modifying the affinity of tRNA and mRNA to translation. It should be noted however, that the codon optimality in the following and related studies may not follow the static delineation of optimized/non-optimized codons. Instead these codons in defined sets of transcripts are specially selected for to be optimal only under specific

conditions. Nonetheless, it would be beneficial to take into account the findings of these and future studies when investigating the system of codon bias under different physiological conditions.

tRNA can be subjected to a multitude of RNA modifications. Indeed, more than a hundred tRNA modifications have been identified [167, 168]. From this, more than 700 modified RNA sequences have been identified, a vast majority of which can be found on tRNA [168]. Wobble base modifications in particular have been shown to facilitate the recognition between codons and wobble-read tRNA globally [169-171] (**Fig. 6a**). This phenomenon has been researched in several organisms under various growth conditions. In yeast, under oxidative stress, Trm4-catalyzed modification of C at the wobble base of $tRNA^{Leu(CAA)}$ has been demonstrated to increase translation of a certain set of mRNAs in which the majority of encoded leucines are biased towards the codon UUG [172]. This phenomena is similarly echoed in *mycobacterium bovis* under hypoxic conditions in which increases in wobble cmo5U in $tRNA^{Thr(UGU)}$ increases the translation of transcripts containing the ACG codon [173]. The loss of wobble uridine (U34) in *Caenorhabditis elegans* and yeast has been shown to result in translational pausing of ribosomes leading up to the accumulation of protein aggregates, the inability to clear these aggregates and eventual proteotoxic stress [174]. A recent study by Bornelöv and colleagues in human embryonic stem cells revealed that codon optimality of self-renewing and differentiating cells is based on the GC-content of differentially expressed transcripts [175]. In particular, self-renewing stems cells which possess high levels of inosine are dependent on inosine tRNA modifications which generally increase translation efficiency of modified transcripts [175].

Apart from those in wobble tRNA, modifications which affect translation efficiency can be found on protein-coding RNA itself; translation of these codons results in increased translation efficiency and mRNA stability (**Fig. 6B**). Arango and colleagues describe the acetylation of wobble base cytidine to N4-acetylcytidine (ac4C) as an mRNA modification which increases translation efficiency and mRNA stability in human cells [176]. Additionally, Eyler and colleagues utilizing a bacterial translation system and human cells, demonstrated that pseudouridine, a common RNA modification negatively impacts translation by altering the interaction between the ribosome and cognate as well as non/near cognate amino-acylated tRNAs [177]. Additionally, the authors demonstrated that the inclusion of pseudouridine-containing codons in a single type of mRNA resulted in the translation of a variety of peptide products suggesting an alteration in tRNA selection by the ribosome [177]. Another common RNA modification, m6A has been investigated with regard to translation due to its presence in coding sequences yielding mixed results. Research by Mao and colleagues show that inclusion of m6A in protein coding transcripts results in ribosomal pausing, while the removal of m6A modifications further decreased translation [178]. This phenomena was attributed to m6A possessing dual functions; eliciting ribosome stalling and, resolving mRNA structures to facilitate translation via YTHDC2, a RNA helicase m6A reader [178]. These findings were to a certain extent contrasted by an earlier publication which concluded that the majority of m6A in protein-coding regions were non-functional with little conservation in both yeast and humans [179]. However in the latter research, the authors also concluded that the remaining minority of evolutionary conserved m6A modifications were suggested to be functional in nature and should be subjected to future investigations [179]. While pseudouridine and m6A are not specifically wobble base modifications, it would be of interest to investigate how their presence in protein-coding sequences ties into the translation of AU-rich mRNA.

Functional importance and physiological relevance

Codon optimality affords organisms a plethora of ways to regulate protein abundance and functionality. This system has been shown to couple the translation response to an organism's state, stress response and ultimately, adaptation to its environment. While briefly touched upon in the previous sections, we would like to further elaborate on the specific physiological merits conferred to various organisms through codon bias and optimality in the following section. It is therefore important to recall that the *Genome Hypothesis* holds true, in that various organisms utilize a unique repertoire of codons in their own systems of codon optimality,

Under the assumption that tRNA isoacceptor abundances are correlated with frequency of their cognate synonymous codons, a study which modelled the charging states of *E. coli* tRNA isoacceptors revealed that

tRNA isoacceptors are selectively charged during amino acid starvation [36]. In this and further validation experiments, rare codons which are read by less abundant tRNAs were demonstrated to preserve high charging levels for their cognate amino acids under starvation, allowing them to be relatively efficiently translated compared to their frequent codon counterparts hinting at the presence of gene subsets which may be selectively translated under nutrient liming conditions [36, 180] (**Fig. 7a**). A study of yeast and *C. elegans* datasets by Gingold and colleagues also revealed that codons read by rare tRNAs are enriched under various stress responses [181] (**Fig 7b**). Furthermore, based on modified tAI, it has been shown that in yeast, genes with codons that were better adapted to the tRNA pool under stress conditions tended to be associated with functions related to responses to external stimuli [182]. Conversely, genes whose codon adaptation was low, were associated with functions such as amino acid biosynthesis and carbohydrate metabolism suggesting that cells can dynamically regulate their tRNA abundance to produce a myriad of responses to cope with stress [182].

Adaptation to varying tRNA abundances via synonymous codons has also been demonstrated in eukaryotes as a method to control the cell cycle. A study by Frenkel-Morgenstern and colleagues demonstrated the effects of a changing charged tRNA pool during various stages of the cell cycle [183]. Importantly, in the G2 phase, several tRNA synthetases were found to be increased towards the G2/M phase of the cell where charged tRNA pools were at the highest [183]. Correspondingly, mRNA with non-optimal codons were efficiently translated, mirroring this increase in charged tRNA pools [183]. In contrast, in the G1 phase, where charged tRNA was limited, mRNA with optimal codons were selectively translated due to a higher affinity of optimal codons to their corresponding tRNA isoacceptors [183]. Further evidence of such phase-specific control has also been demonstrated in yeast through work by Sabi and colleagues, in which translation elongation efficiency was shown to be changed at various phases of yeast sporulation to ensure the selective translation of important phase-specific proteins [184]. Additionally, research in yeast revealed that genes with similar functions have similar codon compositions allowing them to be regulated at similar levels for synchronous expression [4]. For example, genes involved in glycolysis, or encoding ribosome subunits were enriched in optimal codons while genes encoding pheromone responses and small-subunit processome were enriched in non-optimal codons, possibly as a reflection of their physiological requirements [4]. Intriguingly, tRNA gene sequences in eukaryotes are highly diverse. Goodenbour and colleagues show that tRNA genes which harbour the same anticodon but different sequences elsewhere in the tRNA (tRNA isodecoder genes) are diverse in mammalian genome [185]. Interestingly, isodecoders have been shown to vary in expression in mammalian tissues, with each possessing varying degrees of translational efficiency [185-187]. While these studies may not have been specifically discussed in the light of codon bias and optimality, it would be no surprise to surmise that individual species of tRNA isodecoders may have a limiting impact on the overall decoding of transcripts in different tissues and may present as an extra level of translation regulation.

Interestingly, studies have shown that codon optimization of the *neurospora* FRQ protein, a protein predominantly comprised of non-optimal codons, resulted in the disruption of *neurospora* circadian rhythm, hinting at a physiological role for non-optimal codons [188]. In cyanobacterium *Synechococcus elongate*, while optimization of circadian clock genes *kaiBC*, resulted in enhanced rhythmicity at physiologically permissive temperatures, cellular fitness was compromised at cooler temperatures, indicating a form of selection against the usage of optimal genes; only permitting growth at physiologically permissive temperatures [189].

Saikia and colleagues reported that in mammalian cells, the translation of ubiquitin–proteasome pathway-associated genes was resistant to the effects of amino acid starvation [190]. Analysis of this subset of genes showed that were enriched in non-optimal codons [190]. The authors hypothesized that following amino acid limiting conditions, the functioning of this pathway is vital in ensuring that amino acids are recycled to help the cell adapt to the environment [190]. In comparison, a recent study in mouse embryonic fibroblasts found that mRNAs related to cell proliferation were more strongly translated in rapidly dividing cells compared to cells at a resting state despite individual tRNA expression remaining unchanged [191]. The authors attributed this phenomenon instead, to a global upregulation of tRNAs as compared to individual regulation of tRNAs [191].

The effects of codon optimality are also evident in developmental processes. In maternal-to-zygotic transition of zebrafish, *Xenopus*, *Drosophila* and mice, codon optimality is essential for the clearance of maternal mRNA

during early embryogenesis [6, 62]. Strikingly, the attenuation of the codon optimality-based system has been shown to be important in the neural development of drosophila [192]. It has been hypothesized that this attenuation is necessary to allow neural-specific development programs of mRNA degradation to prevail [192].

Additionally, the role of RNA modifications in stress responses has come under increasing scrutiny of late. Trm4's before-mentioned role in cell survival, increases selective translation of mRNA enriched in the TTG codon, allowing yeast to mount a survival response to an oxidative environment by increasing the expression of ribosomes [172]. Additional high-throughput screens in yeast have demonstrated that another methyltransferase, Trm9 is upregulated in response to the presence of alkylating agents [169, 193] and ionizing radiation [194] by introducing wobble base modifications of several tRNA ($tRNA^{UCU}$ and $tRNA^{UUC}$ among others) [195]. Consequently, AGA and GAA codon-enriched mRNAs, associated with protein synthesis and DNA damage responses were translated more efficiently [196]. Expectedly, trm9Δ cells accrue a DNA damage phenotype suggesting that the modified wobble bases were essential in mounting a survival response [196]. Indeed, these results hint that RNA modifications under a variety of stress conditions allow cells to reprogram tRNA for codon-biased translation of subsets of mRNA to mount a survival response.

Conclusion

Codon optimality has been demonstrated to be a strong multi-faceted force in shaping protein production via a myriad of biochemical processes (**Fig. 8**), consequently impacting cellular fitness. Once again, we would like to stress that while this review explores the effects of codon bias and optimality in separate sections, these processes should be thought of as a fluid network which can be calibrated by a myriad of inputs such as environmental stress and developmental programs among others, to help the cell regulate a functional protein output. Considering the effects of synonymous codon usage even at the level of transcription and splicing, the genetic code in the context of our discussion is a superimposition of several layers of regulatory information. This however increases the complexity in teasing apart the effects of each of the above-mentioned process from each other.

Insofar as transcription and post-transcriptional processes such as splicing or localization are concerned, we have discussed how nucleotide sequence (under constraints by codon and amino acid identity) has an impact on transcript expression. While beyond the scope of this discussion, trans-acting factors such as splicing and export factors as well as base modifying enzymes, by extension can be considered to be part of a multi-layered system designed to tune transcript and consequently, protein levels.

