Charles University
Faculty of Science

Study programme: Special Chemical and Biological Programmes

Branch of study: Molecular Biology and Biochemistry of Organisms

Zuzana Hovorková

Role of small ribosomal proteins forming the decoding site in translation

Úloha vybraných malých ribosomálních proteinů tvořících dekódující místo v translaci

Bachelor´s thesis

Supervisor: Mgr. Leoš Valášek, Ph.D.

Prague, 2020

Declaration

I herein declare that I alone have composed this thesis and have appropriately acknowledged all citations using only the stated resources and literature. No other academic title has been achieved using the content of this thesis.

In Prague,

Signature

Acknowledgement

I would like to thank my supervisor Leoš Shivaya Valášek for being constantly forthcoming and willing to help me, under any circumstances, and for creating such a pleasant workspace. Huge thanks goes to all the lab members, especially to Zuzana Pavlíková and Petra Beznosková, for bearing with my never-ending questions and always being very kind, helpful and compassionate. Finally, I would like to thank my family for constantly believing in me and their endless support.

Abstract

Translation is one of the key mechanisms occurring in the cell during every second of its existence. It is a very complex process ensured by three main actors: tRNAs, mRNAs and ribosomes. Despite of being thoroughly studied over decades, the understanding of some of its functional aspects is still rather poor. This bachelor thesis focuses on four small ribosomal proteins listed below that are reaching to the decoding centre of the small ribosomal subunit. It raises awareness of the structure and function of uS12, uS19, eS25 and eS30, their evolution, role within the ribosome, and the influence they have on various stages of translation. In particular, this thesis specifically reviews the importance of these four proteins for the stop codon readthrough. This phenomenon occurs when a near-cognate aminoacyl-tRNA or a natural suppressor tRNA wins with eRF1 over the corresponding stop codon and thus protein synthesis is continued resulting in the existence of a longer protein.

It summarizes our current knowledge of its origin, molecular details of its mechanism, its existence in different species, benefits and disadvantages it brings to the life of a cell or even an organism, and finally it sums up all available knowledge for potential future use of readthrough in therapeutics.

Key words: translation, stop codon readthrough, ribosome, ribosomal proteins, uS12, uS19, eS25, eS30

Abstrakt

Translace je jeden z klíčových mechanismů probíhajících v buňce každou sekundu po dobu celé její existence. Jde o velice složitý proces, který je zajišťován třemi hlavními aktéry: tRNA, mRNA a ribosomy. I přesto, že už studován po několik desetiletí, znalosti týkající se několika funkčních aspektů jsou stále mizivé. Tato bakalářská práce se zaměřuje na čtyři malé ribozomální proteiny vyjmenované níže, které dosahují k dekódovacímu centru malé ribozomální podjednotky. Zvyšuje povědomí o struktuře a funkci uS12, uS19, eS25 a eS30, jejich evoluci, roli v ribozomu a jejich vlivu na různé fáze translace. Konkrétně tato práce specificky zkoumá důležitost těchto čtyř ribosomálních proteinů na pročtění stop kodonu. K tomuto fenoménu dochází, když blízce příbuzná aminacylová-tRNA nebo přirozeně supresorová tRNA vyhraje s eRF1 nad příslušným stop kodonem, a tak syntéza protein pokračuje vedoucí k existence delšího protein. Shrnuje aktuální poznání jeho původu, mechanismu, existence v různých živočišných druzích, přinos i nevýhody, které buňce či dokonce organismu poskytuje, a na závěr shrnuje dostupné znalosti o potenciálním využití pročtení stop kodonu v lékařství.

Klíčová slova: translace, pročtění stop kodonu, ribozom, ribozomální proteiny, uS12, uS19, eS25, eS30

List of abbreviations

A site – acceptor site

aa-tRNA- aminoacyl tRNA

ABCE1 - ATP Binding Cassette Subfamily E Member 1

ATP - Adenosine triphosphate

BCLL - B-cell chronic Lymphocytic Leukemia

BTTD - Brachycephaly, trichomegaly, and developmental delay disease

CDKN2B - Cyclin Dependent Kinase Inhibitor 2B

CTD – C terminal domain

DNA - Deoxyribonucleic acid

E site – exit site

EBNA - EBV Nuclear Antigen

EBV - Epstein-Barr virus

eEF - eukaryotic elongation factor

eIF - eukaryotic initiation factor

EM - electron microscopy

eRF – eukaryotic release factor

eS25 - Eukaryotic small ribosomal protein 25

eS30 - Eukaryotic small ribosomal protein 30

GTP - Guanosine triphophate

HIV - human immunodeficiency virus

IRES - internal ribosome entry site

LDHB - lactate dehydrogenase, subunit B

LEPROTL1 - Leptin Receptor Overlapping Transcript Like 1

L-MPZ – large myelin protein zero

LSU – large subunit

MDH1 - malate dehydrogenase

MPZ – myelin protein zero

mRNA - Messenger ribonucleic acid

NMD – nonsense mediated decay

P site – peptidyl site

PVRL3 - Poliovirus Receptor-Related 3

PTC – premture termination codon

RNA - Ribonucleic acid

RP – ribosomal protein

RPS – ribosomal protein small

rRNA - Ribosomal ribonucleic acid

SFTA2 - Surfactant Associated 2

SSU – small subunit

TC – ternary complex

TMV – tobacco mosaic virus

tRNA - Transfer ribonucleic acid

$tRNA^{Arg}$ – arginine tRNA

$tRNA^{Cys}$ – cysteine tRNA

$tRNA^{Gln}$ – glutamine tRNA

$tRNA^{i\text{-}Met}$ – initiatiator methionyl tRNA

$tRNA^{Lys}$ – lysine tRNA

$tRNA^{Trp}$ – tryptophane tRNA

$tRNA^{Tyr}$ – tyrosine tRNA

uS12 - Universal small ribosomal protein 12

uS19 - Universal small ribosomal protein 19

UTR – untranslated region

VDR – vitamin D receptor

VEGFA - vascular endothelial growth factor A

Table of Contents

1 Introduction

The central dogma of molecular biology defines the flow of information coded in DNA through RNA to proteins. The presented thesis is going to concentrate on the importance of the final part of the process, the translation, during which the language of nucleotides is converted into the language of amino acids. Even though, translation is a conserved process among all life domains (Archaea, Bacteria and Eukarya), only mechanisms in eukaryotes will be discussed.

Translation is an essential process for every living cell occuring in the cytoplasm. Its entire mechanism is very complicated and consists of four key steps: initiation, elongation, termination and recycling. During the initiation the start codon AUG is recognised and the 80S ribosomal complex is formed. The elongation proceeds by step-by-step addition of amino acids according to the specific code in mRNA thus leading to formation of the entire protein. Ribosomes represent the main functional machinery harbouring protein synthesis. They are composed of the large and small subunit and formed by rRNA and ribosomal proteins. The accuracy of translation is overlooked in the decoding centre, which comprises ribosomal RNA and specific small ribosomal proteins (RPs), namely uS12, uS19, eS25 and eS30. Strikingly, these four small RPs have influence over accurate decoding during elongation and termination, stop codon readthrough and other processes related to translation. Understanding the role of small RPs forming the decoding site in translation is crucial and of great importance for the future. Further knowledge in this area could be very beneficial mainly for people suffering from diseases caused by specific deficiencies in the mechanism of translation. Thus, the main aim of this thesis is to review the function of ribosomal proteins uS12, uS19, eS25 and eS30 during translation.

After the ribosome encounters with one of the stop codons the process is terminated, and the whole complex dissociates and reforms. However, in some cases the termination signal is ignored, and the synthesis continues. This phenomenon is described as stop codon readthrough. Although many aspects remain to be understood, there are several known factors, which influence readthrough. This is going to be another perspective considered in this thesis.

2 Translation in eukaryotes

Translation is an essential process in every cell. It gives cell the power of converting the genetic code into proteins. The main phases of translation include initiation, elongation, termination and recycling. Interestingly, the crucial factor in inaccuracy of the cellular proteom is translation of mRNA. Its error rates are by far the highest (about 10^{-4}) compared for instance to error rates of DNA replication (about 10^{-8}) [1].

2.1 Translation initiation

Translation initiation is the process of decoding the AUG start codon in mRNA by methionyl initiator tRNA ($tRNA^{i\text{-}Met}$) [2]. $tRNA^{i\text{-}Met}$ is delivered to the ribosome as a part of the so-called ternary complex (TC) together with the translation initiation factor 2 (eIF2) bound to a GTP molecule [3] in a reaction promoted by translation initiation factor 1, 1A, 5 (eIF1, eIF1A, eIF5), and the multisubunit translation initiation factor 3 (eIF3) [2]. Completion of this step results in a formation of the so-called 43S pre-initiation complex (PIC) [3], which attaches to the 5´end of mRNA–preactivated by association with translation initiation factor 4F (eIF4F) at the 5´ 7-methyl guanosine cap and poly(A)-binding protein bound to the poly(A) tail–to form the 48S PIC [2]. Another co-operative role of eIFs 1, 1A, 3 and 5 is to prepare the small ribosomal subunit for mRNA docking by opening the 40S mRNA binding channel [3]. The 5´untranslated region (5´UTR) is then scanned base-by-base for complementarity to the anticodon of $tRNA^{i\text{-}Met}$ as successive triplets enter the peptidyl (P) decoding site of the 40S subunit [2] till the start codon (usually the first AUG) has been recognized. Scanning requires the action of helicases such as eIF4A and DHX29 to unwind mRNA secondary structures for the ribosome to move smoothly [3]. Nevertheless, the first AUG can be skipped if its surrounding sequence context does not conform to the 'Kozak' consensus, featuring a purine at position -3 and guanine at +4 relative to the AUG (at +1) [2]. The AUG recognition triggers a series of intricate events and conformational changes in the PIC involving irreversible GTP hydrolysis [3] in the TC and release of eIF2–GDP from $tRNA^{i\text{-}Met}$, which is followed by joining of the large (60S) subunit, stimulated by eIF5B, to form an 80S initiation complex [2]. Ejection of eIF5B hand in hand with eIF1A marks the end of the initiation phase leaving the 80S initiation complex behind poised for elongation [3].

2.2 Translation elongation

In elongation individual amino acid residues are connected via the peptide bond in order to make the primary structure of the full-length protein. An elongation cycle can be subdivided into three steps: (1) delivery of aminoacyl-tRNA (aa-tRNA) to the acceptor (A) site, which involves decoding and accommodation, (2) peptide-bond formation in the peptidyl (P) site, and (3) tRNA translocation to the exit (E) site. Translational GTPase elongation factors help to overcome the energy barrier and together are responsible for speed and accuracy of protein synthesis [4]. eEF1A is

a GTPase mediating the recruitment of aminoacyl-tRNAs to the A- site of the elongating ribosome, eEF1B is GEF for eEF1A and eEF2 is a GTPase promoting translocation of the 80S ribosome by one triplet at a time, the purpose of which is to add one amino acid residue per each triplet following AUG into the growing polypeptide chain [3].

In eukaryotes, translation initiation culminates with formation of an 80S initiation complex in which $tRNA^{i\text{-}Met}$ is bound in the P site of the ribosome. The anticodon of the $tRNA^{i\text{-}Met}$ is base-paired with the start codon of the mRNA [5]. The second codon of the open reading frame (ORF) is present in the A site of the ribosome awaiting binding of the cognate aminoacyl-tRNA according to the decoding rules (further discusses in chapter 3.3.1) [6]. Moreover, it is important to add that the anticodon loop is used to detect the appropriate codon and therefore also the transfer of the correct amino acid onto the growing polypeptide chain.

The eukaryotic translation elongation factor eEF1A, like its bacterial ortholog EF-Tu, is activated [5] and binds aminoacyl-tRNA in a GTP-dependent manner forming a TC and then directs the tRNA to the A site of the ribosome [6]. Such position of tRNA•eEF1A•GTP complex within ribosome is reffered as hybrid A/T state. Furthermore, it is known that during this state tRNA is bent in its anticodon stem approximately at position 42 [7]. Base pairing interactions between the anticodon of the aminoacyl-tRNA and the A-site codon trigger GTP hydrolysis by eEF1A. The eEF1A•GDP complex is released and the aminoacyl-tRNA is accommodated into the A site and transits to the A/A state [5].

Following accommodation of the amino-acyl-tRNA into the A site, peptide bond formation with the P-site peptidyl-tRNA occurs rapidly [6]. Methionine is transferred from the initiator P site tRNA onto the next A site aminoacyl-tRNA and thus a peptide bond is formed between the carbonyl group of methionine and the α-amino group of the next amino acid. This is catalysed by the peptidyltransferase activity of the large ribosomal subunit [5].

Following peptide bond formation, ratcheting of the ribosomal subunits triggers movement of the tRNAs into so-called hybrid P/E and A/P states with the acceptor ends of the tRNAs in the E and P-sites and the anticodon loops remaining in the P and A sites, respectively. Elongation factor 2 (eEF2) is required for the translocation of the tRNAs to the canonical E and P sites[6].

Simultaneously the ribosome moves three nucleotides downstream and therefore it is “reset” with free A site capable of accepting a new aa-tRNA according to the next codon-anticodon interaction. This happens with the help of eEF1A and the whole cycle repeats until the last amino acid is incorporated into the polypeptide chain.

Interestingly, the speed of elongation of different polypeptide chains can vary [8], which, beside other factors, can be affected by codon bias [9].

2.3 Translation termination

Termination is the final step of protein synthesis, when one of the three stop codons (UGA, UAG, UAA amino acid) is present in the A-site. Importantly, translation termination is not merely determined by the presence of the stop-codon. In fact, the nucleotide sequences both upstream and downstream, in close proximity or even from great distance, can strongly affect the decision to terminate protein synthesis.

Termination in eukaryotes is catalyzed by two protein factors, eRF1 and eRF3, that appear to collaborate in the process and enter the decoding site of the 80S ribosome in a joined complex. The class I factor, eRF1, is responsible for high-fidelity stop codon recognition [6].eRF1 has an amino-terminal domain (N) that extends into the 40S subunit´s decoding centre and is responsible for the recognition of the stop codon in the A site [6], [10]; a middle domain (M) containing a universally conserved apical GGQ (glycine-glycine-glutamine) motif that enables hydrolysis of the peptidyl-tRNA bond [11] and induces release of the nascent polypeptide from peptidyl-tRNA in the ribosomal P site through the addition of a water molecule rather than the amino acid [11]; and a carboxy-terminal domain (C) that binds to eRF3 and ABCE1,and contains a mini-domain that affects stop codon specificity [6], [12]The class II factor, eRF3, is a translational GTPase. It has a variable amino terminus and a more conserved carboxyl terminus that is directly involved in interactions with eRF1 [6]. eRF3 binds to the GTPase-associated centre (GAC), between the sarcin–ricin loop (SRL) of the 60S subunit and helices h5 and h14 of 18S ribosomal RNA (rRNA) on the 40S subunit [6], [13].

eRF1 enhances binding of GTP to eRF3 by acting as a GTP dissociation inhibitor, promoting formation of a stable eRF1/eRF3•GTP complex [6], [14]. The ternary complex, eRF1:eRF3:GTP, next engages the ribosome, triggering GTP hydrolysis [6], [15] ultimately leading to the deposition of the M domain of eRF1 in the peptidyl transferase centre. The presence of a stop codon in the A site can be evaluated by eRF1 [6].

2.4 Translational readthrough

Under certain circumstances, translational readthrough (also known as nonsense suppression) might occur. During this process near-cognate aa-tRNAs can be incorporated at the stop codon in the ribosomal A site instead of the release factor. Proteosynthesis then terminates on the next stop-codon and results in a C-terminally extended protein isoform.

Readthrough can be spontaneous happening with a frequency less than 0,1%. Programmed or functional translational readthrough can reach even 20% and has firstly been observed in viruses in order to create C-terminus extensions on proteins and thus increasing their diversity [16], [17].

Significantly, readthrough may occur also at a premature termination codon (PTC) that is a cause of existence of many diseases. Readthrough of a PTC allows insertion of a near-cognate tRNA

into the ribosomal A-site, leading to synthesis of a full-length protein from otherwise defective mRNA [18], the degradation of which by the nonsense-mediated decay (NMD) pathway is thus avoided. Nowadays a few readthrough-inducing drugs are in use to treat PTC-related diseases [19] (also discussed in chapter 2.4.4).

