BBA - Gene Regulatory Mechanisms xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

BBA - Gene Regulatory Mechanisms

journal homepage: www.elsevier.com/locate/bbagrm

Review

Post-transcriptional pseudouridylation in mRNA as well as in some major types of noncoding RNAs☆

Hironori Adachi, Meemanage D. De Zoysa, Yi-Tao Yu*

University of Rochester Medical Center, Department of Biochemistry and Biophysics, Center for RNA Biology, 601 Elmwood Avenue, Rochester, NY 14642, USA

A B S T R A C T

Pseudouridylation is a post-transcriptional isomerization reaction that converts a uridine to a pseudouridine (Ψ) within an RNA chain. Ψ has chemical properties that are distinct from that of uridine and any other known nucleotides. Experimental data accumulated thus far have indicated that Ψ is present in many different types of RNAs, including coding and noncoding RNAs. Ψ is particularly concentrated in rRNA and spliceosomal snRNAs, and plays an important role in protein translation and pre-mRNA splicing, respectively. Ψ has also been found in mRNA, but its function there remains essentially unknown. In this review, we discuss the mechanisms and functions of RNA pseudouridylation, focusing on rRNA, snRNA and mRNA. We also discuss the methods, which have been developed to detect Ψs in RNAs. This article is part of a Special Issue entitled: mRNA modifications in gene expression control edited by Dr. Soller Matthias and Dr. Fray Rupert.

1. Introduction

Pseudouridine (Ψ) is a modified ribonucleotide that was first discovered in 1957. Due to its high abundance, Ψ, initially called "?", was thought to be a 5th core nucleotide when it was initially identified [1], but was renamed "pseudouridine" two years later when detailed characterization indicated that it was an isomer (5-ribosyluracil) of uridine (1-ribosyluracil) [2]. Over the past 6 decades, Ψ has drawn much experimental attention, which has yielded some fruitful results.

Ψ is converted from uridine through a post-transcriptional isomerization reaction, called pseudouridylation (Fig. 1A). This reaction begins with the breaking of the N1-C1′ glycosidic bond, allowing the base to be rotated 180° around the C6-N3 axis. A new bond is then established between C5 and C1′, completing the formation of Ψ. In addition to the unusual carbon-carbon bond between base and sugar formed in this reaction, pseudouridylation also results in an additional hydrogen bond donor (the N1 position). Thus, Ψ's chemical properties are distinct from those of uridine, or, indeed, those of any other known nucleotides [3]. Extensive studies have shown that these distinct properties have a significant impact on RNA strands [4]. For instance, Ψ can help rigidify the RNA backbone (locked in a C-3′ endo sugar ring configuration) through a water-mediated hydrogen bond network [3] (Fig. 1B and C), or increase base stacking and thermal stability in an RNA-RNA duplex [5]. It has also been reported that Ψ can alter the local structure of an RNA [6–9].

Ψ is the most abundant and widespread of all modified nucleotides found in RNA. It is present in all species and in many different types of RNAs, including, to date, tRNA, rRNA, snRNA, box H/ACA RNA, box C/D RNA, telomerase RNA, 7SK RNA, RNase MRP RNA, and Steroid Receptor RNA Activator (SRA), as well as a large number of coding RNAs (mRNAs) [10–15]. But how are Ψs introduced into RNAs, and how are they detected? What do we know about the function of RNA pseudouridylation? Here, we will discuss the methods developed for Ψ detection and the mechanisms by which pseudouridylation is introduced into RNAs. We will also highlight some of recent experimental results concerning pseudouridylation, focusing primarily on mRNA pseudouridylation.

2. Methods for detection and quantification of RNA pseudouridylation

A necessary prerequisite for successful pseudouridylation research is a means of detecting Ψ in RNA. However, due to the fact that they are isomers, Ψ and uridine share exactly the same molecular weight. Also, like uridine, Ψ pairs with and is read by adenosine during RNA-templated reverse-transcription. Thus, it is difficult to distinguish between Ψ and uridine in RNA. Indeed, in the early days of pseudouridylation research, identification of Ψ was a lengthy, laborious process, often requiring a number of difficult steps [16,17]. Nonetheless, through great effort, two different methods have been developed to identify and quantify RNA pseudouridylation, and their use is now widespread in the field [18–20].

☆ This article is part of a Special Issue entitled: mRNA modifications in gene expression control edited by Dr. Soller Matthias and Dr. Fray Rupert.