This review has also discussed mRNA degradation and protein folding as co-translational processes contributing to a functional protein output. Besides these, cis-acting elements at the 5' end of the coding sequence demonstrate that translation initiation can have a profound impact on a transcript's translation efficiency and stability. Additionally, the discovery that Dhh1may act as a sensor for monitoring translation elongation and controlling mRNA degradation in yeast demonstrates that ribosome dynamics are crucial. It is therefore plausible that other similar sensing factors exist, particularly in mammals that can link mRNA stability to translation. Given that non-optimal codons have the potential to significantly slow ribosomes, there is a potential for ribosome collisions to occur under suitable conditions. The discovery of disomes (and trisomes) through emerging technologies such as ribosome profiling may thus help determine if a link exists between non-optimal codons and ribosome collisions as well as further shed light on how cells differentiate and resolve these roadblocks.

While we emphasize that optimal and non-optimal codons have been selected for to optimize protein production, many studies have shown that organisms or cell types are able to further alter or bypass this system under various growth or stress conditions. In this sense, codon optimality can also be thought of as dynamic or fluid – different conditions necessitating different 'optimal' or 'non-optimal codons' to mount a suitable physiological response. Finally, it would be interesting to consider how these insights, gleaned from how mRNA and protein levels are calibrated in cells, can be applied at an industrial level; knowledge of the transcription and translation

machineries of cells can be harnessed to produce large quantities of functional recombinant proteins efficiently and effectively.

Compliance with ethical standards

Conflict of interest. The authors have no conflicts of interest to report.

Acknowledgments. The authors thank all members of our laboratory for discussions. This work is supported by Japan Society for the Promotion of Science (JSPS) KAKENHI Grant Numbers JP18H05278 and 20F20115, and Takeda Science Foundation.

References

1. Goel, N.S., et al., *A method for calculating codon frequencies in DNA.* J Theor Biol, 1972. **35**(3): p. 399-457.
2. Post, L.E., et al., *Nucleotide sequence of the ribosomal protein gene cluster adjacent to the gene for RNA polymerase subunit beta in Escherichia coli.* Proc Natl Acad Sci U S A, 1979. **76**(4): p. 1697-701.
3. Grantham, R., et al., *Codon catalog usage and the genome hypothesis.* Nucleic Acids Res, 1980. **8**(1): p. r49-62.
4. Presnyak, V., et al., *Codon optimality is a major determinant of mRNA stability.* Cell, 2015. **160**(6): p. 1111-24.
5. Pechmann, S. and J. Frydman, *Evolutionary conservation of codon optimality reveals hidden signatures of cotranslational folding.* Nat Struct Mol Biol, 2013. **20**(2): p. 237-43.
6. Bazzini, A.A., et al., *Codon identity regulates mRNA stability and translation efficiency during the maternal-to-zygotic transition*, in *EMBO J*. 2016. p. 2087-103.
7. R, H. and P. DA, *Selection on Codon Bias.* Annual review of genetics, 2008. **42**.
8. DA, D. and W. CO, *Mistranslation-induced Protein Misfolding as a Dominant Constraint on Coding-Sequence Evolution.* Cell, 2008. **134**(2).
9. Zhou, T., M. Weems, and C.O. Wilke, *Translationally optimal codons associate with structurally sensitive sites in proteins.* Mol Biol Evol, 2009. **26**(7): p. 1571-80.
10. N, S. and E.-W. A, *Synonymous Codon Usage in Escherichia Coli: Selection for Translational Accuracy.* Molecular biology and evolution, 2007. **24**(2).
11. W, R. and H. PG, *Contributions of Speed and Accuracy to Translational Selection in Bacteria.* PloS one, 2012. **7**(12).
12. W, R. and H. PG, *The Influence of Anticodon-Codon Interactions and Modified Bases on Codon Usage Bias in Bacteria.* Molecular biology and evolution, 2010. **27**(9).
13. Shabalina, S.A., N.A. Spiridonov, and A. Kashina, *Sounds of silence: synonymous nucleotides as a key to biological regulation and complexity*, in *Nucleic Acids Res*. 2013. p. 2073-94.
14. Rodnina, M.V., *The ribosome in action: Tuning of translational efficiency and protein folding*, in *Protein Sci*. 2016. p. 1390-406.
15. Hanson, G. and J. Coller, *Codon optimality, bias and usage in translation and mRNA decay.* Nat Rev Mol Cell Biol, 2018. **19**(1): p. 20-30.
16. Brule, C.E. and E.J. Grayhack, *Synonymous codons: Choose wisely for expression.* Trends Genet, 2017. **33**(4): p. 283-97.
17. Dever, T.E., J.D. Dinman, and R. Green, *Translation Elongation and Recoding in Eukaryotes*, in *Cold Spring Harb Perspect Biol*. 2018.
18. ZE, S. and K.-S. C, *Understanding the contribution of synonymous mutations to human disease.* Nature reviews. Genetics, 2011. **12**(10).
19. Chaney, J.L. and P.L. Clark, *Roles for Synonymous Codon Usage in Protein Biogenesis.* http://dx.doi.org/10.1146/annurev-biophys-060414-034333, 2015.
20. Quax, T.E., et al., *Codon Bias as a Means to Fine-Tune Gene Expression.* Mol Cell, 2015. **59**(2): p. 149-61.
21. Tuller, T. and H. Zur, *Multiple roles of the coding sequence 5' end in gene expression regulation.* Nucleic Acids Res, 2015. **43**(1): p. 13-28.
22. Zur, H., et al., *Predictive biophysical modeling and understanding of the dynamics of mRNA translation and its evolution.* Nucleic Acids Research, 2020. **44**(19): p. 9031-9049.
23. Bali, V. and Z. Bebok, *Decoding Mechanisms by which Silent Codon Changes Influence Protein Biogenesis and Function.* Int J Biochem Cell Biol, 2015. **64**: p. 58-74.
24. AA, K., *The Yin and Yang of codon usage.* Human molecular genetics, 2016. **25**(R2).

25. Goz, E., H. Zur, and T. Tuller, *Hidden Silent Codes in Viral Genomes*, in *Evolutionary Biology: Self/Nonself Evolution, Species and Complex Traits Evolution, Methods and Concepts*. p. 87-110.
26. S, B. and T. T, *Widespread non-modular overlapping codes in the coding regions.* Physical biology, 2020. **17**(3).
27. Sharp, P.M. and W.H. Li, *The codon Adaptation Index--a measure of directional synonymous codon usage bias, and its potential applications.* Nucleic Acids Res, 1987. **15**(3): p. 1281-95.
28. dos Reis, M., R. Savva, and L. Wernisch, *Solving the riddle of codon usage preferences: a test for translational selection.* Nucleic Acids Res, 2004. **32**(17): p. 5036-44.
29. Ikemura, T., *Correlation between the abundance of Escherichia coli transfer RNAs and the occurrence of the respective codons in its protein genes.* J Mol Biol, 1981. **146**(1): p. 1-21.
30. Percudani, R., A. Pavesi, and S. Ottonello, *Transfer RNA gene redundancy and translational selection in Saccharomyces cerevisiae.* J Mol Biol, 1997. **268**(2): p. 322-30.
31. Duret, L., *tRNA gene number and codon usage in the C. elegans genome are co-adapted for optimal translation of highly expressed genes.* Trends Genet, 2000. **16**(7): p. 287-9.
32. Sabi, R. and T. Tuller, *Modelling the Efficiency of Codon–tRNA Interactions Based on Codon Usage Bias*, in *DNA Res*. 2014. p. 511-25.
33. Sabi, R., R. Volvovitch Daniel, and T. Tuller, *stAIcalc: tRNA adaptation index calculator based on species-specific weights.* Bioinformatics, 2017. **33**(4): p. 589-591.
34. Zhang, G., et al., *Global and local depletion of ternary complex limits translational elongation.* Nucleic Acids Res, 2010. **38**(14): p. 4778-87.
35. Dana, A. and T. Tuller, *The effect of tRNA levels on decoding times of mRNA codons.* Nucleic Acids Res, 2014. **42**(14): p. 9171-81.
36. Elf, J., et al., *Selective charging of tRNA isoacceptors explains patterns of codon usage.* Science, 2003. **300**(5626): p. 1718-22.
37. Wu, Q., et al., *Translation affects mRNA stability in a codon-dependent manner in human cells.* Elife, 2019. **8**.
38. Forrest, M.E., et al., *Codon and amino acid content are associated with mRNA stability in mammalian cells*, in *PLoS One*. 2020.
39. Narula, A., et al., *Coding regions affect mRNA stability in human cells*, in *RNA*. 2019. p. 1751-64.
40. Hia, F., et al., *Codon bias confers stability to human mRNAs.* EMBO Rep, 2019. **20**(11): p. e48220.
41. Pop, C., et al., *Causal signals between codon bias, mRNA structure, and the efficiency of translation and elongation.* Mol Syst Biol, 2014. **10**: p. 770.
42. Pringle, E.S., C. McCormick, and Z. Cheng, *Polysome Profiling Analysis of mRNA and Associated Proteins Engaged in Translation.* Curr Protoc Mol Biol, 2019. **125**(1): p. e79.
43. Riba, A., et al., *Protein synthesis rates and ribosome occupancies reveal determinants of translation elongation rates*, in *Proc Natl Acad Sci U S A*. 2019. p. 15023-32.
44. Shine, J. and L. Dalgarno, *Determinant of cistron specificity in bacterial ribosomes.* Nature, 1975. **254**(5495): p. 34-8.
45. Marilyn, K., *The scanning model for translation: an update*, in *J Cell Biol*. 1989. p. 229-41.
46. Bentele, K., et al., *Efficient translation initiation dictates codon usage at gene start*, in *Mol Syst Biol*. 2013. p. 675.
47. Goodman, D.B., G.M. Church, and S. Kosuri, *Causes and effects of N-terminal codon bias in bacterial genes.* Science, 2013. **342**(6157): p. 475-9.
48. Kudla, G., et al., *Coding-sequence determinants of gene expression in Escherichia coli.* Science, 2009. **324**(5924): p. 255-8.
49. Verma, M., et al., *A short translational ramp determines the efficiency of protein synthesis.* Nat Commun, 2019. **10**(1): p. 5774.
50. Tuller, T., et al., *An evolutionarily conserved mechanism for controlling the efficiency of protein translation.* Cell, 2010. **141**(2): p. 344-54.
51. Dobrzynski, M. and F.J. Bruggeman, *Elongation dynamics shape bursty transcription and translation.* Proc Natl Acad Sci U S A, 2009. **106**(8): p. 2583-8.
52. Keller, T.E., et al., *Reduced mRNA Secondary-Structure Stability Near the Start Codon Indicates Functional Genes in Prokaryotes.* Genome Biol Evol, 2012. **4**(2): p. 80-8.
53. Grosjean, H. and W. Fiers, *Preferential codon usage in prokaryotic genes: the optimal codon-anticodon interaction energy and the selective codon usage in efficiently expressed genes.* Gene, 1982. **18**(3): p. 199-209.
54. Grosjean, H., et al., *Bacteriophage MS2 RNA: a correlation between the stability of the codon: anticodon interaction and the choice of code words.* J Mol Evol, 1978. **12**(2): p. 113-9.
55. Gauss, D.H. and M. Sprinzl, *Compilation of tRNA sequences.* Nucleic Acids Res, 1981. **9**(1): p. r1-r23.