2.4.1 Factors influencing readthrough

Each stop codon is prone to termination to a different extent. The potential for readthrough is highest for the UGA stop codon which has the lowest termination fidelity. The likelihood for readthrough is lower for the UAG codon and the lowest for the UAA codon [20]. However, it has been shown that not all near-cognate tRNAs can be incorporated at the stop codons. UAG or UAA stop codons can be read through after the incorporation of $tRNA^{Gln}$, $tRNA^{Lys}$ or $tRNA^{Tyr}$, although $tRNA^{Gln}$ has been proven to be the most efficient and $tRNA^{Lys}$ the least efficient [21]–[23]. On the other hand, the UGA stop codon can be decoded by $tRNA^{Arg}$, $tRNA^{Cys}$ or $tRNA^{Trp}$ [23]. However, $tRNA^{Arg}$ was incorporated at very low rates.

It has been proven that the role of many translation factors influences readthrough. Some of them will be discussed in the following chapters. These include the identity of the base immediately following the stop codon, as that has been shown to be one of the most important determinants for readthrough [24]. Furthermore, the consensus of several bases placed subsequently of the stop codon also affects readthrough [25]. On the other hand, termination factor eRF1 [26] but also initiation factor eIF3 [27] have both been shown to participate on the event of readthrough. Finally, the role of glutamine tRNA will be overviewed [21], [22].

2.4.1.1 *The effect of the +4 base*

Even though the exact molecular mechanism behind specific amino acid choices at different stop codons remains unexplained, it has been shown that the crucial influence on the strength of the termination signal is given by the nucleotide located immediately after the stop codon [24]. It has been observed that when UGA is followed by adenine, this tetranucleotide is most likely read through by tryptophane, while when the codon is followed by C or G, cysteine is usually incorporated [26]. As a matter of fact, it seems that positioning of C, as the +4 nucleotide behind any stop codon, induces readthrough most strongly altogether, followed by A, G and U [27]. On the other hand, incorporation of $tRNA^{Tyr}$ is the most efficient at all four possible UAA-N tetranucleotides. It has been shown that the ψ35 of tyrosine tRNA is the most important for the incorporation of tyrosine both at UAA and UAG without regard to the +4 nucleotide [28]. As assumed, UGA-U does not show any preference for either of the near-cognate tRNAs as it has the lowest readthrough rates [26].

2.4.1.2 *The role of the surrounding sequences*

Besides the role of the +4 base, it has been proven that specific groups of nucleotides, which together promote readthrough, also exist. [25]. For instance, in plant cells the consensus sequence CAR YYA (where R = A/G, Y = C/U) from the tobacco mosaic virus (TMV) induces readthrough at all stop codons [29]. Moreover, the UGA_CUAG consensus is also a common motif prone to readthrough throughout all vertebrates [16]. Furthermore, it seems that the last two incorporated amino acids can also influence readthrough [30].

Valášek et al have also concluded that it is highly possible that the tRNA molecules alone carry some distinct feature that is partly responsible for decoding of the stop-codon, probably by being aware of the nucleotides in close proximity of the stop codon [26]. This especially applies for $tRNA^{Cys}$ and $tRNA^{Trp}$, where *Valášek et. al* have found out that neither the modified geometry of the decoding site nor any noncanonical influence of eRF1 have effect on the choice of near-cognate tRNAs in any way [26].

In any case, all these findings expand current knowledge of the decoding rules (also discussed in chapter 3.3.1).

2.4.1.3 *Role of the initiation factor eIF3 in readthrough*

It has long been known that the eukaryotic initiation factor 3 (eIF3) is one of the key players in translation initiation and all its very diverse roles have been studied by many researches, e.g.: [31]–[37].

Surprisingly, it became later evident, that independently of translation initiation, eIF3 also takes part in other translational processes such as translation termination, readthrough or reinitiation [38]. It has become clear that through the eIF3 interference proper decoding of the third stop codon position by the ribosome and/or eRF1 is impaired resulting in premature rejection of the eRF1•eRF3•GTP complex from pre-termination complexes [27]. Moreover, the eIF3-mediated altered property of the third position would promote wobble miscoding by near-cognate tRNAs (specifically $tRNA^{Tyr}$ for UAA or UAG and $tRNA^{Cys}$ and $tRNA^{Trp}$ for UGA) at this specific position to occasionally win the battle over the stop codon predisposed for programmed readthrough, incorporate into the A-site and allow elongation to continue to the next stop codon downstream [27].

eIF3 also influences readthrough through the interaction with residue K108 of the universal small ribosomal protein 3 (uS3/RPS3). *Poncova et. al* postulated that:" We speculate that K108 functions to sense the presence of eRF1 in the pre-termination complex to timely close the 'termination' latch (formed by Rps3 and 18S rRNA helices 18 and 34). We propose that by contacting K108, eIF3 modulates this role in such a way to delay the latch closure to promote programmed stop

codon readthrough. If K108 is substituted with aspartic acid, the functional impairment of the Rps3–a/Tif32 contact, and perhaps also of that between Rps3 and h34, weakens and/or alters the influence of eIF3 and Rps3 over eRF1 and thus allows the termination-specific constriction of the latch to occur 'prematurely'. This would in turn decrease readthrough, as was observed."[39].

2.4.1.4 *Role of eRF1 in readthrough*

The efficiency of readthrough is determined by the competition taking place in the A site of the ribosomal decoding centre between near-cognate tRNAs longing for incorporation and eRF1 and eRF3. As previously stated, the presence of cytosine at the +4 position affects readthrough. Consistently, the essential K residue from the critical TASNIKS motif in the N-terminal domain of eRF1 was in a recent structural study proposed to be located in the proximity of the +4 nucleotide [26], [40] (des Georges et al. 2014) and thus this interaction would also affect readthrough. Furthermore, results of *Beznosková et. al* give the impression that the contact between the K residue and cytosine at the +4 position is qualitatively different from contacts with the other three bases in the same position, which makes it the least stringent base for efficient termination [26].

2.4.1.5 *Interesting role of glutamine tRNA in readthrough*

It has been proven that glutamine tRNA ($tRNA^{Gln}$) is the most efficient at readthrough at the UAG and UAA stop codons [21], [22]. It is generally known that *Saccharomyces cerevisiae* has 10 genes for $tRNA^{Gln}$ which code for several isoforms. However, only the M form carries the anticodon with the CUG sequence pairing with the UAG termination codon. All the other isoforms pair their UUG anticodons to the UAA termination codon. Another interesting thing is that the tQ(UUG)B differs from tQ(CUG)M in six nucleotides, while tQ(UUG)L from tQ(CUG)M only in three nucleotides. This is why the findings of a recent study are so interesting, as they show that only the M isoform induces notable rates of readthrough [28] at all four UAG-N tetranucleotides, meaning that it is very likely that solely the minor variations in these few nucleotides predispose tQ(CUG)M to such significant rates of readthrough.

2.4.2 Significance of readthrough for cells

Readthrough is a clever example of protein expression regulation. By being able to pass the stop-codon, organisms add extra functional domains on proteins, while continuing translation on the same mRNA and in the same reading frame. Molecules with extended C terminus might have slight or even important alterations in their functions in comparison to the mother protein. These extensions often comprise domains important for physiology of the cell. Some protein alterations even take part in the metabolism [17] or on the contrary, are pathological [41]. This can have a profound impact on the whole cell or even organism. The extensions created in the aftermath of

readthrough often contain targeting signals such as to the nucleus or peroxisome, furthermore, transmembrane domains or even prenylation signals [17], [42].

Readthrough can be evolutionarily regarded also as an experimental mechanism of how the cell can try out new functions on an existing protein, without having to develop a new one entirely. Another thing is, from the evolutionary perspective, that the C-terminus is generally more prone to being modified throughout evolution, rather than the protein core [43]. Also the C-terminus is probably preferentially chosen for modifications compared to the N-terminus, as the gene regulatory sites are more likely to be situated on the N-terminus [43].

Yet another function could be the protection against the NMD pathway. It has become apparent that even low readthrough rates can seriously prevent mRNA from being degraded by NMD. Interestingly, it happens to be the case in yeast, mammals and also plants [44]–[46].

The truth is that as readthrough has a role in many cell conditions, so its existence can hardly be considered a pure coincidence.

2.4.3 Readthrough in organisms

Readthrough has been observed in viruses but also in many other different types of organisms. As no studies have been conducted questioning this fact it can be a mere coincidence or maybe this event is somehow conserved among organisms including yeasts but also mammals.

Viruses are among the smallest entities carrying their own genetic information. As their genome usually encompasses only a few hundred bases, they have developed several mechanisms to benefit the most from it. One of these methods seems to be readthrough, through which the coding capacity is expanded. Some examples of viruses capable of readthrough include the tobacco mosaic virus or the Sindbis virus, both alternating their RNA replicase [47].

Phosphodiesterase 2 in yeast is extended by 22 amino acids after being subjected to readthrough. Consequently, the protein is sentenced to degradation dependent on the proteasome instead of being translocated to the nucleus. Furthermore, higher cAMP levels in the cell are reached as the enzymatic activity of the protein is reduced, which consecutively has effect on stress response [48].

Readthrough in Drosophilla was observed during different stages of development in 13 genes including Abd-B, cnc, kel, Sp1, sync and z [49]. The developmental state seems to be the main factor affecting readthrough [50]. Protein extensions created as a result of readthrough seems to contain functional domains [49].

Another protein extended by 22 amino acids is the rabbit beta-globin [51]. Yet, it is not likely that it is conserved in other mammals. On the other hand, the myelin protein zero (MPZ) is conserved from amphibians to humans and contains antigenic sites for neuropathy-associated

antibodies. It has been found to be the basis of the large myelin protein zero (L-MPZ) which is created due to readthrough. Mutations in this protein can result in the Charcot-Marie-Tooth disease or the Dejerine-Sottas disease [52].

It is quite understandable that higher eukaryotes are generally less prone to readthrough than evolutionary older organisms which have smaller genomes and therefore are in demand of maximizing their coding potential.

To this day several human genes capable of readthrough have been found including: CDKN2B, MDH1, LDHB, LEPROTL1, PVRL3, SFTA2 and VEGFA [53], [54]. VEGFA codes the vascular endothelial growth factor A with mainly proangiogenic function. Due to readthrough occurring with the frequency of 7-25%, the protein is extended by 22 amino acids and its function turns antiangiogenic [54]. Concerning the MDH1 (malate dehydrogenase) and LDHB (lactate dehydrogenase, subunit B), human cells make use of readthrough in order to create C-terminal extensions containing peroxisomal targeting signals [17], [42]. It has become apparent that MDH1 is one of the strongest readthrough genes with readthrough rates from 3-4% [16], [17]. It is followed by a sequence present also in other mammals (UGA CUAG) which has high readthrough rates even in non-inducing conditions [17]. In this specific situation, the formation of a malate/lactate shuttle across the membrane of the peroxisome is enabled. [55]. Another protein which has an altered function due to readthrough is the vitamin D receptor (VDR) [56].

The fact that readthrough produces C-terminally extended proteins with full or altered functions has led to the thought of the development of readthrough-inducing drugs as discussed in the following chapter.

2.4.4 Readthrough-inducing drugs as treatment of PTC-casued diseases

Under undefined circumstances a nonsense mutation can occur in the protein-coding gene sequence. This mutation represents change of a sense codon to a stop codon also known as premature termination codon (PTC). During translation the occurrence of such PTC leads to premature abortion of translation and thus to the production of a truncated protein. This can contribute to later deviations in protein functions which can have even deadly consequences.

According to The Human Gene Mutation Database [57] as much as 11,2 % of all mutations linked to human inherited diseases are caused by nonsense mutations [58]. Examples of nonsense-associated genetic disorders include for example cystic fibrosis, haemophilia, Duchenne muscular dystrophy, many forms of cancer, spinal muscular atrophy or even obesity. However, these are only a few among many [59].

The knowledge of the readthrough event, leading to amino acid incorporation at stop codons, has shed new light on devising promising approaches which would be successful in tackling PTC-

driven diseases. By inducing readthrough at the PTC via genetic manipulations or specific drugging of a cell a full-length protein would be created and thus the probability of nonsense-mediated disorders would be eliminated. One technique devised in the sixties is the application of aminoglycoside antibiotics targeting the ribosomal decoding centre. Among the most commonly used is gentamicin [60]. As these antibiotics preferentially bind to the prokaryotic ribosomes, they interfere with the decoding reactions of the mitochondrial ribosome in eukaryotes leading to side effects which beyond all doubt pose a great risk. Therefore the use of such pharmacological agents is not desirable [59]. Nevertheless, research with the aim of developing safe aminoglycoside derivatives is still being conducted and a step-by-step progress is being made [61].

2.4.5 Methods of readthrough analysis

Many labs studying readthrough and related molecular events have developed a vast array of methods to do so.

On the systems-level we recognise at least three different approaches. Ribosome profiling, which identifies the mRNA in an actively translating ribosome [62], [63] and can possibly uncover readthrough as well as other recoding events outside of coding regions [50], [62], [64]. It is a technique based on deep-sequencing of ribosome-protected mRNA fragments, that makes it possible to monitor translation with a depth, speed and accuracy that rivals existing approaches for following mRNA levels and has the potential to identify protein-coding regions [62]. Through this method it was possible to identify hundreds of readthrough candidates in *Drosophila melanogaster*, but also some in yeast or human skin fibroblasts [50]. However, this technique is less appropriate for discovering readthrough in genes with low expression rates or in long extensions as it may be harder to distinguish readthrough from reinitiation. Further there is the phylogenetic approach which estimates the coding potential of genomic sequences based on a multiple alignment of related species. It is suitable for genome-wide searches for unknown coding regions and potential readthrough candidates, because it incorporates and uses prior information of a genome. The phylogenetic approach was also able to uncover many candidates of readthrough genes, especially in *Drosophila melanogaster* [49], [65]. However, this technique may not be suitable for studying small extensions as they might not be sufficient for comparison. Finally, the in-silico regression model is a computational approach that concentrates on the stop codon context in order to discover readthrough in genes. The disadvantage is that it focuses purely on the context of the stop codon failing thus to uncover any other forms of readthrough (for example induced by structure motives) or other recoding events in the 3´UTR. By using this technique, it was possible to uncover over 50 readthrough candidates, some of which have been confirmed [16], [17], [42].

Readthrough efficiency can also be measured by enzymatic reporter systems. One of the most commonly used is the luciferase-based reporter assay, specifically the dual luciferase assay. The principle of this method relies on the stop codon between the firefly and the renilla luciferase genes [66]. Readthrough efficiency is quantified as the ratio of the two light signals (firefly/renilla) above background as firefly, located downstream of the stop codon, can only be expressed if readthrough occurs [66]. Another similar dual reporter assay also uses the combination of luminescence and fluorescence, this time using the the venus/luciferase measurement [17]. Other types of reporter assays use for example the green fluorescent protein (GFP) [67]. In any case it is important to think about the main inducer of readthrough, as it may alter the function of the reporter, so the experiment must always contain several controls in order to avoid the misinterpretation of the results [68].

Last but not least, in order to study readthrough in yeast a new very efficient method was developed called YARIS (Yeast Applied Readthrough Inducing System). This approach is based on systematic examination of readthrough at multiple sequences in distinct genetic backgrounds and under various environmental conditions [28]. Owing to this method, it was possible to monitor readthrough efficiency of 12 distinct stop tetranucleotides (stop codons followed by one of the four bases), in the presence of elevated levels of all near-cognate tRNAs. and describe the decoding rules for readthrough of all 12 tetranucleotides by defined near-cognate tRNAs. This way a novel set of the so-called readthrough-inducing tRNAs (rti-tRNAs) was identified. YARIS was also used to examine the impact of particular tRNA modifications on the effectiveness of readthrough [28].

Either way, the final determination of readthrough in genes should not be based only on one of these methods, but always as a combination of several different approaches.

3 Structure and function of ribosome

Within every living cell it is possible to find ribosomes, ribonucleoprotein complexes that serve as tools capable of translating genes into proteins. There can be up to millions of ribosomes found either bound to the rough endoplasmic reticulum or free in the cytoplasm within a single cell. The importance of the ribosome is highlighted by the fact, that all three domains of life (Bacteria, Archaea and Eukarya) have many key components conserved [69].

A lot of the knowledge of ribosome structure comes from cryo-electron microscopy (EM) [70], [71] and X-ray crystallography [72]–[75]. The main structural core of the ribosome is conserved in both bacterial and eukaryotic ribosomes and is made up of 4400 RNA bases and 34 proteins which form the major functional sites of the ribosome including the tRNA binding sites (A-site), peptidyl transferase centre and decoding site (P-site) [71], [76]–[78] and the E-site where tRNA dissociates from the ribosome.

The eukaryotic ribosome is composed of two subunits, the large 60S (LSU) and the small 40S (SSU), together creating the 80S complex, ranging in sizes from 3,5 MDa in lower eukaryotes to 4,0 MDa in higher eukaryotes. The interactions between the small and the large ribosomal subunits are mediated through seven intersubunit bridges [72], [73], [79], [80]. 80S is made up of approximately 79 proteins out of which 32 have not been discovered in archaeal or bacterial ribosomes, and around 5500 RNA bases creating 4 rRNA chains. In eukaryotes five specific expansion segments [81] and five variable regions are present on the small subunit and 16 expansion segments and two variable regions on the large subunit [82].