* Corresponding author at: Department of Biochemistry and Biophysics, Center for RNA Biology, University of Rochester Medical Center, 601 Elmwood Avenue, Rochester, NY 14642, USA.

E-mail address: yitao_yu@urmc.rochester.edu (Y.-T. Yu).

https://doi.org/10.1016/j.bbagrm.2018.11.002

Received 9 May 2018; Received in revised form 29 October 2018; Accepted 2 November 2018

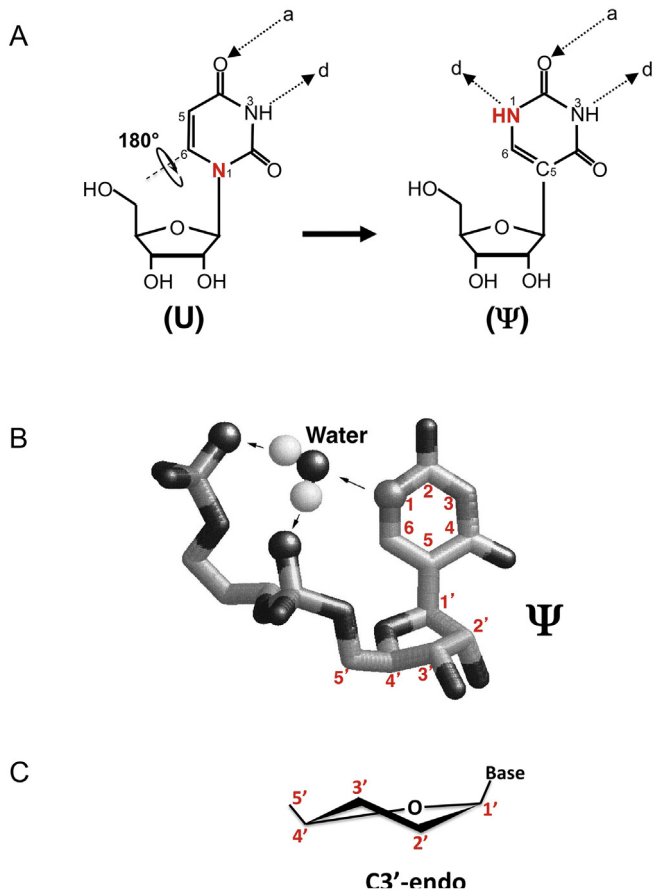

Fig. 1. Pseudouridine properties. (A) U-to-Ψ isomerization. Uridine (U) and pseudouridine (Ψ) are indicated. N1 (red) of the base and the N3-C6 axis are also indicated. a: hydrogen bond acceptor; d: hydrogen bond donor. (B) Water-mediated hydrogen-bond network. Depicted is the Ψ-specific, water-mediated hydrogen-bond network. The water molecule is indicated. The positions of the base and the sugar are also indicated. (C) Sugar ring C3′-endo configuration. Shown is the C3′-endo configuration of the sugar ring, established upon the formation of a Ψ-specific, water-mediated hydrogen bond network. Carbon numbers are indicated.

2.1. Carbodiimide CMC-modification and primer-extension (Primer-extension-based method)

In 1993, Ofengand and colleagues developed a convenient and highly effective pseudouridylation method that can identify Ψ at a one-nucleotide resolution (Fig. 2A) [18]. This method is based on the fact that the organic compound carbodiimide CMC (N-Cyclohexyl-N'-(2-morpholinoethyl)carbodiimide metho-p-toluenesulfonate) can specifically modify (*via* covalent linkage) Ψ, U, and G residues in an RNA. Upon incubation in a bicarbonate buffer (pH 10.4), U-CMC and G-CMC are reversed to unmodified U and G, whereas Ψ-CMC remains covalently linked. Subsequent primer-extension (reverse-transcription) results in stops (or pauses) one nucleotide before the bulky CMC-linked Ψs, thus allowing mapping of Ψs in an RNA. This method is convenient, and in theory can detect multiple Ψs (if there are more than one within an RNA) simultaneously in one reaction. Since it was developed about 25 years ago, this method has been widely used for detecting Ψs within an RNA. Recently, by coupling this method with next-generation sequencing, several labs have developed high-throughput pseudouridine-seq techniques [12–15,21], which have identified a large number of Ψs (see below).

Although extremely useful and effective, CMC-modification followed by primer-extension has some limitations; for instance, this method exhibits relatively low sensitivity and is only semi-quantitative. The limitations are due at least in part to the incomplete nature of both CMC modification and the reversal of U-CMC and G-CMC to unmodified U and G. The former will reduce Ψ signal and the latter will increase false positives in detecting Ψ.