56. Pedersen, S., *Escherichia coli ribosomes translate in vivo with variable rate.* Embo j, 1984. **3**(12): p. 2895-8.
57. Sorensen, M.A. and S. Pedersen, *Absolute in vivo translation rates of individual codons in Escherichia coli. The two glutamic acid codons GAA and GAG are translated with a threefold difference in rate.* J Mol Biol, 1991. **222**(2): p. 265-80.
58. Sorensen, M.A., C.G. Kurland, and S. Pedersen, *Codon usage determines translation rate in Escherichia coli.* J Mol Biol, 1989. **207**(2): p. 365-77.
59. Frumkin, I., et al., *Codon usage of highly expressed genes affects proteome-wide translation efficiency.* Proc Natl Acad Sci U S A, 2018. **115**(21): p. E4940-e4949.
60. Gobet, C., et al., *Robust landscapes of ribosome dwell times and aminoacyl-tRNAs in response to nutrient stress in liver.* Proc Natl Acad Sci U S A, 2020.
61. Carlini, D.B. and W. Stephan, *In vivo introduction of unpreferred synonymous codons into the Drosophila Adh gene results in reduced levels of ADH protein.* Genetics, 2003. **163**(1): p. 239-43.
62. Mishima, Y. and Y. Tomari, *Codon Usage and 3' UTR Length Determine Maternal mRNA Stability in Zebrafish.* Mol Cell, 2016. **61**(6): p. 874-85.
63. Zhou, Z., et al., *Codon usage is an important determinant of gene expression levels largely through its effects on transcription*, in *Proc Natl Acad Sci U S A*. 2016. p. E6117-25.
64. Yu, C.H., et al., *Codon usage influences the local rate of translation elongation to regulate co-translational protein folding.* Mol Cell, 2015. **59**(5): p. 744-54.
65. Jin, H.Y. and C. Xiao, *An Integrated Polysome Profiling and Ribosome Profiling Method to Investigate In Vivo Translatome.* Methods Mol Biol, 2018. **1712**: p. 1-18.
66. Lampson, B.L., et al., *Rare codons regulate KRas oncogenesis.* Curr Biol, 2013. **23**(1): p. 70-5.
67. Fu, J., et al., *Codon usage regulates human KRAS expression at both transcriptional and translational levels*, in *J Biol Chem*. 2018. p. 17929-40.
68. Ingolia, N.T., et al., *Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling.* Science, 2009. **324**(5924): p. 218-23.
69. McGlincy, N.J. and N.T. Ingolia, *Transcriptome-wide measurement of translation by ribosome profiling.* Methods, 2017. **126**: p. 112-129.
70. Charneski, C.A. and L.D. Hurst, *Positively charged residues are the major determinants of ribosomal velocity.* PLoS Biol, 2013. **11**(3): p. e1001508.
71. Ingolia, N.T., L.F. Lareau, and J.S. Weissman, *Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes.* Cell, 2011. **147**(4): p. 789-802.
72. Li, G.W., E. Oh, and J.S. Weissman, *The anti-Shine-Dalgarno sequence drives translational pausing and codon choice in bacteria.* Nature, 2012. **484**(7395): p. 538-41.
73. Qian, W., et al., *Balanced codon usage optimizes eukaryotic translational efficiency.* PLoS Genet, 2012. **8**(3): p. e1002603.
74. Artieri, C.G. and H.B. Fraser, *Accounting for biases in riboprofiling data indicates a major role for proline in stalling translation.* Genome Res, 2014. **24**(12): p. 2011-21.
75. Weinberg, D.E., et al., *Improved ribosome-footprint and mRNA measurements provide insights into dynamics and regulation of yeast translation.* Cell Rep, 2016. **14**(7): p. 1787-99.
76. Gardin, J., et al., *Measurement of average decoding rates of the 61 sense codons in vivo.* Elife, 2014. **3**.
77. Nakahigashi, K., et al., *Effect of codon adaptation on codon-level and gene-level translation efficiency in vivo.* BMC Genomics, 2014. **15**: p. 1115.
78. MV, G. and G. VN, *Ribonuclease selection for ribosome profiling.* Nucleic acids research, 2017. **45**(2).
79. Gerashchenko, M.V. and V.N. Gladyshev, *Translation inhibitors cause abnormalities in ribosome profiling experiments.* Nucleic Acids Res, 2014. **42**(17): p. e134.
80. Wright, G., et al., *Analysis of computational codon usage models and their association with translationally slow codons.* PLoS One, 2020. **15**(4): p. e0232003.
81. A, D. and T. T, *Estimation of ribosome profiling performance and reproducibility at various levels of resolution.* Biology direct, 2016. **11**.
82. Lareau, L.F., et al., *Distinct stages of the translation elongation cycle revealed by sequencing ribosome-protected mRNA fragments*, in *eLife*. 2014.
83. Hussmann, J.A., et al., *Understanding Biases in Ribosome Profiling Experiments Reveals Signatures of Translation Dynamics in Yeast.* PLoS Genet, 2015. **11**(12): p. e1005732.
84. Santos, D.A., et al., *Cycloheximide can distort measurements of mRNA levels and translation efficiency.* Nucleic Acids Res, 2019. **47**(10): p. 4974-4985.
85. Pelechano, V., W. Wei, and L.M. Steinmetz, *Widespread Co-translational RNA Decay Reveals Ribosome Dynamics.* Cell, 2015. **161**(6): p. 1400-12.
86. Pelechano, V., W. Wei, and L.M. Steinmetz, *Genome-wide quantification of 5'-phosphorylated mRNA degradation intermediates for analysis of ribosome dynamics.* Nat Protoc, 2016. **11**(2): p. 359-76.

87. Ingolia, N.T., et al., *The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments.* Nat Protoc. **7**(8): p. 1534-50.
88. Young, D.J., et al., *Rli1/ABCE1 recycles terminating ribosomes and controls translation reinitiation in 3'UTRs in vivo.* Cell, 2015. **162**(4): p. 872-84.
89. Guydosh, N.R. and R. Green, *Dom34 rescues ribosomes in 3' untranslated regions.* Cell, 2014. **156**(5): p. 950-62.
90. Shah, P., et al., *Rate-Limiting Steps in Yeast Protein Translation*, in *Cell*. 2013. p. 1589-601.
91. SK, A., et al., *Dynamics of ribosome scanning and recycling revealed by translation complex profiling.* Nature, 2016. **535**(7613).
92. Zlotorynski, E., *Profiling ribosome dynamics.* Nature Reviews Molecular Cell Biology, 2016. **17**(9): p. 535-535.
93. E, O., et al., *Selective ribosome profiling reveals the cotranslational chaperone action of trigger factor in vivo.* Cell, 2011. **147**(6).
94. Schibich, D., et al., *Global profiling of SRP interaction with nascent polypeptides.* Nature, 2016. **536**(7615): p. 219-223.
95. CV, G., et al., *Selective ribosome profiling to study interactions of translating ribosomes in yeast.* Nature protocols, 2019. **14**(8).
96. Shiber, A., et al., *Cotranslational assembly of protein complexes in eukaryotes revealed by ribosome profiling.* Nature, 2018. **561**(7722): p. 268-272.
97. CC, W., et al., *Ribosome Collisions Trigger General Stress Responses to Regulate Cell Fate.* Cell, 2020. **182**(2).
98. K, I., et al., *Collided ribosomes form a unique structural interface to induce Hel2-driven quality control pathways.* The EMBO journal, 2019. **38**(5).
99. P, T., et al., *Molecular mechanism of translational stalling by inhibitory codon combinations and poly(A) tracts.* The EMBO journal, 2020. **39**(3).
100. A, D., et al., *The extent of ribosome queuing in budding yeast.* PLoS computational biology, 2018. **14**(1).
101. CE, G., et al., *Adjacent Codons Act in Concert to Modulate Translation Efficiency in Yeast.* Cell, 2016. **166**(3).
102. Y, M., et al., *Ubiquitination of stalled ribosome triggers ribosome-associated quality control.* Nature communications, 2017. **8**(1).
103. P, H., et al., *Genome-wide Survey of Ribosome Collision.* Cell reports, 2020. **31**(5).
104. S, M. and G. NR, *Disome and Trisome Profiling Reveal Genome-wide Targets of Ribosome Quality Control.* Molecular cell, 2020.
105. K, R., et al., *Ribosome profiling reveals features of normal and disease-associated mitochondrial translation.* Nature communications, 2013. **4**.
106. Gonzalez, C., et al., *Ribosome Profiling Reveals a Cell-Type-Specific Translational Landscape in Brain Tumors*, in *J Neurosci*. 2014. p. 10924-36.
107. Stern-Ginossar, N. and N.T. Ingolia, *Ribosome Profiling as a Tool to Decipher Viral Complexity.* http://dx.doi.org/10.1146/annurev-virology-100114-054854, 2015.
108. Peltz, S.W., J.L. Donahue, and A. Jacobson, *A mutation in the tRNA nucleotidyltransferase gene promotes stabilization of mRNAs in Saccharomyces cerevisiae.* Mol Cell Biol, 1992. **12**(12): p. 5778-84.
109. Herrick, D., R. Parker, and A. Jacobson, *Identification and comparison of stable and unstable mRNAs in Saccharomyces cerevisiae.* Mol Cell Biol, 1990. **10**(5): p. 2269-84.
110. Kurosaki, T., J.R. Myers, and L.E. Maquat, *Defining nonsense-mediated mRNA decay intermediates in human cells.* Methods, 2019. **155**: p. 68-76.
111. Antic, S., et al., *General and MicroRNA-Mediated mRNA Degradation Occurs on Ribosome Complexes in Drosophila Cells.* Mol Cell Biol, 2015. **35**(13): p. 2309-20.
112. Graille, M. and B. Seraphin, *Surveillance pathways rescuing eukaryotic ribosomes lost in translation*, in *Nat Rev Mol Cell Biol*. 2012: England. p. 727-35.
113. Shoemaker, C.J. and R. Green, *Translation drives mRNA quality control.* Nat Struct Mol Biol, 2012. **19**(6): p. 594-601.
114. Hu, W., et al., *Co-translational mRNA decay in Saccharomyces cerevisiae.* Nature, 2009. **461**(7261): p. 225-9.
115. Coller, J. and R. Parker, *General translational repression by activators of mRNA decapping.* Cell, 2005. **122**(6): p. 875-86.
116. Coller, J.M., et al., *The DEAD box helicase, Dhh1p, functions in mRNA decapping and interacts with both the decapping and deadenylase complexes.* Rna, 2001. **7**(12): p. 1717-27.