The structure of the 80S ribosome illuminates the precise architecture of eukaryote-specific intersubunit bridges, the conformational changes of rRNA and proteins upon relative movements of ribosomal subunits [69].

The main function of the ribosome is carried out by three major centres. The A site, which contains the decoding centre, where aa-tRNA is bound by complementarity of its anticodon to the codon in the mRNA sequence, the P site, where the peptide bond is formed between individual amino acids brought by different tRNAs, and the E site, where the deacylated tRNA leaves the ribosome.

Stress and various other growth conditions can influence the content of the ribosome even within a single species [83]–[86]. One very impressive example is the Plasmodium falciparium, a human parasite causing malaria, which changes the composition of the ribosome depending on the stage of the life cycle [87].

The small subunit is responsible for the decoding process where aminoacyl tRNA is selected according to the mRNA sequence [69].. On the other hand, the LSU serves as the centre for protein synthesis as it accommodates the peptidyl transferase centre which enables the formation of the peptide bond and the polymerization of the peptide chain. The nascent protein exits the ribosome via a tunnel in the LSU [88].

The mutations in DNA sequences coding for ribosome components or in proteins important for the biogenesis and maturation can cause diseases referred to as ribosomopathies [89].

3.1 Large subunit of eukaryotic ribosome

The large 60S subunit has a crown-like shape and contains the central protuberance and the P stalk. This subunit contains 27 specific proteins for eukaryotes as well as several insertions, extensions of conserved proteins or rRNA expansion segments that are located mainly on the periphery creating a ring-like shape around the core [69]. The peptidyl transferase centre and the three tRNA binding sites (A, P and E) are located on the interface side. The peptidyltransferase centre is adjacent to the entrance of a tunnel along which nascent proteins progress before they emerge

from the ribosome on the solvent side [69]. In spite of the fact that the major functional centres of the ribosome are conserved among all life domains, there are, however, important structural differences between the bacterial 50S and the eukaryotic 60S subunits for example, in the organization of the peptide tunnel and the surrounding area, which can be understood in terms of functional divergence [69].

3.2 Small subunit of eukaryotic ribosome

The 40S subunit consists of 32 proteins and 18S rRNA, which together form parts known as the shoulder, beak, body, head and platform (Fig. 1) [69]. The locations of the binding sites for tRNA (A, P and E) and mRNA are on the interface side. mRNA enters and exists the ribosome in specific places. The entry pore through which mRNA can access the ribosome is between the shoulder and the head, whereas the exit pore is between the platform and the head [90], [91]. In the layout of the small subunit it is also possible to find the decoding centre, located on the interface and consisting of domains coming from the shoulder, head and penultimate stem [69].

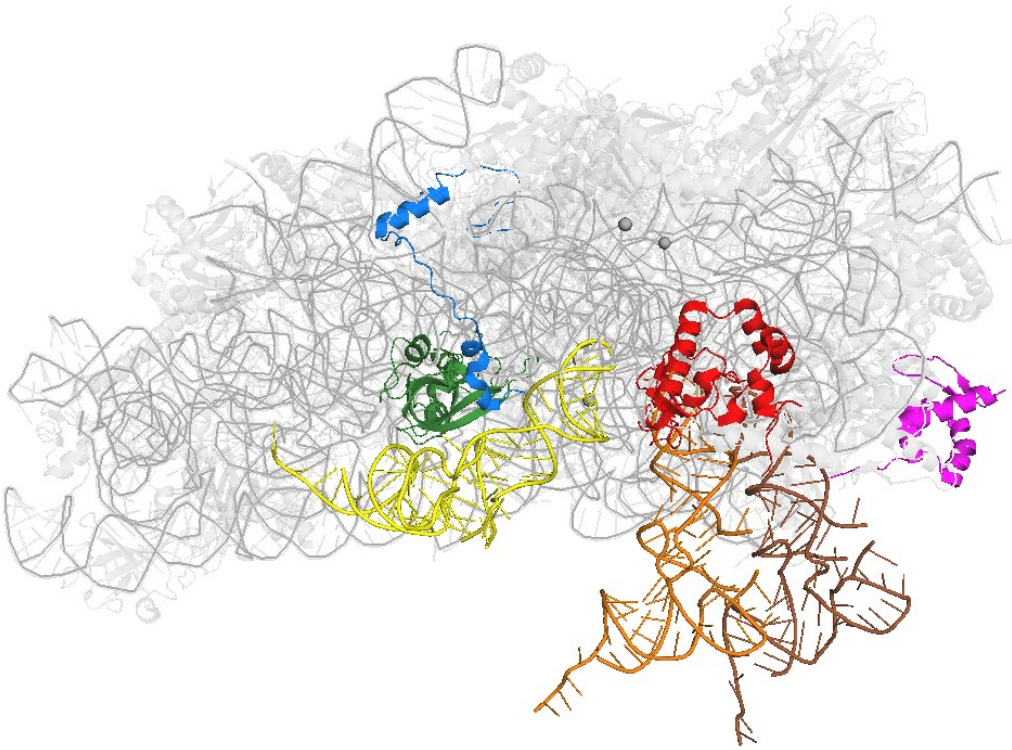

Figure 1 – The 40S ribosomal subunit. *The A site tRNA is depicted in yellow, the P site tRNA is colored in orange and E site tRNA is marked in brown. Furthermore, the location of four ribosomal proteins is shown. uS12 is illustrated in green, uS19 is interpreted in red, eS25 is marked in magenta and eS30 is colored in blue. Adopted and modified PDB entry 5LZS [54]*

Similarly, to the large subunit, the small subunit also contains rRNA expansion segments and proteins exclusive to eukaryotes. Their precise function remains unknown, but they are presumably linked to functional diversity in translation initiation. For example, the solvent side around the mRNA

exit site of the 40S subunit contains many unique proteins and rRNA expansion segments that have no analogy in the 30S subunit. There is some evidence from cross-linking and cryo-EM studies of the 40S ribosomal subunit that this area participates in the binding of eIF3, which in turn mediates mRNA recruitment to the ribosome during translation initiation [69], [92]–[94].

3.3 Decoding centre

As the term reveals, the decoding site is the place, where the ribosome deciphers the information given by the genetic code and conveys it into the language of amino acids. It is located in the A site and is formed by several RPs and 18S rRNA.

RPs have an essential role in decoding and the stabilization of tRNA [79], [80], [95]. This might also be the reason for the diverse positioning of the tRNA in the eukaryotic ribosome compared to the bacterial one [96]. As mentioned above, many of the components of the ribosome are conserved in both eukaryotic as well as archaeal ribosomes implicating that also the general mechanism of decoding is likely conserved too [97], [98].

uS12 is one of the proteins which lies in the proximity of the A site. Its conserved loop seems to be affecting the codon-anticodon duplex, specifically on the second and third positions [97]. Another interaction occurs with the anticodon stem-loop of tRNA in the A and P site with the main participation of uS19 and its C-terminus extension. In addition to these interactions, the A-site might be affected by the presence of the N-terminus of eS30 and eS31. Furthermore, although it has been long thought that eS25 only interacts with the P and E sites [72], [99], [100] it has been observed that eS25 also reaches as far as to the decoding centre (unpublished data).

3.3.1 Decoding rules

The global rules of genetic decoding are defined by the molecular components of the translation machinery [101]. We distinguish 64 nucleotide triplets, described as codons. There are 61 codons recognised by tRNAs carrying one of the 20 amino acids. This implies that the genetic code must be degenerated as several codons are recognised by tRNAs carrying only one type of amino acid. In fact, for some amino acids, the number of codon usage varies from two to six according to codon degeneracy [102]. There is one usual start codon, AUG, which is known to commonly start proteosynthesis by the amino acid methionine. However, the anticodon loop of $tRNA^{iMet}$ allows proteosynthesis to begin also from near-cognate codons, such as GUG or UUG^{36} in bacteria [103].

Finally, there are three stop codons, namely UGA, UAG and UAA, where proteosynthesis is terminated as there are no tRNAs with anticodons complementary to them. In some cases stop-codon readthrough can occur, which is most likely dependent on specific decoding rules, some of which have already been clarified [26]–[28]. Interestingly, pyrrolysine and selenocysteine are

referred as the 21st and 22nd amino acid, however, the molecular mechanisms of their incorporation is out of scope of this thesis.

High fidelity of translation is ensured by nine hydrogen bonds serving as a proofreading mechanism. These bonds are formed between the codon on mRNA, anticodon on tRNA and rRNA. *Ogle et. al* postulated that:" The binding of mRNA and cognate tRNA in the A site induces A1492 and A1493 to flip out of the internal loop of helix 44. This binding also causes the universally conserved base G530, which has been footprinted by A-site tRNA, to switch from the syn conformation present in the native structure to an anti conformation. In their new conformations, A1493 and A1492 interact, respectively, with the first and second base pairs of the codon-anticodon helix, whereas G530 interacts with both the second position of the anticodon and the third position of the codon. The result of these induced changes is that the first two base pairs of the codon-anticodon helix are closely monitored by the ribosome in a way that would be able to discriminate between Watson-Crick base pairing and mismatches, whereas the environment of the third, or "wobble," position appears to be suited for accommodating other base-pairing geometries"[104].

The decoding of a specific codon is highly dependent on the aa-tRNA synthetase that must charge the tRNA with an amino acid, and then the charged tRNA must recognise the correct codon. The preferred codons for the same type of amino acids may differ among organisms.

3.3.2 Wobbling

Translation fidelity can be compromised by the so called "Two-out-of-Three" hypothesis [105] which declares that only the nucleotides on the first two positions are important for codon-anticodon decoding, while the third position is not controlled as carefully.

The contact between the nucleotides of the codon-anticodon duplex is mediated by common Watson-Crick bonding. This means that the existence of wobbling is due to the fact that the nucleotide, situated in the wobble position, may not obey the general Watson-Crick rules, enabling the incorporation of nucleotides which are not in compliance. This solves the puzzle as to why the cell can use only about 40 types of tRNA molecules to translate all mRNA codons, while theoretically 61 would be needed for full codon recognition [106], [107]. It was Francis Crick, who had the original thought that wobbling may induce misdecoding of codons and therefore the wrong inclusion of amino acids into the chain [108]. The phenomenon of misdecoding further sheds light on the likelihood of wobbling on both positions 1 and 3 [102].

Nevertheless, as translation decoding is such a fundamental process for the life of every cell, any common inaccuracies would be absolutely fatal and therefore some other control mechanisms must be ensured. For example, in the prokaryotic ribosome the RNA nucleotides G530, A1492 and

A1493 take part in this process [95], [104]. However, the fact that wobbling increases the chances of misincorporation of amino acids brought by noncognate tRNA remains unquestionable.

4 Small ribosomal proteins forming the decoding site

The four ribosomal proteins discussed in this thesis are all located in the small ribosomal subunit, specifically in the decoding centre or in close proximity of it. uS12 and uS19 are small RPs which can commonly be found in all domains of life. On the other hand, eS25 and eS30 are only present in eukaryotes. The proteins in this thesis are named according to the nomenclature proposed by *Ban et. al* [109]

4.1 Ribosomal protein uS12

uS12 is a small ribosomal protein found in both prokaryotes and eukaryotes. In mammals it is also known as RPS23, while in yeast it is called S28 or Rps23A/B. In bacteria it is identified as S12. All mentioned alternatives of the protein uS12 are homologous to a certain extent. It has been shown that a mutation of the eukaryotic form of the protein uS12 may lead to impeded efficiency of translation termination. In case of the mutations of uS12 in bacteria, translation elongation rate was observed to be reduced [13]. These facts imply that the main function must have been at least partly conserved in evolution.

4.1.1 Analysis of uS12 gene and evolution of uS12 protein

The human gene uS12 is located on the chromosome 5 [[110]; Ensembl: ENSG00000186468]. Due to alternative splicing the uS12 protein can alternatively exist in several forms. The current evidence only refers to the role of the 143 amino acid long alternative of uS12, which is a part of the small ribosomal subunit. However, in the available literature it is not possible to find whether the other forms have any specific functions or localization within the ribosome or even cell. In yeast the uS12 protein is coded by two paralogous genes, RPS23A and RPS23B, that arose due to a whole genome duplication [[110]; Ensembl: YGR118W and YPR132W].

It can be observed that while human [[111] Uniprot entry: P62266] and mouse [[111] Uniprot entry: P62267] uS12 comprises of 143 amino acids, the yeast uS12 protein is made up of 145 residues [[111] Uniprot entry: P0CX29] and the bacterial uS12 of only 124 residues [[111] Uniprot entry: P0A7S3]. After alignment of the uS12 protein across several different organisms (Fig. 2) primary sequences of all vertebrate forms (e.g., human, mouse, zebrafish) of uS12 are almost completely identical. Moreover, the eukaryotic forms of uS12 are identical to almost 80 %. This indicates that the protein is very well conserved in evolution.

The prokaryotic uS12 is homologous to the human uS12 only in less than 35 %. However, two important domains (marked in yellow in Fig. 2) have been conserved across all life domains.

Interestingly, these conserved domains were shown to act in decoding and thus most likely also stay behind the main function of the protein (Fig. 2 and Table 1) [112], [113].

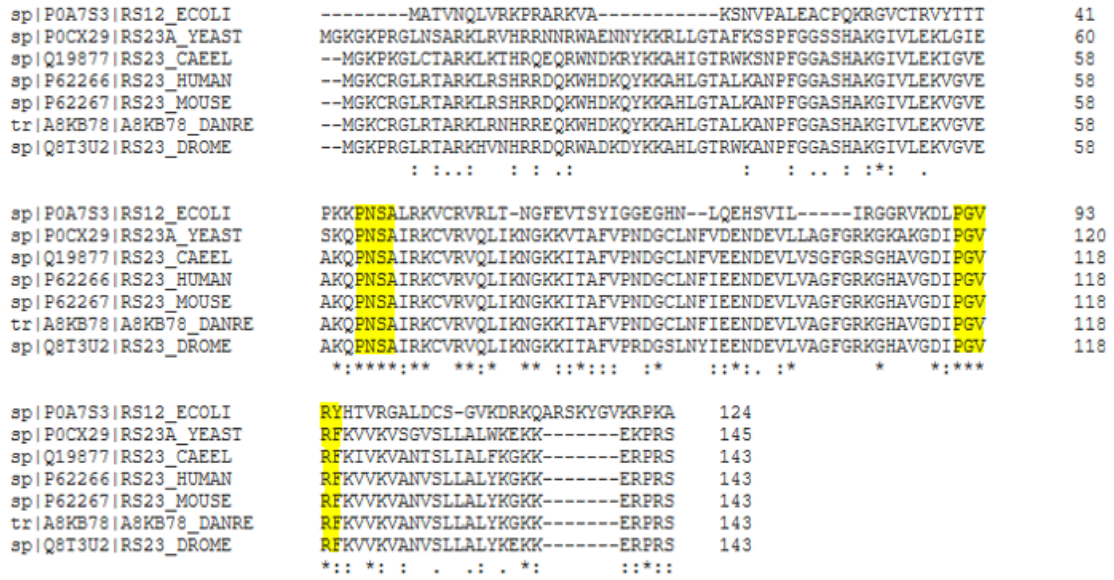

Figure 2 – Multiple alignment of the protein uS12 of various organisms. *Conserved loops containing the PNSA and PGVRY/F domains are marked in yellow. (Escherichia coli – ECOLI, Saccharomyces cerevisiae – YEAST, Ceanorhabditis elegans – CAEEL, Homo sapiens – HUMAN, Mus musculus – MOUSE, Danio rerio – DANRE, Drosophila melanogaster – DROME) [107. 108]*

Table 1 – Percentage of sequence identity of the uS12 protein of different organisms compared to *Homo sapiens*

uS12	Percentage of identity
Escherichia coli	34.78%
Saccharomyces cerevisiae	78.17%
Caenorhabditis elegans	84.62%
Homo sapiens	100.00%
Mus musculus	100.00%
Danio rerio	98.60%
Drosophila melanogaster	90.91%

4.1.2 Interplay between localization of the uS12 within ribosome and the uS12 role in translation

From the studies conducted on bacteria we know that uS12 is located in the vicinity of the mRNA entry channel and is an essential part of the decoding site. Therefore, it is very likely that it takes part in the maintaining of translation fidelity [114], especially through the contact with nucleotides G518 and A1492 (*Escherichia coli* numbering) of 16S rRNA and with the Mg^{2+} cation that interacts with the third nucleotide of the mRNA A site codon [115]. The third nucleotide of the A-site codon is less than 4 Å away from the proline 45 of uS12 [116], [117]. Pro45 adopts a stable conformation with potential implications for mRNA or tRNA translocation and also decoding [114].