2.2. Site-specific cleavage/labeling and TLC (TLC-based method)

To develop a reliable quantitative pseudouridylation assay, Zhao and Yu took advantage of the ability to site-specifically direct RNase H cleavage using 2′-*O*-methyl RNA-DNA chimeric oligonucleotides [22,23]. This technique allows for site-specific cleavage at any potential pseudouridylation site. After cleavage, the test site is exposed at the end (either 5′ or 3′, depending on what position the directing chimeric oligonucleotide is targeting) of a cleaved RNA fragment, allowing site-specific radiolabeling of the test site nucleotide [19,20,24,25]. The end-radiolabeled RNA fragment is digested by nuclease P1 or T2 (depending on which phosphate, 5′ or 3′ of the target pseudouridylation site, is labeled) to completion. The resulting mononucleotides are then separated by TLC (thin-layer chromatography), revealing the radiolabeled test site nucleotide (either uridine or Ψ) (Fig. 2B). This is a quantitative method and has proven to be highly effective [14,19,25]. The disadvantage of this method is that it involves a number of complex steps, and that it is a low throughput approach that tests only one site at a time (Fig. 2B).

Each method, when used alone, has limitations. However, if combined, the two methods are extremely powerful. For instance, one can use high-throughput CMC-modification followed by primer-extension (and deep sequencing) to detect pseudouridylation sites in RNAs. One can then use low-throughput site-specific cleavage and labeling to verify and quantify pseudouridylation at these sites. Using this approach, Li et al. have definitively shown that some of their newly identified sites in rRNA and mRNA are pseudouridylated [14].

3. Pseudouridylation is catalyzed by two different mechanisms

Pseudouridylation can be catalyzed by two different mechanisms, namely, an RNA-independent mechanism and an RNA-dependent mechanism. While the RNA-independent mechanism is conserved across species from bacteria to humans, the RNA-dependent mechanism is unique to eukaryotes and archaea [26].

3.1. RNA-independent mechanism

RNA-independent pseudouridylation involves stand-alone protein enzymes called pseudouridine synthases (Pus enzymes). There are 9 different types of Pus enzymes (Pus1-Pus9) in *S. cerevisiae*, and 10 (Pus 1-Pus10) in higher eukaryotes [27]. All eukaryotic Pus enzymes can be classified into 5 families (TruA, TruB, RluA, TruD, and Pus10) based on their similarities to their bacterial counterparts [28]. The crystal structures of many Pus enzymes have been solved [29–39]. Although they have different sequences and are categorized into different families, all Pus enzymes share a common catalytic core domain with an invariant aspartate residue participating in the catalytic reaction. In order to site-specially convert a uridine into a Ψ, each Pus enzyme is capable of both independently recognizing its substrates and catalyzing the chemical reaction. While some Pus enzymes have only a few specific substrates, others have many. In eukaryotes, tRNA pseudouridylation is catalyzed solely by the Pus enzymes.

Different Pus enzymes are localized to three sub-cellular domains: the nucleus, the cytosol and the mitochondria [27]. While some Pus enzymes reside in only one of these three sub-cellular domains, others are found in two or all of them. For example, in *S. cerevisiae*, Pus1 is localized almost exclusively to the nucleus, Pus 8 to the cytosol, and Pus2 and Pus5 to the mitochondria. On the other hand, Pus3 and Pus7 are found in both the nucleus and the cytosol, Pus4 in both the nucleus and the mitochondria, and Pus6 in both the cytosol and the