117. Sweet, T., C. Kovalak, and J. Coller, *The DEAD-box protein Dhh1 promotes decapping by slowing ribosome movement.* PLoS Biol, 2012. **10**(6): p. e1001342.
118. Harigaya, Y. and R. Parker, *Analysis of the association between codon optimality and mRNA stability in Schizosaccharomyces pombe.* BMC Genomics, 2016. **17**(1): p. 895.
119. de Freitas Nascimento, J., et al., *Codon choice directs constitutive mRNA levels in trypanosomes.* Elife, 2018. **7**.
120. Jeacock, L., J. Faria, and D. Horn, *Codon usage bias controls mRNA and protein abundance in trypanosomatids.* Elife, 2018. **7**.
121. Buschauer, R., et al., *The Ccr4-Not complex monitors the translating ribosome for codon optimality.* Science, 2020. **368**(6488).
122. P, T., et al., *Structure of the 80S ribosome-Xrn1 nuclease complex.* Nature structural & molecular biology, 2019. **26**(4).
123. Radhakrishnan, A., et al., *The DEAD-Box Protein Dhh1p Couples mRNA Decay and Translation by Monitoring Codon Optimality.* Cell, 2016. **167**(1): p. 122-132.e9.
124. F, H., et al., *General decapping activators target different subsets of inefficiently translated mRNAs.* eLife, 2018. **7**.
125. Freimer, J.W., T. Hu, and R. Blelloch, *Decoupling the impact of microRNAs on translational repression versus RNA degradation in embryonic stem cells*, in *eLife*. 2018.
126. Courel, M., et al., *GC content shapes mRNA storage and decay in human cells.* Elife, 2019. **8**.
127. Hanson, G., et al., *Translation elongation and mRNA stability are coupled through the ribosomal A-site*, in *RNA*. 2018. p. 1377-89.
128. Dao Duc, K. and Y.S. Song, *The impact of ribosomal interference, codon usage, and exit tunnel interactions on translation elongation rate variation*, in *PLoS Genet*. 2018.
129. Schwartz, D.C. and R. Parker, *Mutations in Translation Initiation Factors Lead to Increased Rates of Deadenylation and Decapping of mRNAs in Saccharomyces cerevisiae*, in *Mol Cell Biol*. 1999. p. 5247-56.
130. Schwartz, D.C. and R. Parker, *mRNA Decapping in Yeast Requires Dissociation of the Cap Binding Protein, Eukaryotic Translation Initiation Factor 4E*, in *Mol Cell Biol*. 2000. p. 7933-42.
131. Edri, S. and T. Tuller, *Quantifying the Effect of Ribosomal Density on mRNA Stability*, in *PLoS One*. 2014.
132. Chan, L.Y., et al., *Non-invasive measurement of mRNA decay reveals translation initiation as the major determinant of mRNA stability.* Elife, 2018. **7**.
133. Neymotin, B., V. Ettorre, and D. Gresham, *Multiple Transcript Properties Related to Translation Affect mRNA Degradation Rates in Saccharomyces cerevisiae*, in *G3 (Bethesda)*. 2016. p. 3475-83.
134. Purvis, I.J., et al., *The efficiency of folding of some proteins is increased by controlled rates of translation in vivo. A hypothesis.* J Mol Biol, 1987. **193**(2): p. 413-7.
135. Thanaraj, T.A. and P. Argos, *Ribosome-mediated translational pause and protein domain organization.* Protein Sci, 1996. **5**(8): p. 1594-612.
136. Thanaraj, T.A. and P. Argos, *Protein secondary structural types are differentially coded on messenger RNA.* Protein Sci, 1996. **5**(10): p. 1973-83.
137. Krasheninnikov, I.A., A.A. Komar, and I.A. Adzhubei, *Nonuniform size distribution of nascent globin peptides, evidence for pause localization sites, and a contranslational protein-folding model.* J Protein Chem, 1991. **10**(5): p. 445-53.
138. Chartier, M., F. Gaudreault, and R. Najmanovich, *Large-scale analysis of conserved rare codon clusters suggests an involvement in co-translational molecular recognition events*, in *Bioinformatics*. 2012. p. 1438-45.
139. Fluman, N., et al., *mRNA-programmed translation pauses in the targeting of E. coli membrane proteins*, in *eLife*. 2014.
140. Bitran, A., et al., *Cotranslational folding allows misfolding-prone proteins to circumvent deep kinetic traps.* Proc Natl Acad Sci U S A, 2020. **117**(3): p. 1485-1495.
141. Zhao, F., C.H. Yu, and Y. Liu, *Codon usage regulates protein structure and function by affecting translation elongation speed in Drosophila cells.* Nucleic Acids Res, 2017. **45**(14): p. 8484-8492.
142. Oresic, M. and D. Shalloway, *Specific correlations between relative synonymous codon usage and protein secondary structure.* J Mol Biol, 1998. **281**(1): p. 31-48.
143. Adzhubei, I.A., A.A. Adzhubei, and S. Neidle, *An Integrated Sequence-Structure Database incorporating matching mRNA sequence, amino acid sequence and protein three-dimensional structure data.* Nucleic Acids Res, 1998. **26**(1): p. 327-31.
144. Smith, D.W., *Problems of translating heterologous genes in expression systems: the role of tRNA.* Biotechnol Prog, 1996. **12**(4): p. 417-22.

145. Kurland, C. and J. Gallant, *Errors of heterologous protein expression.* Curr Opin Biotechnol, 1996. **7**(5): p. 489-93.
146. Komar, A.A. and R. Jaenicke, *Kinetics of translation of gamma B crystallin and its circularly permutated variant in an in vitro cell-free system: possible relations to codon distribution and protein folding.* FEBS Lett, 1995. **376**(3): p. 195-8.
147. Komar, A.A., T. Lesnik, and C. Reiss, *Synonymous codon substitutions affect ribosome traffic and protein folding during in vitro translation.* FEBS Lett, 1999. **462**(3): p. 387-91.
148. Spencer, P.S., et al., *Silent substitutions predictably alter translation elongation rates and protein folding efficiencies.* J Mol Biol, 2012. **422**(3): p. 328-35.
149. Zhang, G., M. Hubalewska, and Z. Ignatova, *Transient ribosomal attenuation coordinates protein synthesis and co-translational folding.* Nat Struct Mol Biol, 2009. **16**(3): p. 274-80.
150. Zhou, M., et al., *Nonoptimal codon usage influences protein structure in intrinsically disordered regions.* Mol Microbiol, 2015. **97**(5): p. 974-87.
151. Buhr, F., et al., *Synonymous Codons Direct Cotranslational Folding toward Different Protein Conformations.* Mol Cell, 2016. **61**(3): p. 341-351.
152. Yang, J.R., X. Chen, and J. Zhang, *Codon-by-Codon Modulation of Translational Speed and Accuracy Via mRNA Folding*, in *PLoS Biol*. 2014.
153. Faure, G., et al., *Role of mRNA structure in the control of protein folding.* Nucleic Acids Res, 2016. **44**(22): p. 10898-911.
154. PM, S., et al., *Variation in the strength of selected codon usage bias among bacteria.* Nucleic acids research, 2005. **33**(4).
155. Saibil, H., *Chaperone machines for protein folding, unfolding and disaggregation.* Nat Rev Mol Cell Biol, 2013. **14**(10): p. 630-42.
156. X, X., *Maximizing transcription efficiency causes codon usage bias.* Genetics, 1996. **144**(3).
157. E, C., Z. Z, and T. T, *A code for transcription elongation speed.* RNA biology, 2018. **15**(1).
158. Kudla, G., et al., *High guanine and cytosine content increases mRNA levels in mammalian cells.* PLoS Biol, 2006. **4**(6): p. e180.
159. Newman, Z.R., et al., *Differences in codon bias and GC content contribute to the balanced expression of TLR7 and TLR9*, in *Proc Natl Acad Sci U S A*. 2016. p. E1362-71.
160. Mordstein, C., et al., *Codon Usage and Splicing Jointly Influence mRNA Localization*, in *Cell Syst*. 2020. p. 351-362 e8.
161. Fontrodona, N., et al., *Interplay between coding and exonic splicing regulatory sequences.* Genome Res, 2019. **29**(5): p. 711-22.
162. AB, S., et al., *Exonic transcription factor binding directs codon choice and affects protein evolution.* Science (New York, N.Y.), 2013. **342**(6164).
163. Crick, F.H., *Codon--anticodon pairing: the wobble hypothesis.* J Mol Biol, 1966. **19**(2): p. 548-55.
164. Roth, A.C., *Decoding properties of tRNA leave a detectable signal in codon usage bias*, in *Bioinformatics*. 2012. p. i340-8.
165. Gromadski, K.B., T. Daviter, and M.V. Rodnina, *A uniform response to mismatches in codon-anticodon complexes ensures ribosomal fidelity.* Mol Cell, 2006. **21**(3): p. 369-77.
166. Stadler, M. and A. Fire, *Wobble base-pairing slows in vivo translation elongation in metazoans.* Rna, 2011. **17**(12): p. 2063-73.
167. Dedon, P.C. and T.J. Begley, *A System of RNA Modifications and Biased Codon Use Controls Cellular Stress Response at the Level of Translation.* Chem Res Toxicol, 2014. **27**(3): p. 330-7.
168. Boccaletto, P., et al., *MODOMICS: a database of RNA modification pathways. 2017 update*, in *Nucleic Acids Res*. 2018. p. D303-7.
169. Deng, W., et al., *Trm9-Catalyzed tRNA Modifications Regulate Global Protein Expression by Codon-Biased Translation.* PLoS Genet, 2015. **11**(12): p. e1005706.
170. Jaroensuk, J., et al., *Methylation at position 32 of tRNA catalyzed by TrmJ alters oxidative stress response in Pseudomonas aeruginosa.* Nucleic Acids Res, 2016. **44**(22): p. 10834-10848.
171. Gu, C., T.J. Begley, and P.C. Dedon, *tRNA modifications regulate translation during cellular stress.* FEBS Lett, 2014. **588**(23): p. 4287-96.
172. Chan, C.T., et al., *Reprogramming of tRNA modifications controls the oxidative stress response by codon-biased translation of proteins.* Nat Commun, 2012. **3**: p. 937.
173. Chionh, Y.H., et al., *tRNA-mediated codon-biased translation in mycobacterial hypoxic persistence.* Nat Commun, 2016. **7**: p. 13302.
174. Nedialkova, D.D. and S.A. Leidel, *Optimization of Codon Translation Rates via tRNA Modifications Maintains Proteome Integrity.* Cell, 2015. **161**(7): p. 1606-18.
175. Bornelöv, S., et al., *Codon usage optimization in pluripotent embryonic stem cells*, in *Genome Biol*. 2019.