Interestingly, Pro45 of prokaryotic uS12 is part of the conserved region PNSA (marked in yellow in Fig. 2) sharing the identical localization also within the eukaryotic ribosome (Fig. 3). Thus, it indicates potential evolutionary conserved role of uS12 in translation of eukaryotes as described in

bacteria. For illustration of uS12 localization within eukaryotic ribosome and PNSA region close to the decoding site we modified PDB entry: 5LZS [13] (Fig. 3).

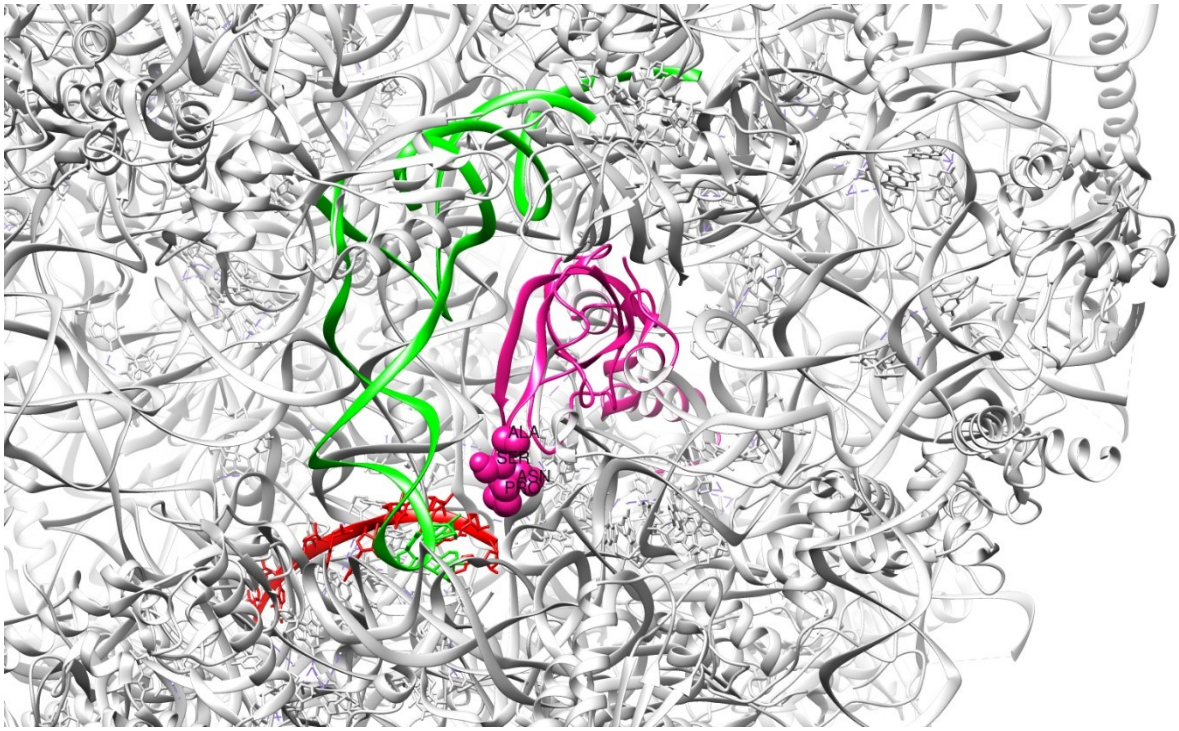

Figure 3 - Localization of the uS12 protein within rabbit ribosome. *Ribosomal protein uS12 is shown in deep pink, A-site tRNA in green and mRNA in red. Evolutionary conserved amino acids (Pro62, Asn63, Ser64, Ala65) of the uS12, shown as spheres, are localized in the vicinity of the decoding site. [modified PDB 5LZS [105] using Chimera 1.14]*

Other interactions of uS12 include those of the human homologue with mRNA via its ribose-phosphate backbone between nucleotides in positions +6/+7 (counting from the first nucleotide in the P site) [118].

Several studies have been conducted in yeast and bacteria on the effect of amino acid substitutions in homologous PNSA (position 45-48 in bacteria, 64-67 in yeast or 62-65 in humans) and PGVRY (position 91-95 in bacteria, or PGVRF on positions 118-122 in yeast and 116-120 in humans) sequences (Fig. 2). It has been shown that alterations in these locations have a negative impact on translation fidelity and hinder mRNA decoding [119]–[121]. Furthermore, the mentioned mutations decrease the effectivity of hydroxylations by human 2-oxoglutarate and Fe(II)-dependent oxygenase domain-containing protein 1 (OGFOD1) [122] which are otherwise also necessary for accurate translation fidelity and for which uS12 is a known substrate (specifically the Pro62 of uS12)[122]. Moreover, the stabilization of uS12 is impaired, which may lead to the existence of a mechanism that prevents ribosomes carrying uS12 mutants to be a part of a translationally active polysomal pool [89]. It is important to note that these mutations affect mainly rRNA production, ribosome assembly and decoding but not the rates of protein synthesis [123]. However, all these mutations lead to ribosomopathies that drive developmental diseases [89].

Interestingly, substitutions of Arg69 following the PNSA sequence had a pronounced effect on translational fidelity in yeast. The Arg69Lys substitution increased the rate at which ribosomes read through UAA and UAG stop codons [89].

It has been shown that the eukaryotic homologue of Pro45 (Pro64 in yeast and Pro62 in humans) undergoes hydroxylation which is necessary for accurate translation fidelity [122], and therefore can alternate the efficiency of translation termination. This event can change interactions between uS12 and the A-site and also the binding energy of the ribosome-mRNA complex [117]. Hydroxylation of Pro64 in yeast is executed by TPA1, which is a hydroxylase homologous to OGFOD1. It was shown that the deletion of TPA1 affects readthrough in a sequence- and context-dependent manner (Fig. 4) [124].

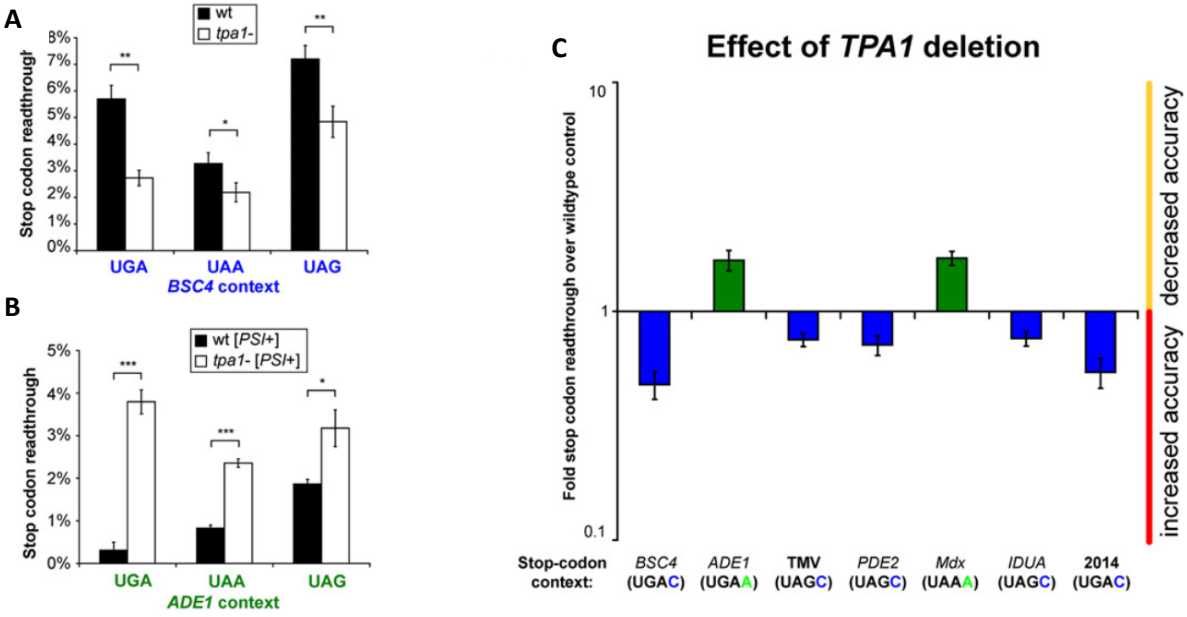

Figure 4 – Effect of TPA1 deletion on readthrough in context- and sequence-dependant manner. A) *Readthrough at stop codons UGA, UAA and UAG in BSC4 context is significantly altered upon the deletion of TPA1* ***B)*** *Readthrough at stop codons UGA, UAA and UAG in ADE1 context is significantly altered upon the deletion of TPA1* ***C)*** *Relative stop codon readthrough of tpa1⁻ compared with WT cells; Readthrough of different stop codon contexts are compared. Adopted and modified from [120]*

Despite the fact that homologous Pro62 in humans is located in the apex of the loop that extends to the tRNA acceptor site of ribosome [125], a crucial position of the decoding site, the effect of hydroxylations of this amino acid on readthrough is disputable [122], [124], [126], [127].

Additional experiments of *Valášek et. al* in yeast detected weaker but highly specific interactions between the purified j/hcr1-C-terminal domain (CTD) and small ribosomal proteins uS5 (RPS2) and uS12 (RPS23) dependent on the last 80 amino acid residues of j/HCR1 and the intact KERR motif (K205-x5-E211R212-x2-R215), which is conserved between eIF3j and the HCR1-like domain of eIF3a across species (23). uS5 was previously shown to occur on the solvent side and uS12 on the interface side of the mRNA entry channel (28). Further experiments in yeast showed that the last 16

amino acids of the human eIF3j (heIF3j) are required for stable binding of eIF3 to the 40S subunit (20), and that binding of heIF3j-CTD occurs in the 40S mRNA entry channel (27). Together these findings suggest that the ribosomal binding site of the CTD of eIF3j might have remained evolutionary conserved and that it thus represents an important functional domain of eIF3j .

Furthermore, uS12, along with uS3 and uS19, cross-links with the internal ribosome entry site (IRES), an RNA element typically present in the 5´untranslated region (UTR) in viruses and sporadically also in eukaryotic mRNAs [118].

4.1.3 Impact of mutations of uS12 protein on disease development

Mutations of the eukaryotic protein uS12 can lead to several diseases. Except of already mentioned ribosomopathies in the chapter 4.1.2 mutations of the human uS12, specifically at positions 67 or 120, can cause Brachycephaly, trichomegaly, and developmental delay disease (BTDD) [[111] Uniprot entry: P62266].

The fact that even such a minor change as the mutation of one amino acid can lead to extensive disorders shows the importance of a fully functioning ribosome.

4.2 Ribosomal protein uS19

Another small ribosomal protein, found in all domains of life, is uS19. This essential protein is also known under the name RPS15 in mammals and yeast. In yeast it is also known as RPS21, while in bacteria it is identified as S19., The main role of uS19 has most likely remained conserved as it is homologous among different species of organisms.

4.2.1 Analysis of RPS15 gene and evolution of uS19 protein

The human RPS15 gene is located on the chromosome 19. uS19 protein can alternatively exist in several forms resulting from translation of alternatively spliced mRNAs. The 145 amino acids long version is present in the ribosome. In the available literature it is not possible to find whether other forms of this protein also have any specific function or localization within the ribosome or even cell [[111] Uniprot entry: P62841]. The yeast RPS15 (uS19) protein comprises 142 amino acids [[110]; Ensembl: YOL040C] and the bacterial uS19 has only 93 amino acids in length [[111] Uniprot entry: P80381]. Structural alignment was carried out on several evolutionary variants of the uS19 protein (Fig. 5, Table 2), all of which were compared to the human uS19. It is apparent that the uS19 sequences of all vertebrates have a very high percentage of sequence identity (over 70 %) and the mammalian forms of the uS19 are completely identical. Interestingly, the uS19 of *Ceanorhabditis elegans* has a higher sequence identity compared to *Danio rerio* even though it is evolutionary older and thus one would expect the protein to have changed more. The other results are as expected with sequential identities decreasing in accordance with the evolutionary age of the organism (Fig. 5 and Table 2) [112], [113].

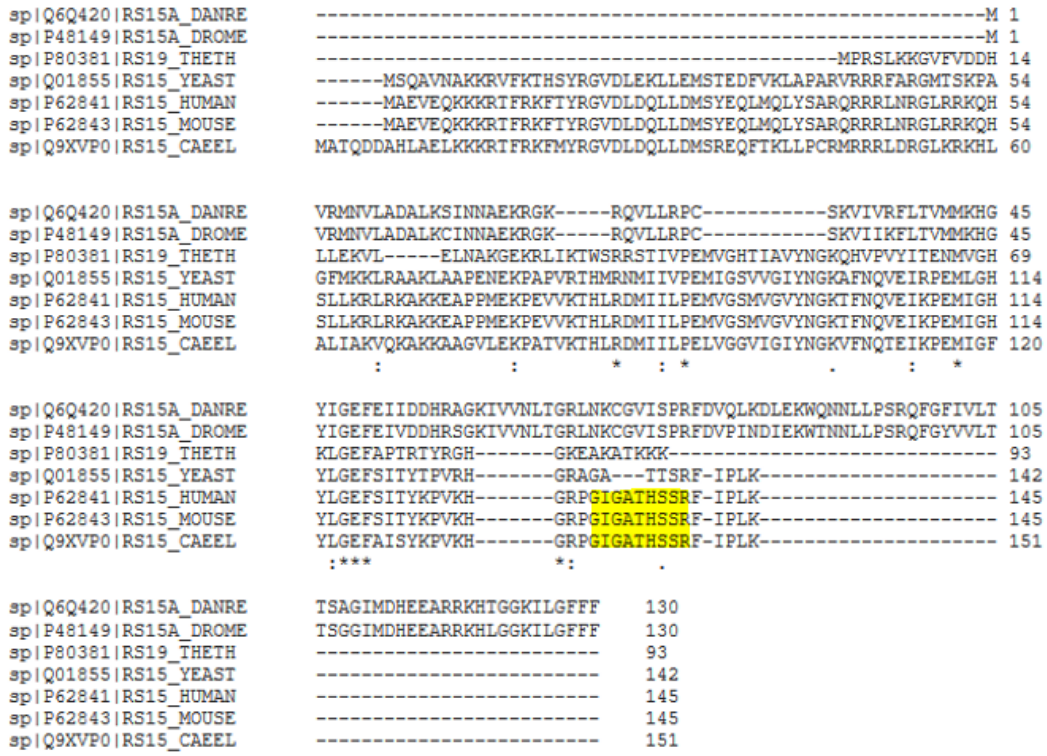

Figure 5 - Multiple alignment of the protein uS19 of various organisms. *The 9 amino acid sequence affecting readthrough, translation fidelity and decoding are highlighted. (Danio rerio – DANRE, Drosophila melanogaster – DROME, Thermus thermophilus – THETH, Saccharomyces cerevisiae – YEAST, Homo sapiens – HUMAN, Mus musculus – MOUSE, Ceanorhabditis elegans – CAEEL) [107, 108]*

Table 2 – Percentage of sequence identity of the uS19 proteins of different organisms compared to *Homo sapiens*

uS19	Percentage of identity
Danio rerio	71.43%
Drosophila melanogaster	62.50%
Thermus thermophilus	40.45%
Saccharomyces cerevisiae	59.31%
Homo sapiens	100.00%
Mus musculus	100.00%
Caenorhabditis elegans	73.94%

4.2.2 Location of uS19 within the ribosome and important interactions

uS19 is bound to the head of the ribosome towards the beak [128]. The exact location of the eukaryotic uS19 is depicted in the Fig. 6. The eukaryotic uS19 comprises of a C-terminal tail that extends up to the decoding centre [13], [129]. The first amino acid of the uS19 contacting the decoding centre is R130 (PDB 5LZS [105]). The C-terminal tail interacts with the riboses of the A site nucleotides located in the +4 and +7 position (counting from the first nucleotide in the P site). The tail contacts the decoding site in this way throughout all phases of translation [118]. The C-terminal tail of human uS19 in contrast to the tail of prokaryotic uS19 is able to reach the ribosomal decoding site and interact with the mRNA due to the larger length of the tail and its eukaryote-specific location in the 40S subunit.

[91], [104], [129]. This fact brings chaos upon the generally accepted hypothesis that the decoding centre is a structural assembly that is conserved among all life domains [118].