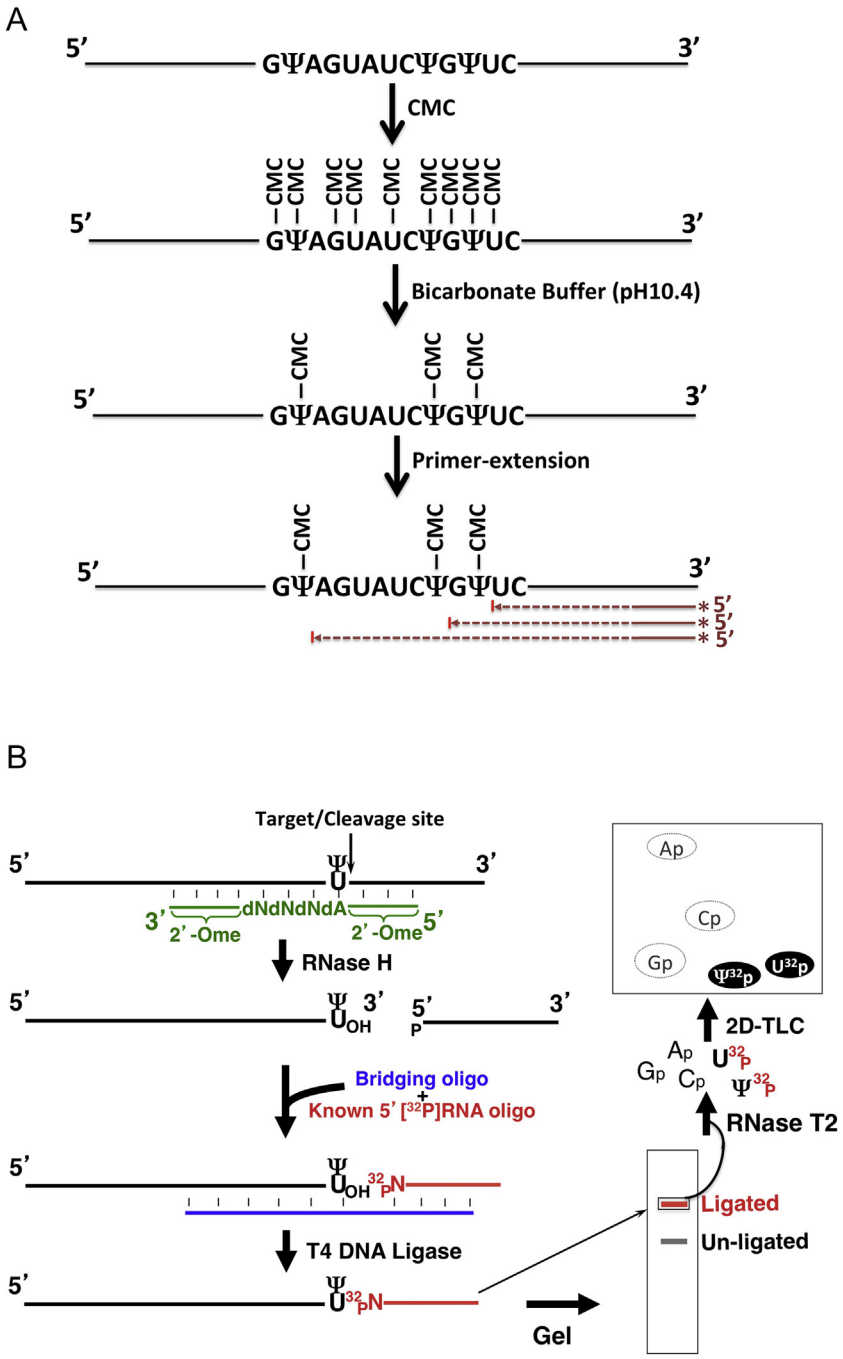

Fig. 2. Pseudouridylation assays. (A) Assay based on CMC-modification followed by primer-extension. The carbodiimide CMC is able to covalently attach to G, U and Ψ. Upon alkaline treatment, G-CMC and U-CMC adducts are reversed, but CMC remains covalently bound to Ψ (Ψ-CMC). Subsequent primer-extension results in stops one nucleotide before the Ψ-CMC adducts. CMC-nucleotide adducts are indicated. Primer (red lines)-extension (dotted red lines) stops are indicated as well. * represents the radio label. (B) Assay based on site-specific cleavage/labeling, nuclease digestion and TLC. An RNA is site-specifically cleaved by RNase H directed by a complementary 2′-*O*-methyl RNA-DNA chimera (here, cleavage is directed at the site 3′ of the U/Ψ site). The 5′ cleaved fragment is ligated to a known, 5′ labeled RNA oligonucleotide. After gel purification, the ligated, radiolabeled RNA is digested with T2. The digestion products $\Psi^{32}p$ and $U^{32}p$ are separated on TLC. The chimeric oligonucleotide (4,2′-deoxy nucleotides flanked by 2′-*O*-methylated nucleotides) is depicted (green). The cleavage site is indicated by a small arrow. The bridging deoxyoligonucleotide (blue) and known 5′ labeled RNA oligonucleotide are also shown. On TLC, the $\Psi^{32}p$ and $U^{32}p$ spots are radiolabeled and visible; the Ap, Gp and Cp are not labeled and not visible.

mitochondria. Interestingly, Pus9 is the only one that resides in all these three sub-cellular domains.