176. Arango, D., et al., *Acetylation of Cytidine in mRNA Promotes Translation Efficiency.* Cell, 2018. **175**(7): p. 1872-1886.e24.
177. Eyler, D.E., et al., *Pseudouridinylation of mRNA coding sequences alters translation*, in *Proc Natl Acad Sci U S A*. 2019. p. 23068-74.
178. Mao, Y., et al., *m(6)A in mRNA coding regions promotes translation via the RNA helicase-containing YTHDC2.* Nat Commun, 2019. **10**(1): p. 5332.
179. Liu, Z. and J. Zhang, *Most m6A RNA Modifications in Protein-Coding Regions Are Evolutionarily Unconserved and Likely Nonfunctional*, in *Mol Biol Evol*. 2018. p. 666-75.
180. KA, D., et al., *Selective Charging of tRNA Isoacceptors Induced by Amino-Acid Starvation.* EMBO reports, 2005. **6**(2).
181. Gingold, H., O. Dahan, and Y. Pilpel, *Dynamic changes in translational efficiency are deduced from codon usage of the transcriptome*, in *Nucleic Acids Res*. 2012. p. 10053-63.
182. Torrent, M., et al., *Cells alter their tRNA abundance to selectively regulate protein synthesis during stress conditions*, in *Sci Signal*. 2018.
183. M, F.-M., et al., *Genes adopt non-optimal codon usage to generate cell cycle-dependent oscillations in protein levels.* Molecular systems biology, 2012. **8**.
184. R, S. and T. T, *Novel insights into gene expression regulation during meiosis revealed by translation elongation dynamics.* NPJ systems biology and applications, 2019. **5**.
185. JM, G. and P. T, *Diversity of tRNA genes in eukaryotes.* Nucleic acids research, 2006. **34**(21).
186. Kutter, C., et al., *Pol III binding in six mammalian genomes shows high conservation among amino acid isotypes, despite divergence in tRNA gene usage.* Nat Genet. **43**(10): p. 948-55.
187. Geslain, R. and T. Pan, *Functional analysis of human tRNA isodecoders.* J Mol Biol, 2010. **396**(3): p. 821.
188. M, Z., et al., *Non-optimal Codon Usage Affects Expression, Structure and Function of Clock Protein FRQ.* Nature, 2013. **495**(7439).
189. Y, X., et al., *Non-optimal Codon Usage Is a Mechanism to Achieve Circadian Clock Conditionality.* Nature, 2013. **495**(7439).
190. M, S., et al., *Codon Optimality Controls Differential mRNA Translation During Amino Acid Starvation.* RNA (New York, N.Y.), 2016. **22**(11).
191. Guimaraes, J.C., et al., *A rare codon-based translational program of cell proliferation*, in *Genome Biol*. 2020.
192. Burrow, D.A., et al., *Attenuated Codon Optimality Contributes to Neural-Specific mRNA Decay in Drosophila.* Cell Rep, 2018. **24**(7): p. 1704-12.
193. Begley, T.J., et al., *Hot spots for modulating toxicity identified by genomic phenotyping and localization mapping.* Mol Cell, 2004. **16**(1): p. 117-25.
194. Bennett, C.B., et al., *Genes required for ionizing radiation resistance in yeast.* Nat Genet, 2001. **29**(4): p. 426-34.
195. Kalhor, H.R. and S. Clarke, *Novel methyltransferase for modified uridine residues at the wobble position of tRNA.* Mol Cell Biol, 2003. **23**(24): p. 9283-92.
196. Begley, U., et al., *Trm9-catalyzed tRNA modifications link translation to the DNA damage response.* Mol Cell, 2007. **28**(5): p. 860-70.

Figure Legend

Figure 1. Synonymous codons regulate translation. (**a**) Diagram depicting the effects of the 'ramp' on ribosome density. Translation is slow at the start (indicated by high ribosome density) but decreases thereafter, possibly indicating an increase in translation speed. (**b**) Usage of rare codons results in decreased protein synthesis rates as their cognate tRNA are less abundant. (**c**) Summary of the effects of optimal and non-optimal codons on translation.

Figure 2. Ribosome profiling for monitoring translation dynamics. (**a**) Outline of the ribosome profiling workflow. Translating ribosomes are firstly immobilized and then subjected to RNase treatment. The resulting fragments are then subjected to protease treatment to remove the ribosomes. High-throughput sequencing follows and the ribosome densities are calculated. (**b**) As non-optimal codons are translated at a slower speed, the ribosome densities at sites of low codon optimality are expected to depict a high ribosome density. (**c**) Simplifed diagram highlighting how ribosome collisions occur. A stalled ribosome is subject to collisions by the trailing ribosome forming a disome. Further collisions may result in further staking of ribosomes.

Figure 3. Synonymous codons can regulate mRNA stability (**a**) Outline of the fate of transcripts during degradation. Transcripts can be degraded from the 5' to 3' direction by Xrn1 after decapping or in the 3' to 5' direction by the exosome complex. (**b**) Dhh1 acts as a sensor of codon optimality by facilitating the degradation of transcripts possessing low codon optimality. (**c**) mRNAs possessing high GC-content are stable and are translated more efficiently that their AU-rich counterparts. (**d**) The competition between mRNA degradation factors and translation initiation factors can determine the fate of the transcript. Transcripts occupied by ribosomes can exclude decay factors.

Figure 4. Both optimal and non-optimal codons are utilized in several ways to ensure translation fidelity and proper folding. (**a**) Structurally sensitive sites of proteins are generally coded for by optimal codons to ensure high translation fidelity and the consequent production of functional proteins. (**c**) Non-optimal codons can be utilized to slow translation to enable folding of the nascent peptide chain into the correct conformation for a functional protein. Substitution of the non-optimal codons with optimal codons alters the folding of the peptide, resulting in incorrectly folded and non-functional proteins.

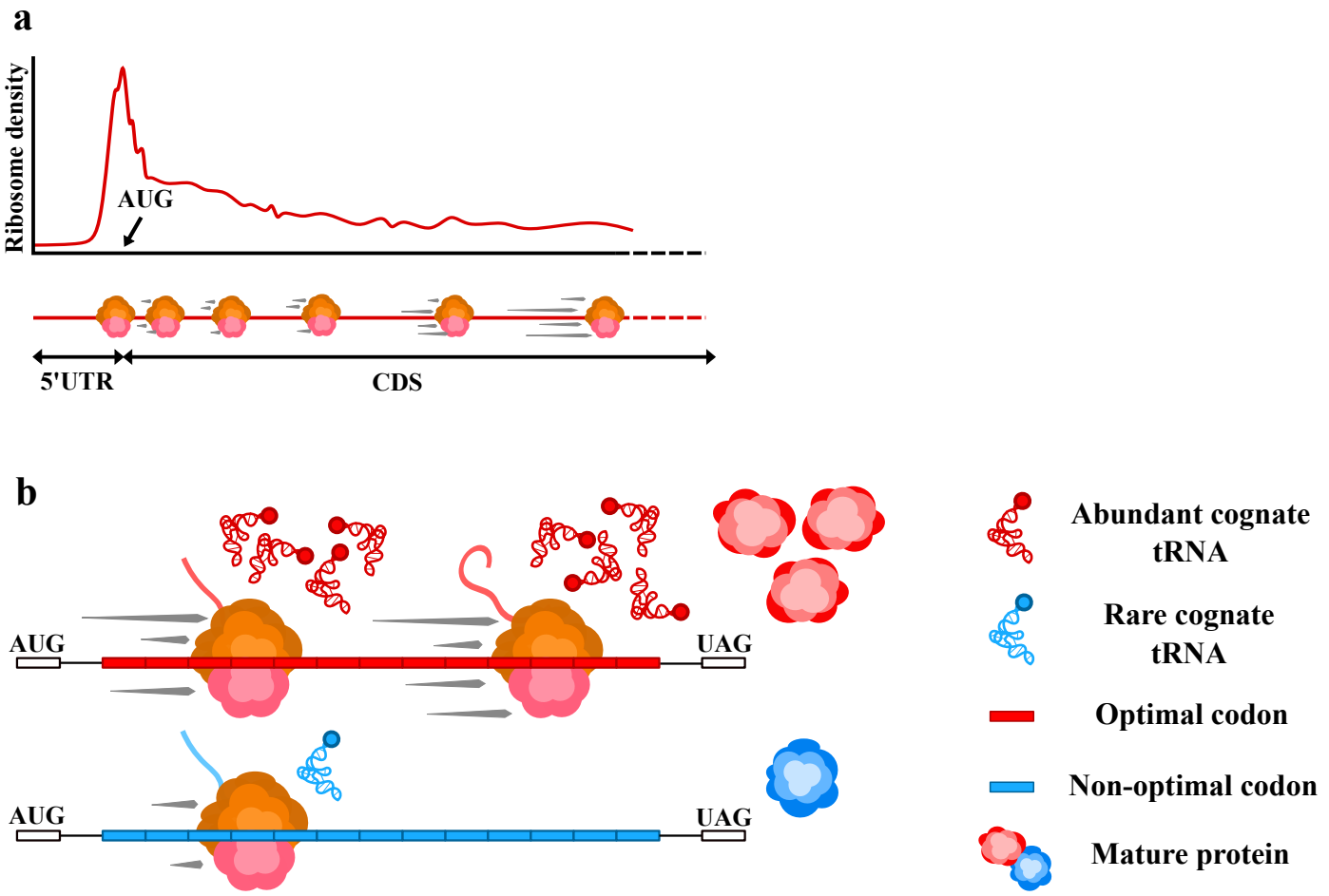
Figure 5. Synonymous codons can regulate transcription. (**a**) H3K9 trimethylation on sequences which possess non-optimal codons can have a silencing effect on transcription. (**b**) Schematic diagram illustrating how splicing, GC-content and 5'GC content affects expression of a gene.

Figure 6. RNA modifications have an effect on translation and mRNA stability. (**a**) Generalized illustration of one of the possible effects of wobble base-modified tRNA. Such an effect can be observed with Trm4-catalyzed modification of C at the wobble base of $tRNA^{Leu(CCA)}$. (**b**) Besides wobble base modifications of tRNA, mRNA can be subject to modifications such as ac4C (at the wobble position), pesudouridine and m6A among others. Modifications such as ac4C can have a mRNA stabilizing effect in addition to increased translation efficiency.