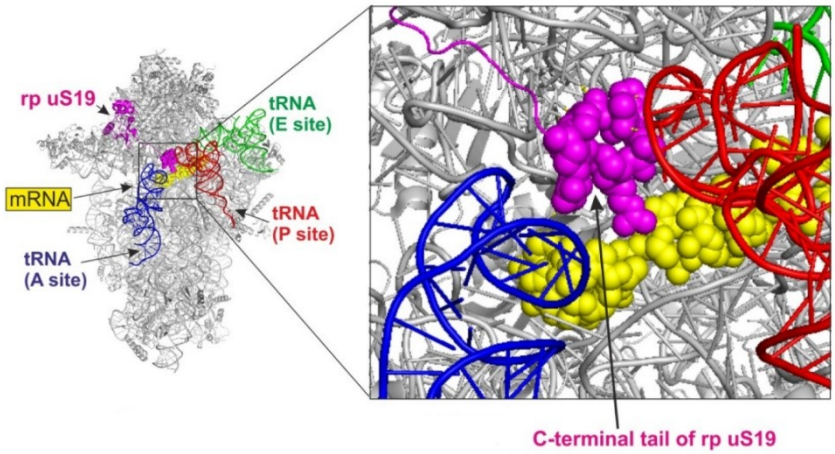

Figure 6 - Location of uS19, mRNA, tRNA and the decoding centre *(A, P, E sites) in the 40S ribosomal subunit; based on the near atomic complex by Shao et al. 2016: On the left the general view of the interface side of the 40S subunit; uS19 illustrated by magenta color, tRNA molecules are depicted in dark blue (A site), red (P site) and green (E site), mRNA is shown in yellow. The zoomed image on the right gives a detailed view of the interactions of the C-terminal tail of uS19 (positions 131-140 are shown as spheres) with the decoding centre, mRNA is also shown in spheres. Adopted and modified from [132]*

Further experiments reveal that functional centres of the ribosomal subunits are connected via intersubunit bridges B1a and B1b/c. In prokaryotes the two bridges are formed by a single protein, uS13, which connects to the Helix 38 in the A site to form the B1a bridge and with uL5 to create the B1b/c bridge. However, in eukaryotes this function has been separated between two proteins, which are connected to each other [72], [130] and most likely cooperate [130]. Whereas, the uS13 binds to the uL5 to build the B1b/c bridge as in prokaryotes, the B1a bridge is constructed using solely uS19, which again partners with the Helix 38. This discovery indicates that the uS19 acts as an essential conveyer of information between the two ribosomal subunits directly via the B1a interbridge and indirectly through the B1b/c bridge with the assistance of the uS13 [130], [131]. Furthermore, the Helix 38 is a large structure interacting also with the D-loop of aa-tRNA, the GTPase associated centre and the peptidyl transferase centre [131]. Therefore, the fact that uS19 interacts with such an important partner might be of great importance.

Structural studies of the uS19 by Bowen´s group have also shed light on the effect of mutations of amino acids in close proximity of the aforementioned proteins (uS13, uL5, uS19). These alterations inhibit the binding of the eEF2 translocase. They also influence translation fidelity by enhancing the affinity of the ribosomes for the tRNAs in the P site and the ternary complexes in the A site by shifting the ribosomes towards the unrotated state [130]. However, uS19 itself does not interact with the P site directly [132].

4.2.3 The role of uS19 in translation

According to the research teams of *Bulygin. et al* and *Sharifulin et. al* there seem to be several implications of the contact between the C-terminal tail of the uS19 and the decoding centre. Firstly, the tail is essential for the formation of polysomes, a critical event in the elongation phase.

Secondly, it is crucial for the accommodation of aa-tRNAs in the A site of the decoding centre, thus improving the efficiency of translation by promoting the fixation of the mRNA codon in the decoding centre. This would also diminish the likelihood of a frameshifting event. Furthermore, the uS19 tail could assist in the identification of stop codons by eRF1 during termination [118], [132].

Important function of the C-terminal tail in translation has been described also by *Bretones et al.* They have clarified that the uS19 mutations between positions 131-140 influence global protein synthesis, translation initiation and translation fidelity. Specifically, alterations of amino acids in positions 131, 136 and 139 corresponding to proline, threonine and serine, respectively, have been shown to decrease rates in protein synthesis. What is more, all mutants with alterations in the positions 132, 134, 136, 137 and 138 have shown a significant increase in readthrough. Specifically mutations in the positions 132, 136 and 138 lead to higher rates of readthrough at the UAA stop codon and alterations in the positions 134 and 137 induced higher readthrough at the UAG stop codon. Thus, the heptapeptide GIGATHS (132-138; Fig. 5 and Fig. 7) plays an important role in translation fidelity, hence affecting the ribosome function [133]. Consistently with the data of *Bretones et al.* direct participation of the tetrapeptide 137-HSSR-140 in decoding has already been described also by other research groups [104], [129].It is important to mention that these mutations do not have any impact on the efficiency of ribosome assembly [133].

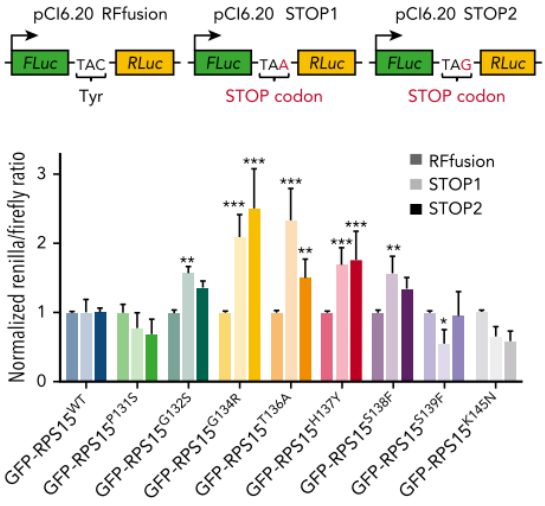

Figure 7 *- **Readthrough rates after the UAA and UAG stop codon in the aftermath of mutations of uS19** in positions 131, 132, 134, 136, 137, 138, 139 and 145. Adopted and modified from [133]*

4.2.4 The role of uS19 in ribosome biogenesis and in nuclear export of pre-40S particles

Over 200 structural components and factors are required for the assembly of the eukaryotic ribosome. The uS19 seems to have an important function in the biogenesis of the 40S subunit, specifically in the final maturation stage. The C-terminal tail of the uS19 lies within interaction distance of the SSU assembly factors, such as Rio2 and Tsr1, factors necessary for 20S pre-rRNA processing in the cytoplasm. [134]. Therefore, there is a possibility the uS19 recruits Rio2 and Tsr1 [134], binds them with another pair, Ltv1 and Enp1 [128], and through the binding of its C-terminal tail is responsible for the maturation of the 40S subunit [132].

Furthermore, the uS19 has also been discovered to act as a factor important for the export of the pre-40S particle from the nucleus to the cytoplasm [134]. It is also possible that it interacts with the nuclear export machinery in a way that is not clearly understood, yet [134]. However, it has

become clear that this function of the uS19 is dependent on the exportin Crm1, another factor essential for nuclear export of pre-ribosomes in vertebrates [135]. As uS19 is highly homologous between all eukaryotes, Roquette's team has concluded, that this function has also been conserved [136]. According to Leger-Silvestre other functions of uS19 may also include chaperoning pre-rRNA folding.

4.2.5 Impact of mutations of uS19 protein on disease development

Mutations in the uS19 gene, leading to alterations in the C-terminal end of the protein, have been associated with B-cell chronic Lymphocytic Leukemia (BCLL) [137]. *Bretones et al.* further investigated these mutations of the uS19 C-tail as already described in the chapter 4.2.2. Their study suggests that decreased translation fidelity causing altered protein expression patterns might be new molecular mechanism involved in the BCLL development [133].

Moreover, uS19 along with other ribosomal proteins has been shown to regulate the Mdm2-MdmX pathway. The Mdm2-MdmX pathway tightly regulates the levels of p53, a transcription factor associated with over a half of human cancers, including BCLL. The mechanism of Mdm2 and p53 stabilization appears to be through inhibiting the E3 ubiquitin ligase activity of Mdm2 [138] due to the interaction of ectopically expressed uS19 with the central Zinc finger region or with the N-terminal region of Mdm2. This leads to the inhibition of Mdm2 and p53 degradation and thus to the p53 activation. It is also capable of regulating the MdmX activity in a way that is not exactly understood [138].

Mutations in uS19 have also been linked with Diamond-Blackfan anemia a hereditary disease identified as a predisposition to cancer [139].

4.3 Ribosomal protein eS25

eS25 is a ribosomal protein found in the 40S subunit and it is unique for eukaryotes. It is formerly known as protein S25 or RPS25.

4.3.1 Analysis of RPS25 gene and evolution of eS25 protein

In humans the eS25 protein is coded by the RPS25 gene located on the chromosome 11. RPS25 pre-mRNA might be post transcriptionally modified into six transcripts, and even though three are protein coding, two of them are sentenced to the NMD pathway. The longest protein consists of 125 amino acids and ensemble within ribosome [[110]; Ensembl: ENSG00000118181]. On the contrary, in yeast eS25 is coded by two paralogous genes RPS25A and RPS25B. These two paralogs were created as a result of a whole genome duplication. Transcripts of both paralogs are translated into a functional 108 amino acids long protein. These two variants differ in one amino acid in position 105. RPS25A contains a threonine, while RPS25B encompasses an alanine [[110]; Ensembl: YGR027C and YLR333C].

The following sequence alignments (Fig. 8 and Table 3) show that the protein eS25 is very well conserved among all vertebrates as there is over 90 % sequential identity between them. The eS25 of *Drosophila melanogaster* is identical to the human eS25 sequence in over 70 %, *Ceanorhabditis elegans* in over 65 % and *Saccharomyces cerevisiae* has almost 50 % of sequential identity. The fact that almost 50 % of the protein sequence has been conserved throughout evolution indicates that function of the eS25 may be of notable importance [112], [113].

```
sp|Q3E792|RS25A_YEAST      MPPKQQLSK--------AAKAAAALAGGKKSKKKWSKKSMKDRAQHAVILDQEKYDRILK 52
sp|P62851|RS25_HUMAN       MPPKDDKKKKDAGKSAKKDKDPVNKSGGKAKKKKWSKGKVRDKLNNLVLFDKATYDKLCK 60
sp|P62852|RS25_MOUSE       MPPKDDKKKKDAGKSAKKDKDPVNKSGGKAKKKKWSKGKVRDKLNNLVLFDKATYDKLCK 60
sp|Q6PBI5|RS25_DANRE       MPPKDSKQKKDAGKS-KKDKDPVNKSGGKAKKKKWSKGKVRDKLNNLVLFDKATYDKLYK 59
sp|P48588|RS25_DROME       MPPKKDAKSSAK-QPQKTQKKKEGSGGGKAKKKKWSKGKVRDKLNNQVLFDKATYEKLYK 59
sp|P52821|RS25_CAEEL       MPPKKDPKGGKA-PP---SKKKEGSGGGKAKKKKWSKGKVRDKLNNMVLFDQATYDKLYK 56
                           ****.. .          *     .*** .****** .::*: :: *::*: .*::: *

sp|Q3E792|RS25A_YEAST      EVPTYRYVSVSVLVDRLKIGGSLARIALRHLEKEGIIKPISKHSKQAIYTRATASE---- 108
sp|P62851|RS25_HUMAN       EVPNYKLITPAVVSERLKIRGSLARAALQELLSKGLIKLVSKHRAQVIYTRNTKGGDAPA 120
sp|P62852|RS25_MOUSE       EVPNYKLITPAVVSERLKIRGSLARAALQELLSKGLIKLVSKHRAQVIYTRNTKGGDAPA 120
sp|Q6PBI5|RS25_DANRE       EVPNYKLITPAVVSERLKIRGSLARAALQELLGKGLIKLVSKHRAQVIYTRNTKGTDEAA 119
sp|P48588|RS25_DROME       EVPAYKLITPSVVSERLKIRGSLAKRALIELREKGLIKQVVQHHSQVIYTRATKGDEA-- 117
sp|P52821|RS25_CAEEL       EVITYKLITPSVVSERLKVRASLAKAGLKELQAKGLVKCVVHHHGQVVYTRATKEADVIV 116
                           **  *: :: :*: :***: .***: .* .*  :*::* : :*  *.:*** *

sp|Q3E792|RS25A_YEAST      -----        108
sp|P62851|RS25_HUMAN       AGEDA        125
sp|P62852|RS25_MOUSE       AGEDA        125
sp|Q6PBI5|RS25_DANRE       PEKEA        124
sp|P48588|RS25_DROME       -----        117
sp|P52821|RS25_CAEEL       E----        117
```

Figure 8 - Multiple alignment of the protein eS25 of various organisms. *(Saccharomyces cerevisiae – YEAST, Homo sapiens – HUMAN, Mus musculus – MOUSE, Danio rerio – DANRE, Drosophila melanogaster – DROME, Ceanorhabditis elegans – CAEEL) [107, 108]*

Table 3 – Percentage of sequence identity of the eS25 proteins of different organisms compared to *Homo sapiens*

eS25	Percentage of identity
Saccharomyces cerevisiae	49.38%
Homo sapiens	100.00%
Mus musculus	100.00%
Danio rerio	90.40%
Drosophila melanogaster	70.34%
Caenorhabditis elegans	65.25%

4.3.2 Location of eS25 within the ribosome and important interactions

Compared to the prokaryotic ribosome, the eukaryotic homolog has several additional components. Therefore, naturally, some of the interactions in the ribosomes may differ including also the interactions in the decoding

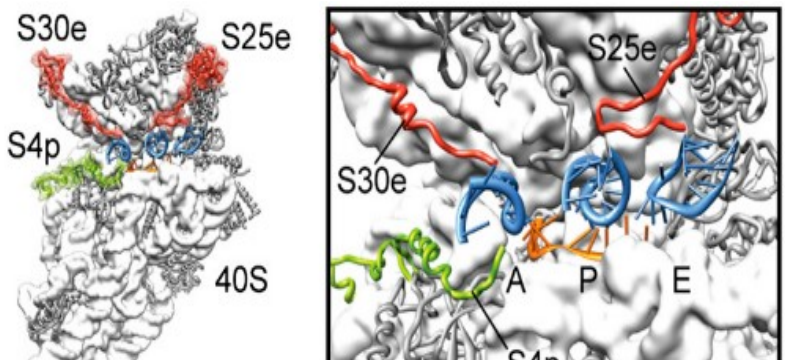

Figure 9 – Location of specifically eukaryote proteins eS30 and eS25 of the 40S subunit and their interaction with the A, P and E sites. *eS30 contacts the anticodon stem of the A site tRNA in position 28 and 29. eS25 is located between the P and E site and contacts the tRNA in E site. Adopted and modifief from [99]*

centre and tRNA-binding sites. This is due to the presence of 4 exclusively eukaryotic proteins (eS30, eS31, eS25 and eS1), which stretch their extensions towards the A, P and E site [140]. The protein eS25 is placed in the head of the 40S ribosomal subunit on the solvent side, specifically near the mRNA exit site [69], [99]. The N-terminal extension of the eS25 was originally thought to reach between the P and E sites (Fig. 9) [72], [99], [100]. However, recently it has been observed that N-terminus of eS25 stretches as far as to the decoding centre of the A site (Thomas Becker and Roland Beckmann, personal communication). eS25 is also in the vicinity of uS7 (S5), which interacts with E site tRNA [80].

Furthermore, eS25 along with another exclusively eukaryotic protein, eS19, has been identified to interact with eEF3 [99].

4.3.3 Role of eS25 in translation

Functions of eS25 have been summarised in several studies. One important role is the effect of eS25 on readthrough. Previous studies claim that the depletion in eS25 does not significantly affect programmed frameshifting or miscoding and improves the efficiency of translation termination [141]. Further studies in this area would be interesting to exactly define which specific residues of eS25 take part in these processes.

In addition, it has been shown that eS25 is not an essential protein for the biogenesis of the 40S subunit, though the general ribosome profile is altered in eS25 absence. The results indicate that the association of the 40S with the 60S subunit in order to form the 80S ribosome might be impaired and/or that the 80S might be degraded due to its instability [142].

Another study suggests that as eS25 is not an essential protein, its function could be the formation of specialized ribosomes within the cell that could translate a specific population of mRNAs [141]. Altogether, the question of ribosome heterogeneity has been a very discussed topic in recent years [143].

Further unexpected discovery has uncovered that the knockdown of Rps25 provides the cell with resistance to ricin, a highly potent toxin. The authors speculate that under certain stress conditions eS25 may participate in translation regulation of transcripts that affect ricin sensitivity [144].

As eS25 is adjacent to the E site, *Landry et al*. have suggested that eS25 may participate in the E site tRNA binding. They have proposed two models. The first hypothesis suggests that if eS25 takes part in the binding of E site tRNA, then the translocation rate would be slower in ribosomes devoid of eS25. The second model proposes that the interaction between the codon and anticodon in the E site is crucial for the maintaining the reading frame through the contact of eS25 to E site tRNA [141].