Figure 7. Functional importance of synonymous codons on physiological processes. (**a**) Selective charging of tRNA isoacceptors after starvation. Highly abundant tRNA isoacceptor (major tRNA, red) with a given anticodon $(NNN)_1$ may possess lower charging levels after amino acid starvation. Conversely a less abundant isoacceptor (minor tRNA, blue) with a given anticodon $(NNN)_2$ can maintain high charging levels to ensure transcripts enriched with this codon can be rapidly transated to produce a stress response. (**b**) Organisms can induce changes to their charged tRNA pool (and isoacceptor abundance) to ensure a timely response to stress. Under optimal growth conditions, transcripts which possess codons which require a major tRNA isoacceptor (red) are translated efficiently. Under stress, organisms can facilitate changes to tRNA abundances to ensure the efficient translation of a separate set of stress response transcripts which utilize minor tRNA isoacceptors (blue).

Figure 8. Summary of how codon optimality affects mRNA and protein regulation. Organisms have evolved a system of codon optimality to fine-tune mRNA and protein levels. It should be noted that certain processes may be organism specific and may differ in the extent of their impact.

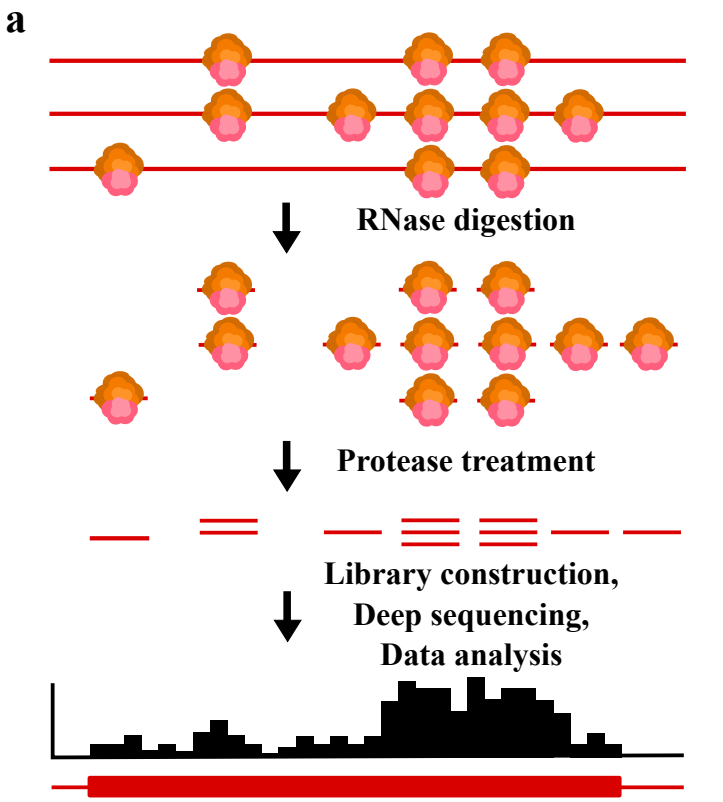
Figure 1

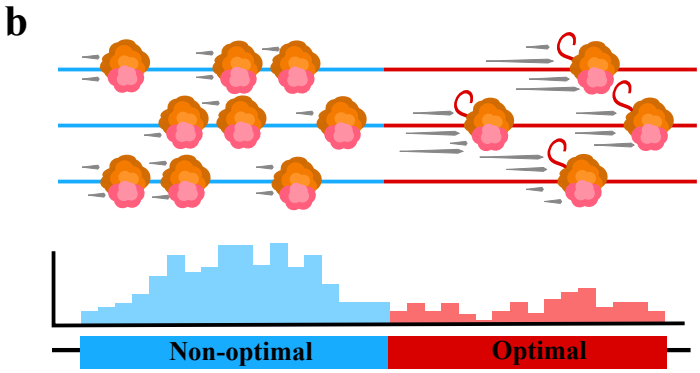


c

	Optimal	Non-optimal
Translation	Fast	Slow
Cognate tRNA	Abundant	Rare
Protein abundance	High	Low

Figure 2


a
RNase digestion
Protease treatment
Library construction,
Deep sequencing,
Data analysis


b
Non-optimal
Optimal

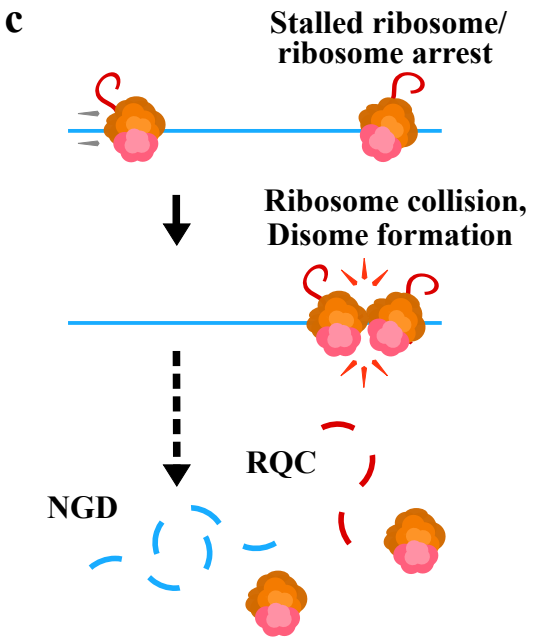
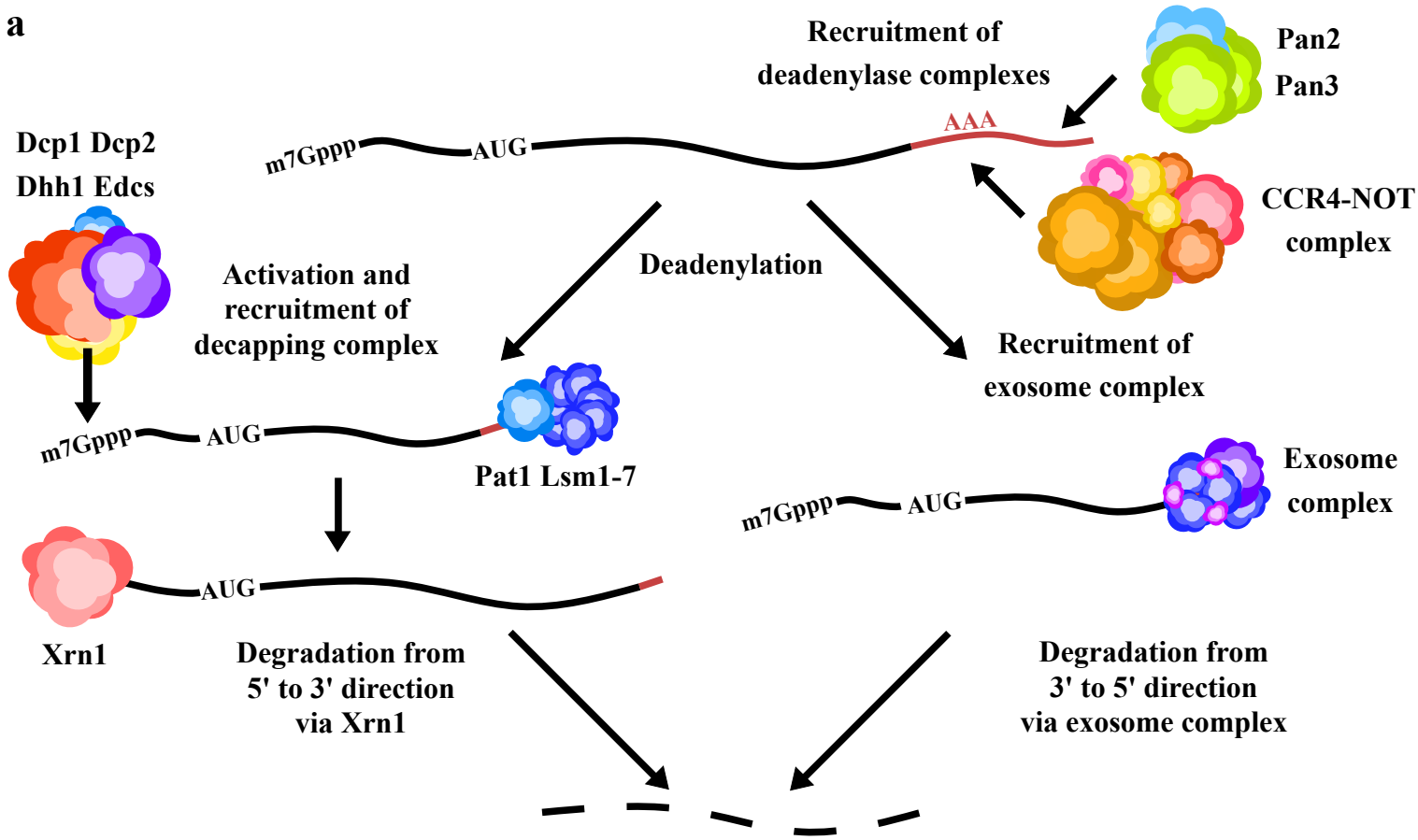
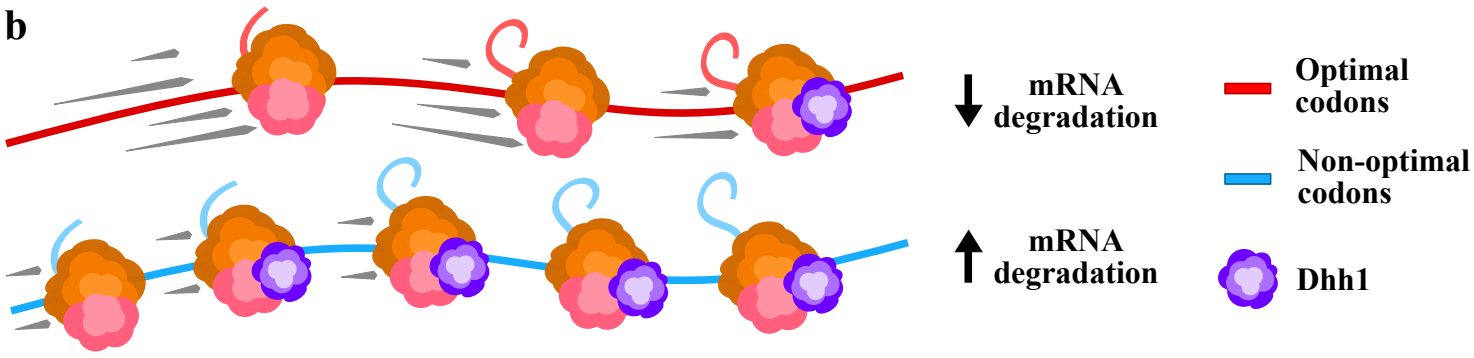

c
Stalled ribosome/
ribosome arrest
Ribosome collision,
Disome formation
RQC
NGD

Figure 3





c

	GC-rich	AU-rich
Translation	Efficient	Less efficient
mRNA Stability	High	Low
Protein abundance	High	Low

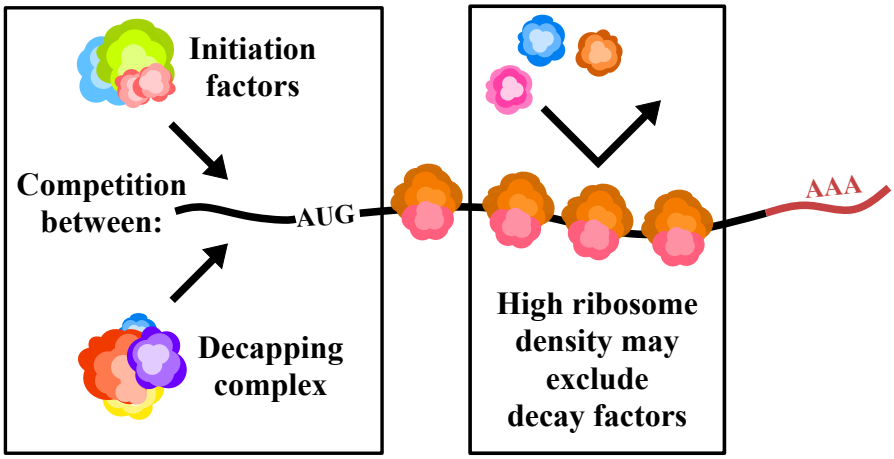


Figure 4

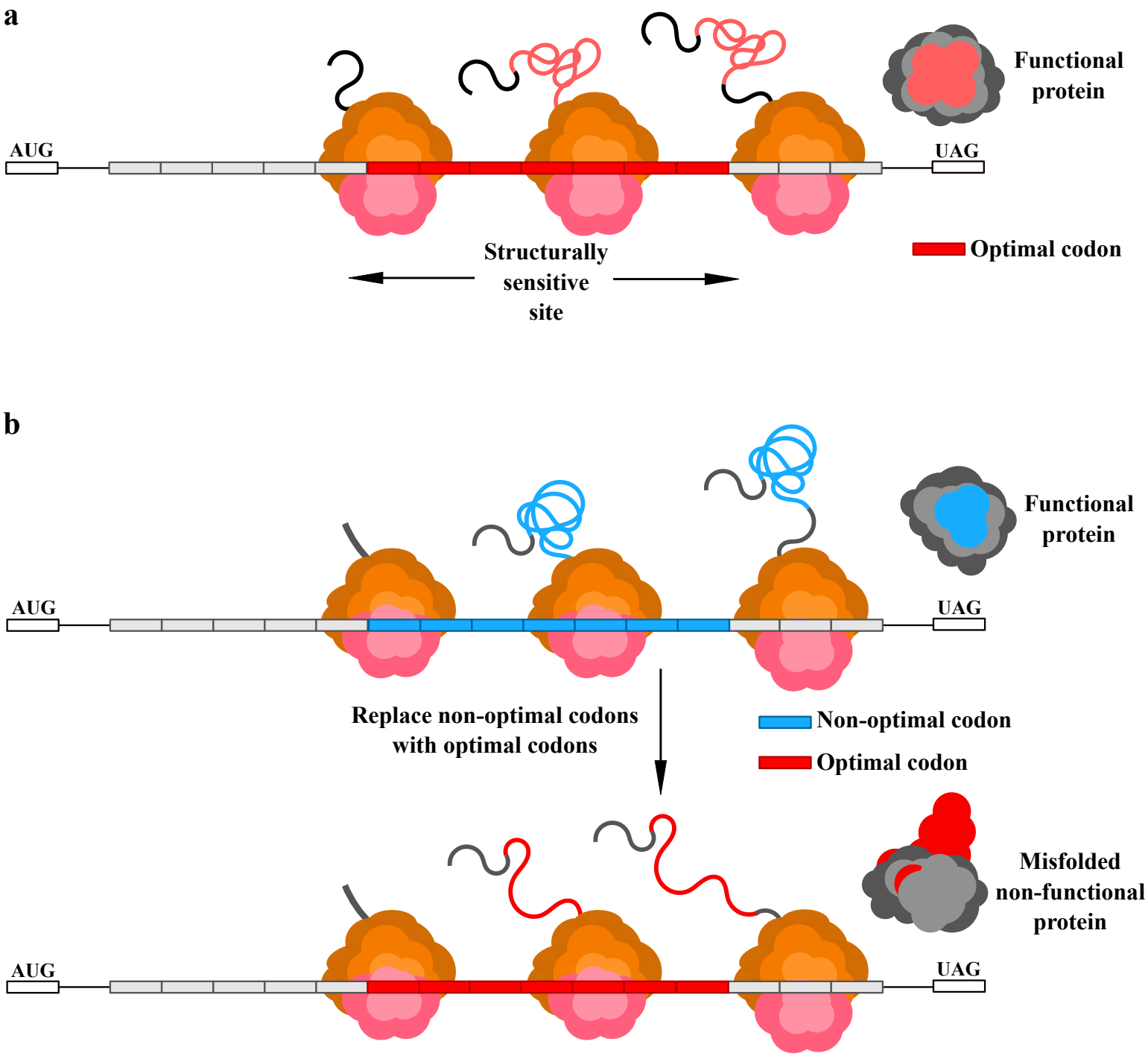

a
Functional protein
AUG
UAG
Optimal codon
Structurally sensitive site
b
Functional protein
AUG
UAG
Replace non-optimal codons with optimal codons
Non-optimal codon
Optimal codon
Misfolded non-functional protein
AUG
UAG

Figure 5

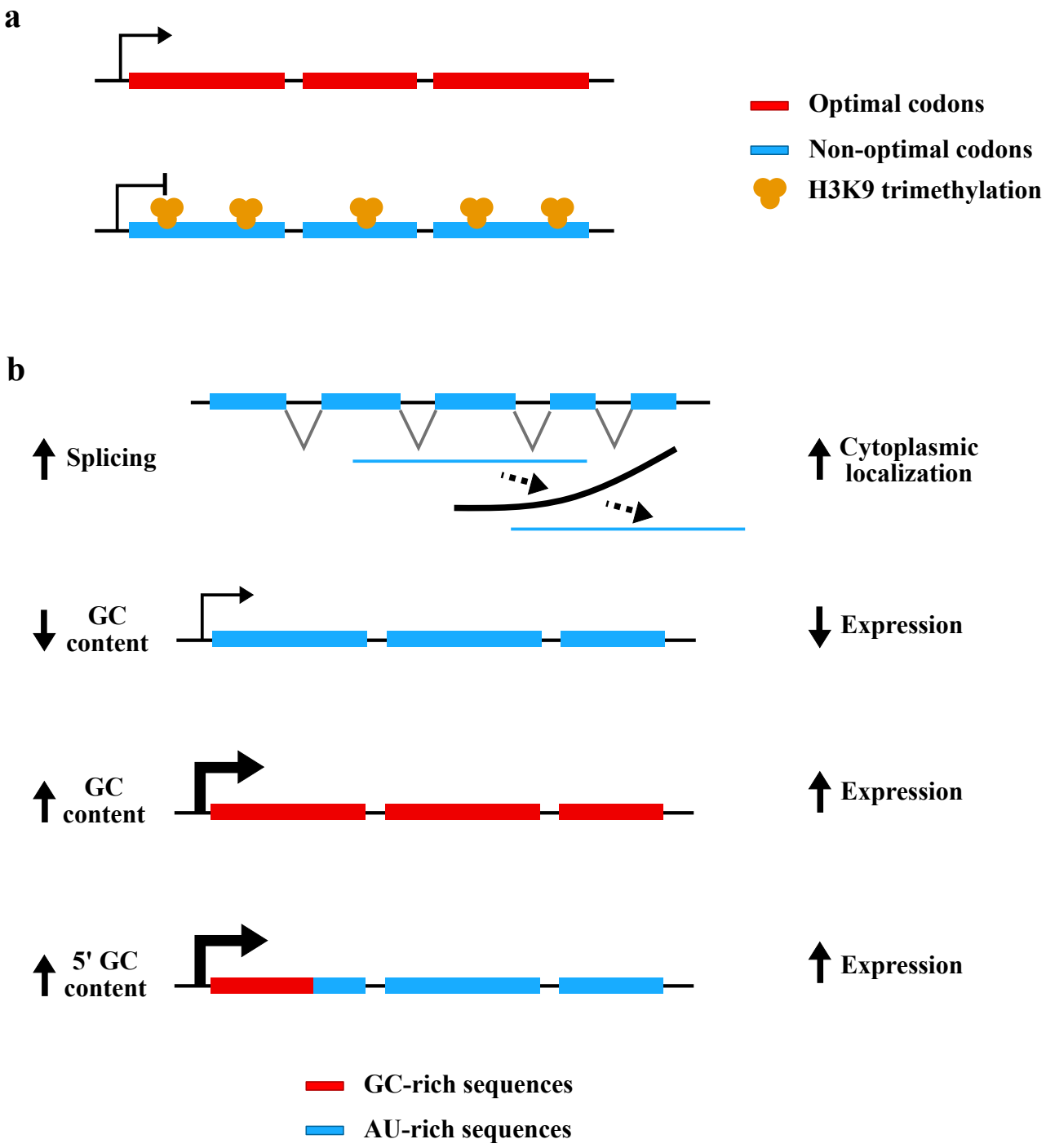

a
Optimal codons
Non-optimal codons
H3K9 trimethylation
b
Splicing
Cytoplasmic localization
GC content
Expression
GC content
Expression
5' GC content
Expression
GC-rich sequences
AU-rich sequences