Majority of studies concerning eS25 has drawn an attention to its role in translation initiation [141], [145], [146]. While the loss of eS25 does not affect cap-dependant translation initiation, eS25 exhibit significant participation in two alternative initiation routes, particularly in internal ribosome entry site (IRES) initiation and ribosome shunting [156]. It has been suggested that both of the alternative routes employ a different mechanism to recruit the 40S subunit, but require eS25 for further steps in initiation, such as the joining of subunits to form the 80S complex [142], [147]. Furthermore, eS25 seems to be crucial for the induction of the conformational change allowing the contact of IRES with the ribosome [141], [147]. All in all, it seems that eS25 is an essential protein only for some viruses and proliferation of some cells that under specific conditions might depend one of the two alternative routes of translation initiation described above. That would mean that eS25 would be a worthy target for a broad spectrum of therapeutic agents.

4.3.4 Impact of mutations of eS25 protein on disease development

Despite the fact there is no direct evidence that the mutation of the eS25 protein causes disorders, several links have been proposed. One meaningful role of eS25 is in regulation of the Mdm2-MdmX pathway, which has already been reviewed in 4.2.5. Similarly, p53 is stabilized through ectopic expression of the eS25 protein. In nucleus, eS25 inhibits the E3 ubiquitin ligase activity through binding to the central region of Mdm2. This leads to a decline in Mdm2-mediated p53 ubiquitination and thus to p53 activation. Interestingly, feedback regulation of eS25 by p53 has also been observed. p53 specifically binds to the eS25 promoter hence supressing its expression. This feedback loop may have a critical role in cancer development, however, that awaits to be examined [148].

Secondly, eS25 has been found to interact with the Epstein-Barr virus (EBV), specifically with the EBV Nuclear Antigen 1 (EBNA1). This is not a unique interaction as many RPs have been linked to EBV before. EBV has also been associated with the development of several types of cancer, however the link between RPs and tumorigenesis through EBV has not been clearly established. Nevertheless, the EBNA-eS25 interaction is of interest and it has been proposed that it may be an alternative route for EBNA1-mediated destabilization of p53 and thus EBV-mediated carcinogenesis [149].

Finally, it has been observed that eS25 has a strong targeting signal on the N-terminus that navigates it to the nucleus with a strong preference for the nucleolus. There are two regions of highly basic residues which also encompass alanine in position 17 that has been determined as the main potential targeting signal. However, most importantly, these nucleolar targeting signals are a common feature of Rev proteins, RNA binding transactivators, which are also possessed by the human immunodeficiency virus (HIV). Thus, clear understanding of eS25 may be of great importance for further understanding of HIV and AIDS development [150].

4.4 Ribosomal protein eS30

eS30 is a protein solely found in eukaryotes in the small ribosomal subunit. It has previously been known as S30.

4.4.1 Analysis of FAU gene and evolution of eS30 protein

In humans eS30 is coded by the FAU gene located on the chromosome 11 [[110]; Ensembl: ENSG00000149806]. Transcript of the FAU gene is translated into a fusion protein consisting of two proteins, the N-terminal fubi and the C-terminal eS30. The fubi is an ubiquitin-like protein with unknown exact function. The mature eS30 is created in the aftermath of a posttranslational cleavage of the 133 amino acids long protein into fubi (74 amino acids) and eS30 itself (59 amino acids) [151]. On the other hand, in yeast eS30 is coded by two genes, RPS30A and RPS30B, created as the outcome of a whole genome duplication. Both genes code for functional 63 amino acids long protein [[110]; Ensembl: YLR287C-A and YOR182C].

After sequence alignment (Fig. 10 and Table 4) it is obvious that all vertebrate forms of eS30 have remained almost completely conserved as their sequences are identical in over 96 %. The yeast eS30 variant is identical in over 62 % and *Ceanorhabditis elegans* in 76 % compared to the human sequence of eS30 [112], [113].

```
sp|P0CX33|RS30A_YEAST       ------------------------------------------------------------        0
tr|Q18231|Q18231_CAEEL      MQIFLLGLDNTTHTLDVDASTTLSAIKGVIGAGEE-----FSISYGSKVLSEELTLGECQ        55
sp|P62861|RS30_HUMAN        ------------------------------------------------------------        0
sp|P62862|RS30_MOUSE        ------------------------------------------------------------        0
tr|Q6PC01|Q6PC01_DANRE      MQLFLRA--QSLHTLEVSGNETVRDIKAHVEALEGVCVEDQLLLLSGAPLQDDSTLLNCG        58

sp|P0CX33|RS30A_YEAST       --------------MAKVHGSLARAGKVKSQTPKVEKTEKPKKPKGRAYKRLLYTRRFVN        46
tr|Q18231|Q18231_CAEEL      IESLSTLSVNGRLLGGKVHGSLARAGKVRAQTPKVDKQDKKKKKRGRAFRRVQYTRRYVN        115
sp|P62861|RS30_HUMAN        ----------------KVHGSLARAGKVRGQTPKVAKQEKKKKKTGRAKRRMQYNRRFVN        44
sp|P62862|RS30_MOUSE        ----------------KVHGSLARAGKVRGQTPKVAKQEKKKKKTGRAKRRMQYNRRFVN        44
tr|Q6PC01|Q6PC01_DANRE      ITEFCTLEVSGRLLGGKVHGSLARAGKVRGQTPKVDKQEKKKKKTGRAKRRIQYNRRFVN        118
                                        ************:.***** * :* **  *** :*: *.**:**

sp|P0CX33|RS30A_YEAST       VTLVNGKRRMNPGPSVQ   63
tr|Q18231|Q18231_CAEEL      VASGPGKKRGPNSNS--   130
sp|P62861|RS30_HUMAN        VVPTFGKKKGPNANS--   59
sp|P62862|RS30_MOUSE        VVPTFGKKKGPNANS--   59
tr|Q6PC01|Q6PC01_DANRE      VVPTFGKKKGPNANS--   133
                            *.   **::   . *
```

Figure 10 - Multiple alignment of the protein eS30 of various organisms. *(Saccharomyces cerevisiae – YEAST, Ceanorhabditis elegans – CAEEL, Homo sapiens – HUMAN, Mus musculus – MOUSE, Danio rerio – DANRE). [107, 108]*

Table 4 – Percentage of sequence identity of the eS30 proteins of various organisms compared to *Homo sapiens*

eS30	Percentage of identity
Saccharomyces cerevisiae	61.67%
Caenorhabditis elegans	76.27%
Homo sapiens	100.00%
Mus musculus	100.00%
Danio rerio	96.61%

4.4.1 Interplay between localization of eS30 within ribosome and eS30 role in translation

eS30 is located at the entry site of the mRNA channel of the eukaryotic SSU. It is made up of α helices that are not tightly attached to each other. Interestingly, the area where the α-helical portion of eS30 is localized is in bacteria occupied by the N-terminus of uS4 (corresponding to rpS9e) and the h16 RNA helix, which has a different position in 40S [100]. eS30 protrudes between eS23 and the h18 RNA helix with its N-terminus and between h16 and h18 RNA helices with the helical C-terminus [71].

In view of the fact that eS30 has no clear bacterial homolog it is likely that it has evolved in order to make the decoding event even more precise. As stated in chapter 4.3.2, the existence of four ribosomal proteins, unique for eukaryotes, slightly changes the interactions of proteins within the decoding centre in comparison to prokaryotic ribosome.

The N-terminal extensions of eS30 and eS31 that reach into the A site are additional interactions on the SSU maintained in eukaryotic 80S ribosomes [72], [99], [100], [140]. eS30 stretches towards a groove between two flipped-out decoding bases and towards the phosphate backbone of the anticodon +1 position to establish possibly stabilizing contacts. These contacts are created once a cognate aa-tRNA is present. An important residue taking part in these interactions is a conserved His76, that is located on the N terminus of eS30 [13]. Furthermore, the N-terminus of eS30 is involved in the interaction with the A site tRNA anticodon stem loop between positions 28-29 [96], where it probably assists uS19 in controlling the A site codon in remaining in a conformation required for correct recognition [13], [152]. Along with uS19, eS30 could take part in stabilizing and fixing the aa-tRNAs to mRNA codons, hence enhancing the effectivity of translation. There is no current knowledge whether the interactions of uS19 and eS30 are also essential for elongation, which could be likely. Further investigations will be necessary in order to fully understand the role of RPs in the mechanism of translation in humans [152].

eS30 could provide a structural link between the outer and inner ends of the mRNA entry tunnel, suggesting that it may play a functional role in escorting the mRNA through the 40S subunit [4]. This would be mediated by a contact on the solvent side of the ribosome, where eS30 could interact with the 3´end of mRNA and it would be accomplished through the eS30 C-terminus. Moreover, eS30 extends conserved basic lysine residues into the mRNA entry channel, indicating that along with uS3 it acts as a factor unwinding the secondary mRNA structure [140].

Along with eS31 and several 18S rRNA sites, eS30 interacts with the N domain of eRF1 through its N tail [40]. As eRF1 controls the stop codon decoding, it is possible that eS30 could also participate in termination or readthrough. The efficiency of translation could further be improved by the contact with eIF1A [153] and eEF2 [154].

There is no current knowledge whether mutations in the eS30 protein induce any type of diseases or disorders or contribute to their development in any significant way.

5 Summary of interactions and functions of four unique small ribosomal proteins in the decoding site & Perspectives

The process of translation has been studied for several decades as it is a fundamental mechanism for the existence of every living cell. The main “producer” in translation is the ribosome ensuring translation accuracy through monitoring proper and accurate decoding. Altogether four ribosomal proteins of the small subunit, namely uS12, uS19, eS25 and eS30, create the decoding center together with 18S rRNA and it seems that their roles in translation are rather functionally interconnected.

Sequence alignments of these four ribosomal proteins revealed that either they have been conserved to a large extent as whole proteins or at least the sequences of their functional domains have been preserved. These facts imply that their roles within the ribosome are of great importance or even indispensable. That is beyond all doubt a strong reason for further research of these proteins. Besides, clear understanding of their functions could also be helpful for the development of methods to cure many types of ribosomapathies or other diseases.

Interestingly, all four proteins have been associated with stop codon readthrough. This is important, since an artificially imposed event of readthrough of PTCs naturally occurring in critical genes may guarantee the synthesis of a full-length protein rather than of its defective truncation, which often gives rise to various diseases. This has undoubtedly been the main driving force behind numerous research activities aiming at leveraging readthrough potential, in which these four proteins may play a significant role, to develop specific non-invasive treatments of disorders caused by premature termination and nonsense mutations. Therefore, it is clear that further knowledge of readthrough and all related molecular events with these four proteins in mind is acutely required.

Here I briefly summarize all so far identified roles of the aforementioned ribosomal proteins in translation termination and readthrough. Firstly, in case of uS19, the C-terminal tail is critical for the correct identification of stop codons by eRF1 [118], [132] and interestingly, mutations of amino acid residues between positions 131-140 in this tail have been described to have a pronounced effect on readthrough [133]. Secondly, it is known that eS30 also interacts with eRF1but through its N-terminal domain. Therefore, it would be of interest to examine its role during termination and especially readthrough. Thirdly, hydroxylations of Pro64 of uS12 in yeast have been proven to affect readthrough in context and sequence dependent manner. Nevertheless, the consequences of Pro62 hydroxylations in humans remain to be clarified and thus offer a perspective field for future research [124]. Finally, it has been shown that the absence of eS25 improves the efficiency of translation termination but does not significantly affect programmed frameshifting or miscoding[141]. Despite this progress it is obvious that there are still many open questions to be answered. For example, a

perspective approach to understanding readthrough would be to uncover the exact role of nucleotides forming the anticodon stem of the tRNA molecule, which is a direction of basic research where I was involved experimentally during my bachelor work.

To conclude, this thesis has attempted to extend the knowledge of translation regulation by particapation of the aforementioned small ribosomal proteins during translation, and most importantly, during stop codon readthrough.

6 Literature

All reviews are marked with *

[1] * Iwona J. Fijalkowska, R. M. Schaaper, and P. Jonczyk, "DNA replication fidelity in Escherichia coli: a multi-DNA polymerase affair," *FEMS Microbiol Rev.*, vol. 36, no. 6, pp. 1105–1121, 2012.

[2] * Alan G. Hinnebusch, "Structural Insights into the Mechanism of Scanning and Start Codon Recognition in Eukaryotic Translation Initiation," *Trends Biochem. Sci.*, vol. 42, no. 8, pp. 589–611, 2017.

[3] * Leoš Shivaya Valášek *et al.*, "Survey and summary: Embraced by eIF3: Structural and functional insights into the roles of eIF3 across the translation cycle," *Nucleic Acids Res.*, vol. 45, no. 19, pp. 10948–10968, 2017.

[4] T. V. Budkevich *et al.*, "Regulation of the mammalian elongation cycle by subunit rolling: A eukaryotic-specific ribosome rearrangement," *Cell*, vol. 158, no. 1, pp. 121–131, 2014.

[5] * Thomas E. Dever, J. D. Dinman, and R. Green, "Translation elongation and recoding in eukaryotes," *Cold Spring Harb. Perspect. Biol.*, vol. 10, no. 8, pp. 1–25, 2018.

[6] * Thomas E. Dever and R. Green, "The elongation, termination, and recycling phases of translation in eukaryotes," *Cold Spring Harb. Perspect. Biol.*, vol. 4, no. 7, pp. 1–16, 2012.

[7] M. G. Gagnon, J. Lin, and T. A. Steitz, "Elongation factor 4 remodels the A-site tRNA on the ribosome," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 113, no. 18, pp. 4994–4999, 2016.

[8] S. Varenne, J. Buc, R. Lloubes, and C. Lazdunski, "Translation is a non-uniform process. Effect of tRNA availability on the rate of elongation of nascent polypeptide chains," *J. Mol. Biol.*, vol. 180, no. 3, pp. 549–576, 1984.

[9] J. F. Curran and M. Yarus, "Rates of aminoacyl-tRNA selection at 29 sense codons in vivo," *J. Mol. Biol.*, vol. 209, no. 1, pp. 65–77, 1989.

[10] G. Bertram, H. A. Bell, D. W. Ritchie, G. Fullerton, and I. Stansfield, "Terminating eukaryote translation: Domain 1 of release factor eRF1 functions in stop codon recognition," *RNA*, vol. 6, no. 9, pp. 1236–1247, 2000.

[11] L. Y. Frolova *et al.*, "Mutations in the highly conserved GGQ motif of class I polypeptide release factors abolish ability of human eRF1 to trigger peptidyl-tRNA hydrolysis," *RNA*, vol. 5, no. 8, pp. 1014–1020, 1999.

[12] H. Song *et al.*, "The crystal structure of human eukaryotic release factor eRF1 - Mechanism of stop codon recognition and peptidyl-tRNA hydrolysis," *Cell*, vol. 100, no. 3, pp. 311–321, 2000.

[13] S. Shao, J. Murray, A. Brown, J. Taunton, V. Ramakrishnan, and R. S. Hegde, "Decoding Mammalian Ribosome-mRNA States by Translational GTPase Complexes," *Cell*, vol. 167, no. 5, pp. 1229-1240.e15, 2016.

[14] V. P. Pisareva, A. V. Pisarev, C. U. T. Hellen, M. V. Rodnina, and T. V. Pestova, "Kinetic analysis of interaction of eukaryotic release factor 3 with guanine nucleotides," *J. Biol. Chem.*, vol. 281, no. 52, pp. 40224–40235, 2006.

[15] L. Frolova, X. L. E. Goff, G. Zhouravleva, E. Davydova, M. Philippe, and L. Kisselev, "Eukaryotic polypeptide chain release factor eRF3 is an eRF1- and ribosome-dependent guanosine triphosphatase," *RNA*, vol. 2, no. 4. pp. 334–341, 1996.

[16] G. Loughran *et al.*, "Evidence of efficient stop codon readthrough in four mammalian genes," *Nucleic Acids Res.*, vol. 42, no. 14, pp. 8928–8938, 2014.

[17] F. Schueren *et al.*, "Peroxisomal lactate dehydrogenase is generated by translational readthrough in mammals," *Elife*, vol. 3, pp. 1–24, 2014.

[18] B. Roy, J. D. Leszyk, D. A. Mangus, and A. Jacobson, "Nonsense suppression by near-cognate tRNAs employs alternative base pairing at codon positions 1 and 3," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 112, no. 10, pp. 3038–3043, 2015.

[19] * Liat Linde and B. Kerem, "Introducing sense into nonsense in treatments of human genetic diseases," *Trends Genet.*, vol. 24, no. 11, pp. 552–563, 2008.

[20] J. Sun, M. Chen, J. Xu, and J. Luo, "Relationships among stop codon usage bias, its context, isochores, and gene expression level in various eukaryotes," *J. Mol. Evol.*, vol. 61, no. 4, pp. 437–444, 2005.

[21] K. Zerfass and H. Beier, "The leaky UGA termination codon of tobacco rattle virus RNA is suppressed by tobacco chloroplast and cytoplasmic tRNAs(Trp) with CmCA anticodon.," *EMBO J.*, vol. 11, no. 11, pp. 4167–4173, 1992.

[22] M. Nilsson and M. Rydén-Aulin, "Glutamine is incorporated at the nonsense codons UAG and UAA in a suppressor-free Escherichia coli strain," *Biochim. Biophys. Acta - Gene Struct. Expr.*, vol. 1627, no. 1, pp.

1–6, 2003.
[23] S. Blanchet, D. Cornu, M. Argentini, and O. Namy, “New insights into the incorporation of natural suppressor tRNAs at stop codons in Saccharomyces cerevisiae,” *Nucleic Acids Res.*, vol. 42, no. 15, pp. 10061–10072, 2014.
[24] W. P. Tate *et al.*, “Translational termination efficiency in both bacteria and mammals is regulated by the base following the stop codon.,” *Biochem. Cell Biol.*, vol. 73, no. 11–12, pp. 1095–1103, 1995.
[25] O. Namy, I. Hatin, and J. P. Rousset, “Impact of the six nucleotides downstream of the stop codon on translation termination,” *EMBO Rep.*, vol. 2, no. 9, pp. 787–793, 2001.
[26] P. Beznosková, S. Gunišová, and L. S. Valášek, “Rules of UGA-N decoding by near-cognate tRNAs and analysis of readthrough on short uORFs in yeast,” *RNA*, vol. 22, no. 3, pp. 456–466, Mar. 2016.
[27] P. Beznosková, S. Wagner, M. E. Jansen, T. Von Der Haar, and L. S. Valášek, “Translation initiation factor eIF3 promotes programmed stop codon readthrough,” *Nucleic Acids Res.*, vol. 43, no. 10, pp. 5099–5111, Apr. 2015.
[28] P. Beznosková, Z. Pavlíková, J. Zeman, C. Echeverría Aitken, and L. S. Valášek, “Yeast applied readthrough inducing system (YARIS): an invivo assay for the comprehensive study of translational readthrough,” *Nucleic Acids Res.*, vol. 47, no. 12, pp. 6339–6350, Jul. 2019.
[29] J. M. Skuzeski, L. M. Nichols, R. F. Gesteland, and J. F. Atkins, “The signal for a leaky UAG stop codon in several plant viruses includes the two downstream codons,” *J. Mol. Biol.*, vol. 218, no. 2, pp. 365–373, 1991.
[30] D. M. Janzen, L. Frolova, and A. P. Geballe, “Inhibition of Translation Termination Mediated by an Interaction of Eukaryotic Release Factor 1 with a Nascent Peptidyl-tRNA,” *Mol. Cell. Biol.*, vol. 22, no. 24, pp. 8562–8570, 2002.
[31] C. Echeverría Aitken *et al.*, “Eukaryotic translation initiation factor 3 plays distinct roles at the mRNA entry and exit channels of the ribosomal preinitiation complex,” *Elife*, vol. 5, pp. 1–37, 2016.
[32] L. Valášek, A. A. Mathew, B. Shin, K. H. Nielsen, B. Szamecz, and A. G. Hinnebusch, “and eIF5 make critical connections with the 40S ribosome in vivo,” *Genes Dev.*, vol. 3, pp. 786–799, 2003.
[33] L. Valášek, K. H. Nielsen, and A. G. Hinnebusch, “Direct eIF2-eIF3 contact in the multifactor complex is important for translation initiation in vivo,” *EMBO J.*, vol. 21, no. 21, pp. 5886–5898, 2002.
[34] L. Valasek, K. H. Nielsen, F. Zhang, C. A. Fekete, and A. G. Hinnebusch, “Interactions of Eukaryotic Translation Initiation Factor 3 (eIF3) Subunit NIP1/c with eIF1 and eIF5 Promote Preinitiation Complex Assembly and Regulate Start Codon Selection,” *Mol. Cell. Biol.*, vol. 24, no. 21, pp. 9437–9455, 2004.
[35] T. Kouba *et al.*, “Small ribosomal protein RPS0 stimulates translation initiation by mediating 40S-binding of eIF3 via its direct contact with the eIF3a/TIF32 subunit,” *PLoS One*, vol. 7, no. 7, pp. 1–9, 2012.
[36] T. Kouba, E. Rutkai, M. Karásková, and L. S. Valášek, “The eIF3c/NIP1 PCI domain interacts with RNA and RACK1/ASC1 and promotes assembly of translation preinitiation complexes,” *Nucleic Acids Res.*, vol. 40, no. 6, pp. 2683–2699, 2012.
[37] S. Khoshnevis *et al.*, “Structural integrity of the PCI domain of eIF3a/TIF32 is required for mRNA recruitment to the 43S pre-initiation complexes,” *Nucleic Acids Res.*, vol. 42, no. 6, pp. 4123–4139, 2014.
[38] P. Beznosková *et al.*, “Translation Initiation Factors eIF3 and HCR1 Control Translation Termination and Stop Codon Read-Through in Yeast Cells,” *PLoS Genet.*, vol. 9, no. 11, Nov. 2013.
[39] K. Poncová *et al.*, “uS3/Rps3 controls fidelity of translation termination and programmed stop codon readthrough in co-operation with eIF3,” *Nucleic Acids Res.*, vol. 47, no. 21, pp. 11326–11343, 2019.
[40] A. Des Georges *et al.*, “Structure of the mammalian ribosomal pre-termination complex associated with eRF1•eRF3•GDPNP,” *Nucleic Acids Res.*, vol. 42, no. 5, pp. 3409–3418, 2014.
[41] Y. Yamaguchi, A. Hayashi, C. W. Campagnoni, A. Kimura, T. Inuzuka, and B. Hiroko, “L-MPZ , a Novel Isoform of Myelin P0 , Is Produced by Stop Codon Readthrough,” *J. Biol. Chem.*, vol. 287, no. 21, pp. 17765–17776, 2012.
[42] A. C. Stiebler *et al.*, “Ribosomal Readthrough at a Short UGA Stop Codon Context Triggers Dual Localization of Metabolic Enzymes in Fungi and Animals,” *PLoS Genet.*, vol. 10, no. 10, pp. 1–9, 2014.
[43] J. Weiner, F. Beaussart, and E. Bornberg-Bauer, “Domain deletions and substitutions in the modular protein evolution,” *FEBS J.*, vol. 273, no. 9, pp. 2037–2047, 2006.
[44] J. R. Hogg and S. P. Goff, “Upf1 senses 3’UTR length to potentiate mRNA decay,” *Cell*, vol. 143, no. 3, pp. 379–389, 2010.
[45] K. M. Keeling *et al.*, “Leaky termination at premature stop codons antagonizes nonsense-mediated mRNA decay in S . cerevisiae,” *RNA*, vol. 10, pp. 691–703, 2004.
[46] T. Nyiko *et al.*, “Expression of the eRF1 translation termination factor is controlled by an autoregulatory

circuit involving readthrough and nonsense-mediated decay in plants Nyik o," *Nucleic Acids Res.*, vol. 45, no. 7, pp. 4174–4188, 2017.
[47] H. R. B. Pelham, "Leaky UAG termination codon in tobacco mosaic virus RNA," *Nature*, vol. 272, no. 5652. pp. 469–471, 1978.
[48] O. Namy, G. Duchateau-Nguyen, and J. P. Rousset, "Translational readthrough of the PDE2 stop codon modulates cAMP levels in Saccharomyces cerevisiae," *Mol. Microbiol.*, vol. 43, no. 3, pp. 641–652, 2002.
[49] I. Jungreis *et al.*, "Evidence of abundant stop codon readthrough in Drosophila and other metazoa," *Genome Res.*, vol. 21, pp. 2096–2113, 2011.
[50] J. G. Dunn, C. K. Foo, N. G. Belletier, E. R. Gavis, and J. S. Weissman, "Ribosome profiling reveals pervasive and regulated stop codon readthrough in Drosophila melanogaster," *Elife*, vol. 2013, no. 2, pp. 1–32, 2013.
[51] H. S. Chittum *et al.*, "Rabbit β-globin is extended beyond its UGA stop codon by multiple suppressions and translational reading gaps," *Biochemistry*, vol. 37, no. 31, pp. 10866–10870, 1998.
[52] K. Hayasaka *et al.*, "De novo mutation of the myelin P0 gene in Dejerine-Sottas disease (hereditary motor and sensory neuropathy type III).," *Nat. Genet.*, vol. 4, pp. 266–268, 1993.
[53] A. V Anzalone, S. Zairis, A. J. Lin, R. Rabadan, and W. Virginia, "Interrogation of Eukaryotic Stop Codon Readthrough Signals by in Vitro RNA Selection," *Biochemistry*, vol. 58, no. 8, pp. 1167–1178, 2019.
[54] S. M. Eswarappa *et al.*, "Programmed translational readthrough generates antiangiogenic VEGF-Ax," *Cell*, vol. 157, no. 7, pp. 1605–1618, 2014.
[55] * Sven Thoms, "Import of proteins into peroxisomes: Piggybacking to a new home away from home," *Open Biol.*, vol. 5, no. 11, pp. 0–3, 2015.
[56] G. Loughran *et al.*, "Stop codon readthrough generates a C-terminally extended variant of the human Vitamin D receptor with reduced calcitriol response," *J. Biol. Chem.*, vol. 293, no. 12, pp. 4434–4444, 2018.
[57] * Peter D. Stenson *et al.*, "The Human Gene Mutation Database: towards a comprehensive repository of inherited mutation data for medical research, genetic diagnosis and next-generation sequencing studies," *Hum. Genet.*, vol. 136, no. 6, pp. 665–677, 2017.
[58] P. D. Stenson *et al.*, "The human gene mutation database: 2008 update," *Genome Med.*, vol. 1, no. 1, pp. 1–6, 2009.
[59] * Hui Ling Rose Lee and J. P. Dougherty, "Pharmaceutical therapies to recode nonsense mutations in inherited diseases," *Pharmacology and Therapeutics*, vol. 136, no. 2. pp. 227–266, Nov-2012.
[60] W. J. Friesen *et al.*, "The minor gentamicin complex component, X2, is a potent premature stop codon readthrough molecule with therapeutic potential," *PLoS One*, vol. 13, no. 10, pp. 1–18, 2018.
[61] L. Bidou, O. Bugaud, V. Belakhov, T. Baasov, and O. Namy, "Characterization of new-generation aminoglycoside promoting premature termination codon readthrough in cancer cells," *RNA Biol.*, vol. 14, no. 3, pp. 378–388, 2017.
[62] N. T. Ingolia, L. F. Lareau, and J. S. Weissman, "Ribosome Profiling of Mouse Embryonic Stem Cells Reveals the Complexity of Mammalian Proteomes," *Cell*, vol. 147, no. 4, pp. 789–802, 2011.
[63] N. T. Ingolia, S. Ghaemmaghami, J. R. S. Newman, and J. S. Weissman, "Genome-Wide Analysis in Vivo of Translation with Nucleotide Resolution Using Ribosome Profiling," *Science (80-.).*, vol. 324, no. 5924, pp. 218–223, 2009.
[64] A. S. Kleppe and E. Bornberg-Bauer, "Robustness by intrinsically disordered C-termini and translational readthrough," *Nucleic Acids Res.*, vol. 46, no. 19, pp. 10184–10194, 2018.
[65] I. Jungreis, C. S. Chan, R. M. Waterhouse, G. Fields, M. F. Lin, and M. Kellis, "Evolutionary dynamics of abundant stop codon readthrough," *Mol. Biol. Evol.*, vol. 33, no. 12, pp. 3108–3132, 2016.
[66] G. Grentzmann, J. A. Ingram, P. J. Kelly, R. F. Gesteland, and J. F. Atkins, "A dual-luciferase reporter system for studying recoding signals," *RNA*, vol. 4, no. 4, pp. 479–486, 1998.
[67] P. J. Halvey, D. C. Liebler, and R. J. C. Slebos, "A reporter system for translational readthrough of stop codons in human cells," *FEBS openbio*, vol. 2, pp. 56–59, 2012.
[68] D. S. Auld, N. Thorne, W. F. Maguire, and J. Inglese, "Mechanism of PTC124 activity in cell-based luciferase assays of nonsense codon suppression," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 106, no. 9, pp. 3585–3590, 2009.
[69] * Sergey Melnikov, A. Ben-Shem, N. Garreau De Loubresse, L. Jenner, G. Yusupova, and M. Yusupov, "One core, two shells: Bacterial and eukaryotic ribosomes," *Nat. Struct. Mol. Biol.*, vol. 19, no. 6, pp. 560–567, 2012.
[70] J. Frank *et al.*, "A model of the translational apparatus based on a three-dimensional reconstruction of

the Escherichia coli ribosome.," *Biochem. cell Biol.*, vol. 73, no. 11–12, pp. 757–765, 1995.
[71] C. M. T. Spahn *et al.*, "Structure of the 80S ribosome from Saccharomyces cerevisiae - tRNA-ribosome and subunit-subunit interactions," *Cell*, vol. 107, no. 3, pp. 373–386, 2001.
[72] A. Ben-Shem, N. G. de Loubresse, S. Melnikov, L. Jenner, G. Yusupova, and M. Yusupov, "The Structure of the Eukaryotic Ribosome at 3.0 Å Resolution," *Science (80-.).*, vol. 334, pp. 1524–1529, 2011.
[73] Ben-Shem et al., "Crystal Structure of the Eukaryotic Ribosome," *Science (80-.).*, vol. 330, pp. 1203–1209, 2010.
[74] S. Klinge, F. Voigts-Hoffmann, M. Leibundgut, S. Arpagaus, and N. Ban, "Crystal structure of the eukaryotic 60S ribosomal subunit in complex with initiation factor 6," *Science (80-.).*, vol. 334, no. 6058, pp. 941–948, 2011.
[75] M. Weisser, F. Voigts-hoffmann, J. Rabl, M. Leibundgut, and N. Ban, "The crystal structure of the eukaryotic 40S ribosomal subunit in complex with eIF1 and eIF1A," *Nat. Struct. Mol. Biol.*, vol. 1708, pp. 1–4, 2013.
[76] * Susan A. Gerbi, "The evolution of eukaryotic ribosomal DNA," *BioSystems*, vol. 19, no. 4, pp. 247–258, 1986.
[77] O. Lecompte, R. Ripp, J. C. Thierry, D. Moras, and O. Poch, "Comparative analysis of ribosomal proteins in complete genomes: An example of reductive evolution at the domain scale," *Nucleic Acids Res.*, vol. 30, no. 24, pp. 5382–5390, 2002.
[78] T. F. Smith, J. C. Lee, R. R. Gutell, and H. Hartman, "The origin and evolution of the ribosome," *Biol. Direct*, vol. 3, no. 16, pp. 1–13, 2008.
[79] M. Selmer *et al.*, "Structure of the 70S ribosome complexed with mRNA and tRNA," *Science (80-.).*, vol. 313, no. 5795, pp. 1935–1942, 2006.
[80] M. M. Yusupov *et al.*, "Crystal structure of the ribosome at 5.5 Å resolution," *Science (80-.).*, vol. 292, no. 5518, pp. 883–896, 2001.
[81] S. Gerbi, "Expansion segments: regions of variable size that interrupt the universal core secondary structure of ribosomal RNA," *Ribosomal RNA Struct. Evol. Process. Funct. Protein Synth.*, no. eds: R.A. Zimmermann and A.E. Dahlberg. Telford-CRC Press, Boca Raton, FL, pp. 71–87, 1996.
[82] J. J. Cannone *et al.*, "The Comparative RNA Web (CRW) Site: An online database of comparative sequence and structure information for ribosomal, intron, and other RNAs," *BMC Bioinformatics*, vol. 3, pp. 1–31, 2002.
[83] * Wendy V. Gilbert, "Functional Specialization of Ribosomes?," *Trends Biochem Sci.*, vol. 36, no. 3, pp. 127–132, 2011.
[84] K. B. McIntosh and J. R. Warner, "Yeast Ribosomes: Variety Is the Spice of Life," *Cell*, vol. 131, no. 3, pp. 450–451, 2007.
[85] E. M. Panina, A. A. Mironov, and M. S. Gelfand, "Comparative genomics of bacterial zinc regulons: Enhanced ion transport, pathogenesis, and rearrangement of ribosomal proteins," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 100, no. 17, pp. 9912–9917, 2003.
[86] O. Vesper *et al.*, "Selective translation of leaderless mRNAs by specialized ribosomes generated by MazF in Escherichia coli," *Cell*, vol. 147, no. 1, pp. 147–157, 2011.
[87] J. H. Gunderson *et al.*, "Structurally Distinct, Stage-Specific Ribosomes Occur in Plasmodium," *Science (80-.).*, pp. 933–938, 1987.
[88] N. R. Voss, M. Gerstein, T. A. Steitz, and P. B. Moore, "The Geometry of the Ribosomal Polypeptide Exit Tunnel," *J. Mol. Biol.*, vol. 360, no. 4, pp. 893–906, 2006.
[89] N. A. Paolini *et al.*, "A Ribosomopathy Reveals Decoding Defective Ribosomes Driving Human Dysmorphism," *Am. J. Hum. Genet.*, vol. 100, no. 3, pp. 506–522, 2017.
[90] L. B. Jenner, N. Demeshkina, G. Yusupova, and M. Yusupov, "Structural aspects of messenger RNA reading frame maintenance by the ribosome," *Nat. Struct. Mol. Biol.*, vol. 17, no. 5, pp. 555–560, 2010.
[91] G. Z. Yusupova, M. M. Yusupov, J. H. D. Cate, and H. F. Noller, "The path of messenger RNA through the ribosome," *Cell*, vol. 106, no. 2, pp. 233–241, 2001.
[92] B. Siridechadilok, C. S. Fraser, R. J. Hall, J. A. Doudna, and E. Nogales, "Molecular biology: Structural roles for human translation factor eIF3 in initiation of protein synthesis," *Science (80-.).*, vol. 310, no. 5753, pp. 1513–1515, 2005.
[93] Y. Yu, I. S. Abaeva, A. Marintchev, T. V. Pestova, and C. U. T. Hellen, "Common conformational changes induced in type 2 picornavirus IRESs by cognate trans-acting factors," *Nucleic Acids Res.*, vol. 39, no. 11, pp. 4851–4865, 2011.
[94] J. Zeman *et al.*, "Binding of eIF3 in complex with eIF5 and eIF1 to the 40S ribosomal subunit is accompanied by dramatic structural changes," *Nucleic Acids Res.*, vol. 47, no. 15, pp. 8282–8300, 2019.