Figure 6

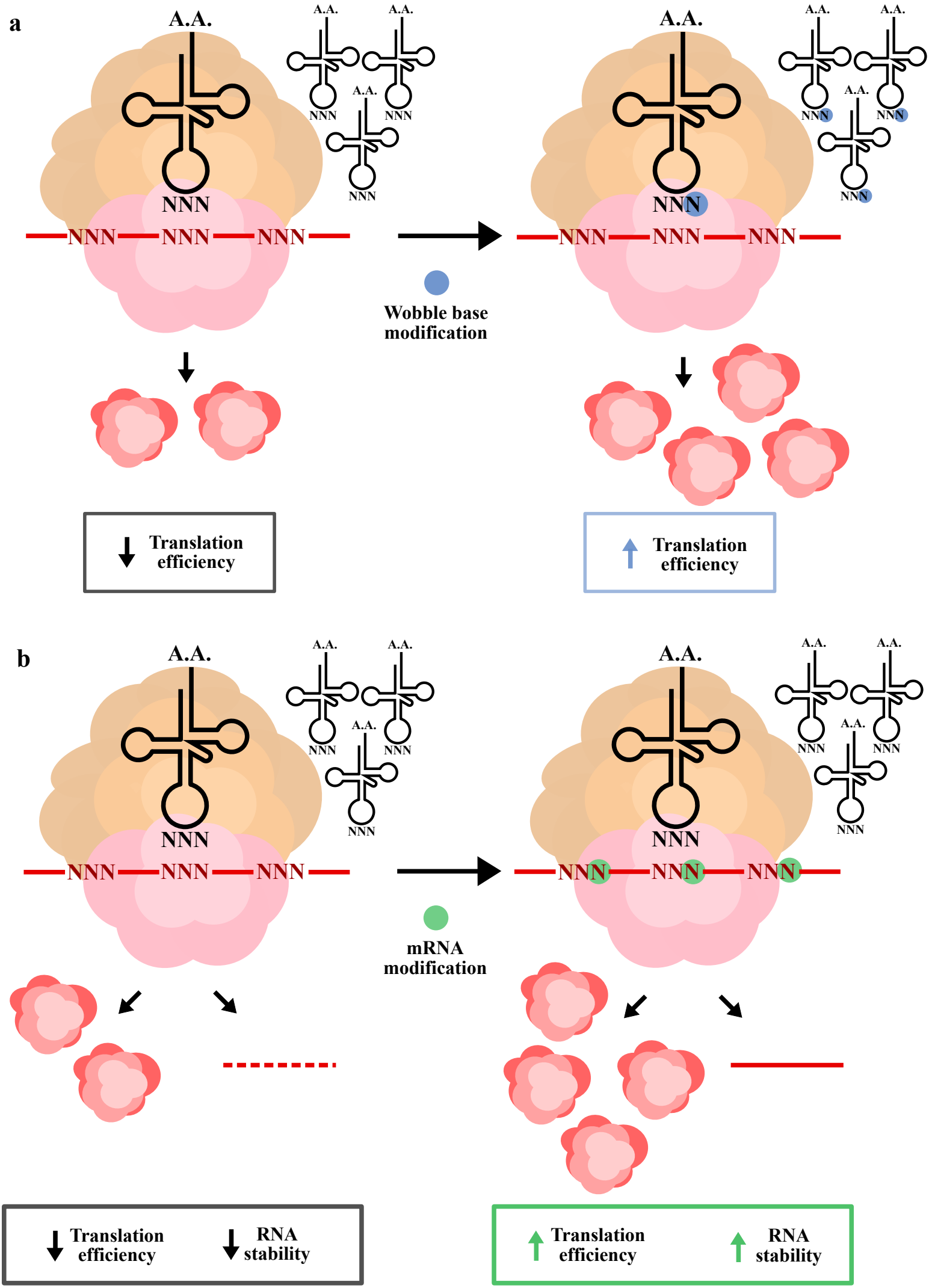

a
A.A.
NNN
NNN NNN NNN
Wobble base modification
Translation efficiency
Translation efficiency
b
mRNA modification
Translation efficiency
RNA stability
Translation efficiency
RNA stability

Figure 7

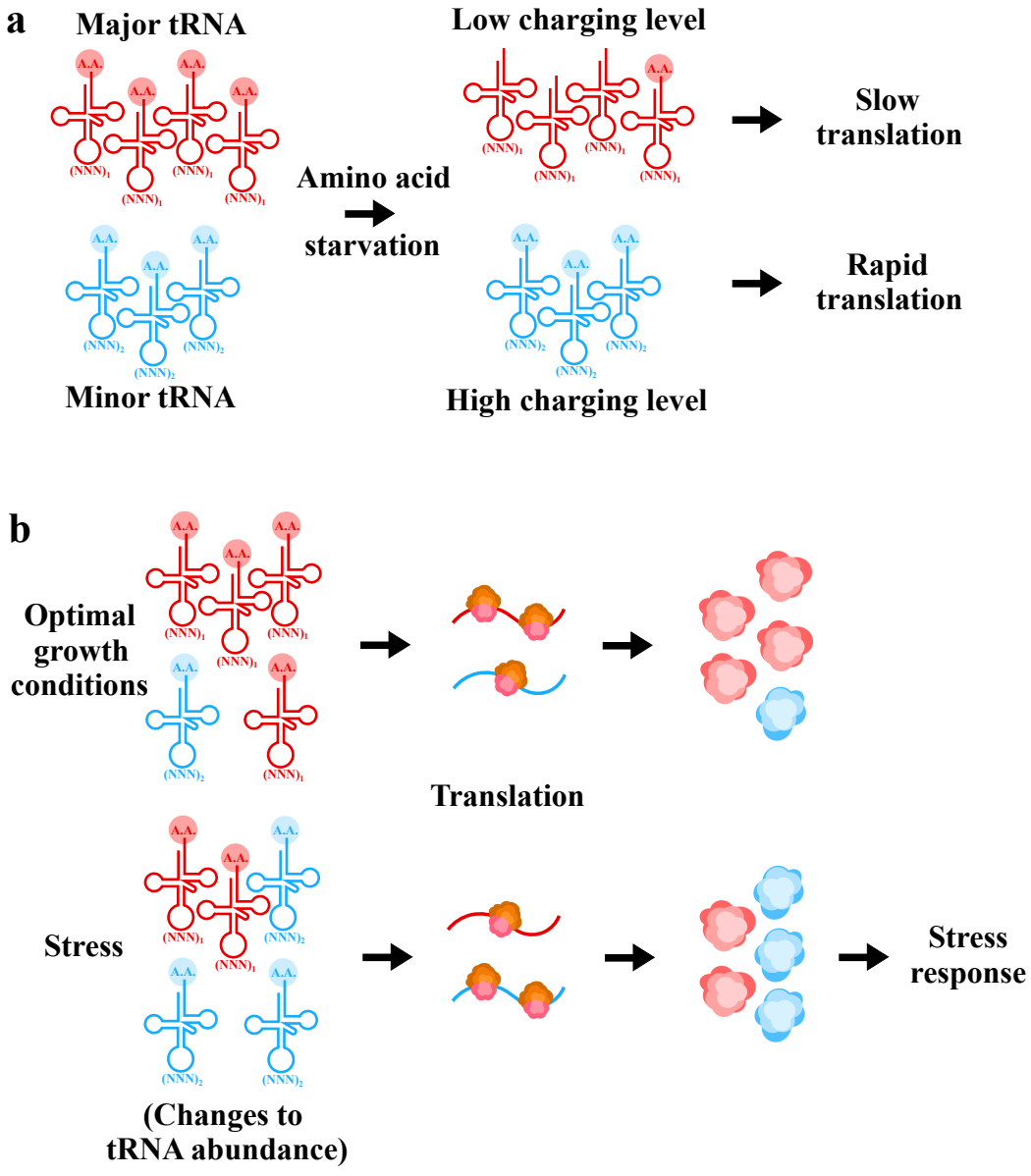
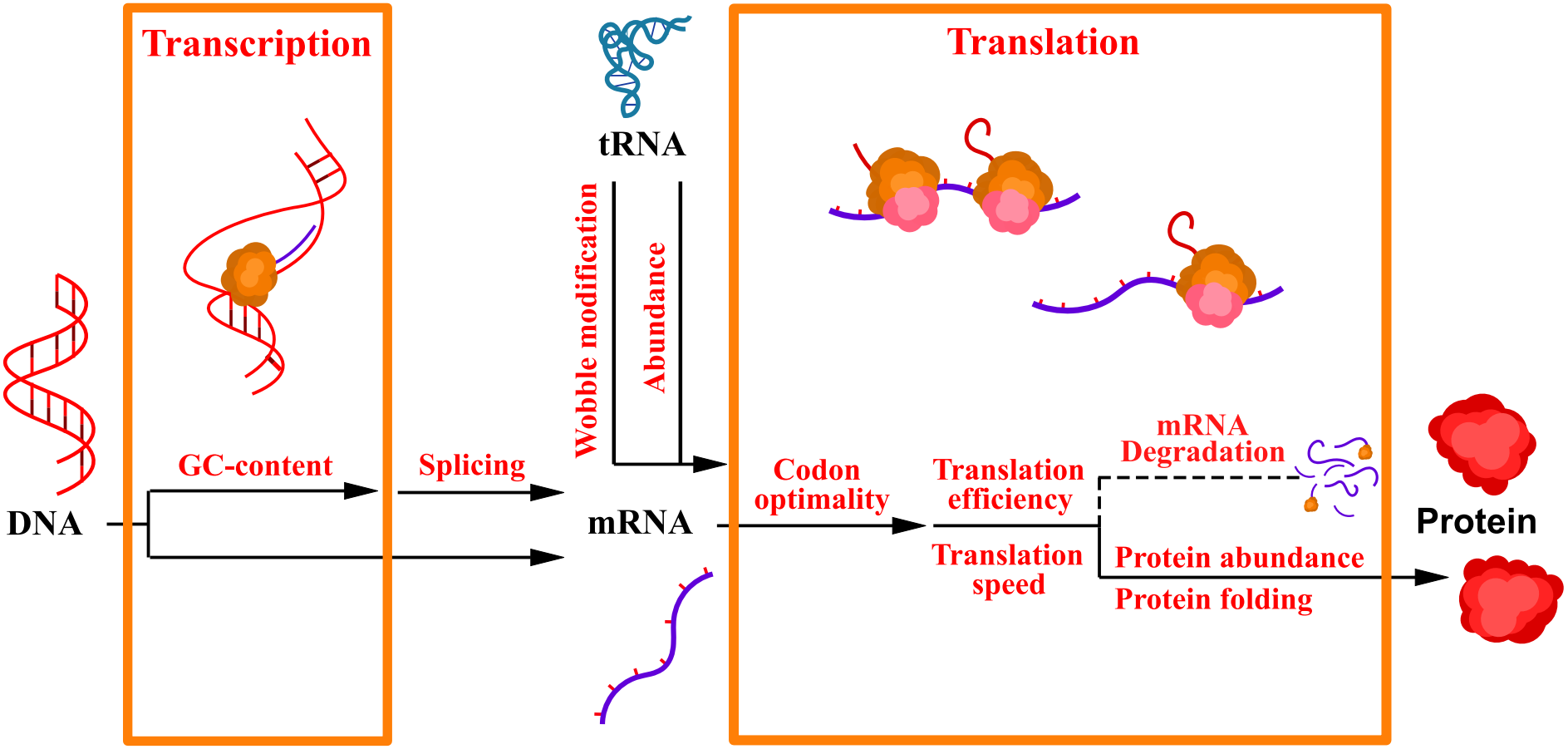

a
Major tRNA
Low charging level
Amino acid
starvation
Slow
translation
Rapid
translation
Minor tRNA
High charging level
b
Optimal
growth
conditions
Translation
Stress
Stress
response
(Changes to
tRNA abundance)

Figure 8


Transcription
Translation
tRNA
Wobble modification
Abundance
DNA
GC-content
Splicing
mRNA
Codon optimality
Translation efficiency
Translation speed
mRNA Degradation
Protein abundance
Protein folding
Protein