[95] L. Jenner, N. Demeshkina, G. Yusupova, and M. Yusupov, “Structural rearrangements of the ribosome at the tRNA proofreading step,” *Nat. Struct. Mol. Biol.*, vol. 17, no. 9, pp. 1072–1078, 2010.
[96] T. Budkevich *et al.*, “Structure and dynamics of the mammalian ribosomal pre-translocation complex,” *Mol. Cell*, vol. 44, no. 2, pp. 214–224, 2011.
[97] * James M. Ogle and V. Ramakrishnan, “Structural Insights Into Translational Fidelity,” *Annu. Rev. Biochem.*, vol. 74, no. 1, pp. 129–177, 2005.
[98] T. M. Schmeing, R. M. Voorhees, A. C. Kelley, and V. Ramakrishnan, “How mutations in tRNA distant from the anticodon affect the fidelity of decoding,” *Nat. Struct. Mol. Biol.*, vol. 18, no. 4, pp. 432–437, 2011.
[99] J.-P. et al. Armache, “Localization of eukaryote-specific ribosomal proteins in a 5.5-Å cryo-EM map of the 80S eukaryotic ribosome,” *Proc Natl Acad Sci*, vol. 107, no. 46, pp. 19754– 19759, 2010.
[100] J. Rabl, M. Leibundgut, S. F. Ataide, A. Haag, and N. Ban, “Crystal structure of the eukaryotic 40S ribosomal subunit in complex with initiation factor 1,” *Science (80-.).*, vol. 331, no. 6018, pp. 730–736, 2011.
[101] P. V. Baranov, J. F. Atkins, and M. M. Yordanova, “Augmented genetic decoding: Global, local and temporal alterations of decoding processes and codon meaning,” *Nat. Rev. Genet.*, vol. 16, no. 9, pp. 517–529, 2015.
[102] * Xumin Ou, J. Cao, A. Cheng, M. P. Peppelenbosch, and Q. Pan, “Errors in translational decoding: tRNA wobbling or misincorporation?,” *PLoS Genet.*, vol. 15, no. 3, pp. 1–15, 2019.
[103] O. Uhlenbeck, J. Baller, and P. Doty, “Complementary oligonucleotide binding to the anticodon loop of fMet-transfer RNA,” *Nature*, vol. 225, pp. 508–510, 1970.
[104] J. M. Ogle, D. E. Brodersen, J. Clemons, M. J. Tarry, A. P. Carter, and V. Ramakrishnan, “Recognition of cognate transfer RNA by the 30S ribosomal subunit,” *Science (80-.).*, vol. 292, no. 5518, pp. 897–902, 2001.
[105] U. Lagerkvist, “‘Two out of three’: An alternative method for codon reading,” *Biochemistry*, vol. 75, no. 4, pp. 1759–1762, 1978.
[106] S. Alkatib *et al.*, “The Contributions of Wobbling and Superwobbling to the Reading of the Genetic Code,” *PLoS Genet.*, vol. 8, no. 11, pp. 1–16, 2012.
[107] * Paul F. Agris, E. R. Eruysal, A. Narendran, V. Y. P. Väre, S. Vangaveti, and S. V. Ranganathan, “Celebrating wobble decoding: Half a century and still much is new,” *RNA Biol.*, vol. 15, no. 4–5, pp. 537–553, 2018.
[108] F. H. C. Crick, “Codon—anticodon pairing: The wobble hypothesis,” *J. Mol. Biol.*, vol. 19, no. 2, pp. 548–555, 1966.
[109] N. Ban *et al.*, “A new system for naming ribosomal proteins,” *Curr Opin Struct Biol.*, vol. 24, pp. 165–169, 2014.
[110] D. R. Zerbino *et al.*, “Ensembl 2018,” *Nucleic Acids Res.*, vol. 46, no. D1, pp. D754–D761, 2018.
[111] A. Bateman, “UniProt: A worldwide hub of protein knowledge,” *Nucleic Acids Res.*, vol. 47, no. D1, pp. D506–D515, 2019.
[112] F. Sievers *et al.*, “Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega,” *Mol. Syst. Biol.*, vol. 7, no. 539, 2011.
[113] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman, “Basic local alignment search tool,” *J. Mol. Biol.*, vol. 215, no. 3, pp. 403–410, 1990.
[114] J. Noeske, M. R. Wasserman, D. S. Terry, R. B. Altman, S. C. Blanchard, and J. H. D. Cate, “High-resolution structure of the Escherichia coli ribosome,” *Nat. Struct. Mol. Biol.*, vol. 22, no. 4, pp. 336–341, 2015.
[115] * James M. Ogle, A. P. Carter, and V. Ramakrishnan, “Insights into the decoding mechanism from recent ribosome structures,” *Trends Biochem. Sci.*, vol. 28, no. 5, pp. 259–266, 2003.
[116] Yury S. Polikanov, T. A. Steitz, and C. A. Innis, “A proton wire to couple aminoacyl-tRNA accommodation and peptide bond formation on the ribosome,” *Nat Struct Mol Biol.*, vol. 21, no. 9, pp. 787–793, 2014.
[117] Y. S. Polikanov *et al.*, “Amicoumacin A inhibits translation by stabilizing mRNA interaction with the ribosome,” *Mol Cell*, vol. 56, no. 4, pp. 531–540, 2014.
[118] D. E. Sharifulin *et al.*, “Molecular contacts of ribose-phosphate backbone of mRNA with human ribosome,” *Biochim. Biophys. Acta - Gene Regul. Mech.*, vol. 1849, no. 8, pp. 930–939, 2015.
[119] D. Sharma, A. R. Cukras, E. J. Rogers, D. R. Southworth, and R. Green, “Mutational Analysis of S12 Protein and Implications for the Accuracy of Decoding by the Ribosome,” *J. Mol. Biol.*, vol. 374, no. 4, pp. 1065–1076, 2007.
[120] R. A. Anthony and S. W. Liebman, “Alterations in ribosomal protein RPS28 can diversely affect

translational accuracy in Saccharomyces cerevisiae," *Genetics*, vol. 140, no. 4, pp. 1247–1258, 1995.
[121] D. Synetos, C. P. Frantziou, and L. E. Alksne, "Mutations in yeast ribosomal proteins S28 and S4 affect-the accuracy of translation and alter the sensitivity of the ribosomes to paromomycin," *Biochim. Biophys. Acta*, vol. 1309, pp. 156–166, 1996.
[122] R. S. Singleton *et al.*, "OGFOD1 catalyzes prolyl hydroxylation of RPS23 and is involved in translation control and stress granule formation," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 111, no. 11, pp. 4031–4036, 2014.
[123] M. Zakari, R. Trimble Ross, A. Peak, M. Blanchette, C. Seidel, and J. L. Gerton, "The SMC Loader Scc2 Promotes ncRNA Biogenesis and Translational Fidelity," *PLoS Genet.*, vol. 11, no. 7, pp. 1–25, 2015.
[124] C. Loenarz *et al.*, "Hydroxylation of the eukaryotic ribosomal decoding center affects translational accuracy," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 111, no. 11, pp. 4019–4024, 2014.
[125] P. Chandramouli *et al.*, "Structure of the Mammalian 80S Ribosome at 8.7Å Resolution," *Structure*, vol. 16, no. 4, pp. 535–548, 2008.
[126] K. M. Keeling, J. Salas-Marco, L. Z. Osherovich, and D. M. Bedwell, "Tpa1p Is Part of an mRNP Complex That Influences Translation Termination, mRNA Deadenylation, and mRNA Turnover in Saccharomyces cerevisiae," *Mol. Cell. Biol.*, vol. 26, no. 14, pp. 5237–5248, 2006.
[127] J. Henri, D. Rispal, E. Bayart, H. Van Tilbeurgh, B. Séraphin, and M. Graille, "Structural and functional insights into Saccharomyces cerevisiae Tpa1, a putative prolylhydroxylase influencing translation termination and transcription," *J. Biol. Chem.*, vol. 285, no. 40, pp. 30767–30778, 2010.
[128] M. G. Campbel and K. Karbstein, "Protein-protein interactions within late pre-40S ribosomes," *PLoS One*, vol. 6, no. 1, pp. 1–10, 2011.
[129] J. Khairulina, D. Graifer, K. Bulygin, A. Ven'yaminova, L. Frolova, and G. Karpova, "Eukaryote-specific motif of ribosomal protein S15 neighbors A site codon during elongation and termination of translation," *Biochimie*, vol. 92, no. 7, pp. 820–825, Jul. 2010.
[130] A. M. Bowen, S. Musalgaonkar, C. A. Moomau, S. P. Gulay, M. Mirvis, and J. D. Dinman, "Ribosomal protein uS19 mutants reveal its role in coordinating ribosome structure and function," *Translation*, vol. 3, no. 2, pp. 1–8, 2015.
[131] R. Rakauskaite and J. D. Dinman, "An Arc of Unpaired 'Hinge Bases' Facilitates Information Exchange among Functional Centers of the Ribosome," *Mol. Cell. Biol.*, vol. 26, no. 23, pp. 8992–9002, 2006.
[132] K. Bulygin, A. Malygin, A. Gopanenko, D. Graifer, and G. Karpova, "The functional role of the C-terminal tail of the human ribosomal protein uS19," *BBA - Gene Regul. Mech.*, vol. 1863, no. 3, pp. 1–8, 2020.
[133] G. Bretones *et al.*, "Altered patterns of global protein synthesis and translational fidelity in RPS15-mutated chronic lymphocytic leukemia," *Blood*, vol. 132, no. 22, pp. 2375–2388, Nov. 2018.
[134] I. Léger-Silvestre *et al.*, "The ribosomal protein Rps15p is required for nuclear exit of the 40S subunit precursors in yeast," *EMBO J.*, vol. 23, no. 12, pp. 2336–2347, 2004.
[135] F. Thomas and U. Kutay, "Biogenesis and nuclear export of ribosomal subunits in higher eukaryotes depend on the CRM1 export pathway," *J. Cell Sci.*, vol. 116, no. 12, pp. 2409–2419, 2003.
[136] J. Rouquette, V. Choesmel, and P. E. Gleizes, "Nuclear export and cytoplasmic processing of precursors to the 40S ribosomal subunits in mammalian cells," *EMBO J.*, vol. 24, no. 16, pp. 2862–2872, 2005.
[137] D. A. Landau *et al.*, "Mutations driving CLL and their evolution in progression and relapse," *Nature*, vol. 526, no. 7574, pp. 525–530, 2015.
[138] L. Daftuar, Y. Zhu, X. Jacq, and C. Prives, "Ribosomal Proteins RPL37, RPS15 and RPS20 Regulate the Mdm2-p53-MdmX Network," *PLoS One*, vol. 8, no. 7, pp. 1–13, Jul. 2013.
[139] H. T. Gazda *et al.*, "Ribosomal Protein L5 and L11 Mutations Are Associated with Cleft Palate and Abnormal Thumbs in Diamond-Blackfan Anemia Patients," *Am. J. Hum. Genet.*, vol. 83, no. 6, pp. 769–780, 2008.
[140] * Daniel N. Wilson and J. H. D. Cate, "The structure and function of the eukaryotic ribosome," *Cold Spring Harb Perspect Biol*, vol. 4, pp. 1–17, 2012.
[141] D. M. Landry, M. I. Hertz, and S. R. Thompson, "RPS25 is essential for translation initiation by the Dicistroviridae and hepatitis C viral IRESs," *Genes Dev.*, vol. 23, no. 23, pp. 2753–2764, Dec. 2009.
[142] S. Robledo, R. A. Idol, D. L. Crimmins, J. H. Ladenson, P. J. Mason, and M. Bessler, "The role of human ribosomal proteins in the maturation of rRNA and ribosome production," *RNA*, vol. 14, no. 9, pp. 1918–1929, 2008.
[143] * Naomi R. Genuth and M. Barna, "The discovery of ribosome heterogeneity and its implications for gene regulation and organismal life," *Mol Cell*, vol. 71, no. 3, pp. 364–374, 2018.
[144] M. C. Bassik *et al.*, "A Systematic Mammalian Genetic Interaction Map Reveals Pathways Underlying Ricin Susceptibility Michael," *Cell*, vol. 152, no. 4, pp. 909–922, 2013.

[145] T. Nishiyama, H. Yamamoto, T. Uchiumi, and N. Nakashima, “Eukaryotic ribosomal protein RPS25 interacts with the conserved loop region in a dicistroviral intergenic internal ribosome entry site,” *Nucleic Acids Res.*, vol. 35, no. 5, pp. 1514–1521, 2007.

[146] F. Carvajal *et al.*, “Structural domains within the HIV-1 mRNA and the ribosomal protein S25 influence cap-independent translation initiation,” *FEBS J.*, vol. 283, no. 13, pp. 2508–2527, 2016.

[147] M. I. Hertz, D. M. Landry, A. E. Willis, G. Luo, and S. R. Thompson, “Ribosomal Protein S25 Dependency Reveals a Common Mechanism for Diverse Internal Ribosome Entry Sites and Ribosome Shunting,” *Mol. Cell. Biol.*, vol. 33, no. 5, pp. 1016–1026, 2013.

[148] X. Zhang, W. Wang, H. Wang, M.-H. Wang, W. Xu, and R. Zhang, “Identification of ribosomal protein S25 (RPS25)–MDM2–p53 regulatory feedback loop,” *Oncogene*, vol. 32, no. 22, pp. 2782–2791, 2013.

[149] E. U. H. Sim and S. P. Talwar, “In silico evidence of de novo interactions between ribosomal and Epstein - Barr virus proteins,” *BMC Mol. Cell Biol.*, vol. 20, no. 1, pp. 1–12, 2019.

[150] S. Kubota, T. D. Copeland, and R. J. Pomerantz, “Nuclear and nucleolar targeting of human ribosomal protein S25: Common features shared with HIV-1 regulatory proteins,” *Oncogene*, vol. 18, no. 7, pp. 1503–1514, 1999.

[151] R. T. Baker, N. A. Williamson, and R. E. H. Wettenhall, “The yeast homolog of mammalian ribosomal protein S30 is expressed from a duplicated gene without a ubiquitin-like protein fusion sequence: Evolutionahy implications,” *J. Biol. Chem.*, vol. 271, no. 23, pp. 13549–13555, 1996.

[152] E. S. Babaylova *et al.*, “mRNA regions where 80S ribosomes pause during translation elongation in vivo interact with protein uS19, a component of the decoding site,” *Nucleic Acids Res.*, vol. 48, no. 2, pp. 912–923, 2020.

[153] I. B. Lomakin and T. A. Steitz, “The initiation of mammalian protein synthesis and mRNA scanning mechanism,” *Nature*, vol. 500, no. 7462, pp. 307–311, 2013.

[154] A. M. Anger *et al.*, “Structures of the human and Drosophila 80S ribosome,” *Nature*, vol. 497, no. 7447, pp. 80–85, 2013.