3.2. RNA dependent mechanism

RNA-dependent pseudouridylation is catalyzed by a family of box H/ACA RNPs [40–42]. Each RNP consists of one unique box H/ACA RNA and four common proteins: Cbf5 (called NAP57 in rodents and Dyskerin in human), Nhp2, Nop10 and Gar1. Cbf5 is the catalytic subunit (pseudouridylase) of box H/ACA RNPs, and a member of the TruB pseudouridine synthase family. Despite sequence differences, eukaryotic box H/ACA RNAs all fold into unique variants of a common "hairpin-hinge-hairpin-tail" structure in which each of the two hairpins contains a pseudouridylation pocket that base-pairs with the substrate sequence and positions the target uridine at the base of the upper stem (Fig. 3). Here, the RNA component of each box H/ACA RNP recognizes the substrate and specifies the target uridine, whereupon Cbf5, the catalytic subunit, catalyzes the pseudouridylation reaction, converting the specified uridine into Ψ. Unlike tRNA pseudouridylation, pseudouridylation of rRNAs and spliceosomal snRNAs is catalyzed almost exclusively by the RNA-dependent mechanism in eukaryotes. In yeast and mammals, almost all of the box H/ACA RNPs responsible for pseudouridylation of snRNAs and rRNAs have been identified [43–48].

Using NMR, the details of guide-substrate base-pairing interactions in box H/ACA RNPs have been clarified [49,50]. According to NMR solution structures, the interactions between the guide sequence and substrate are not regular Watson-Crick A-form helices. Instead, when pairing, the substrate interacts on one face of the guide sequence without threading through the guide pocket to form an A-form helix [49,50]. Soon after the NMR studies were completed, the crystal

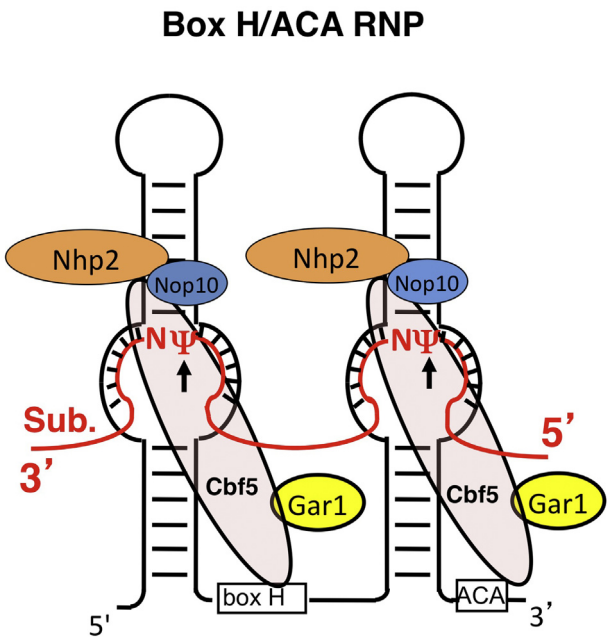

Fig. 3. Box H/ACA guided RNA pseudouridylation. The components of a box H/ACA RNP, including one box H/ACA RNA (black line with box H and box ACA) and four core proteins (Nhp2, Nop10, Gar1 and Cbf5), are shown. The red line represents the substrate RNA. N represents any nucleotide; Ψ (indicated by an arrow) is the modified nucleotide.

structures of several versions of archaeal and yeast box H/ACA RNPs (with and without substrate) became available, depicting, at the molecular level, how box H/ACA RNA-guided pseudouridylation occurs [51–59].

Box H/ACA RNPs are localized to the nuclear sub-compartments, including the nucleolus and the Cajal bodies. Specifically, while the snRNA-specific box H/ACA RNPs are targeted to the Cajal bodies, the rRNA-specific box H/ACA RNPs are concentrated in the nucleoli [60–65]. Thus, it is believed that pseudouridylation of snRNAs and rRNAs occurs in the Cajal bodies and the nucleoli, respectively [65].

4. Noncoding RNA pseudouridylation

Ψ was initially identified in stable and highly abundant noncoding RNAs, such as rRNAs, spliceosomal snRNAs and tRNAs. In eukaryotes, rRNAs and snRNAs are the most extensively pseudouridylated of all cellular RNAs. For instance, there are about 50 Ψs in *S. cerevisiae* rRNAs, and double that number in mammalian rRNAs [46,47,66]. Likewise, there are 6 Ψs in *S. cerevisiae* snRNAs, and more than 20 in mammalian snRNAs [26,41,67]. In this subsection, we will provide a brief overview of noncoding RNA pseudouridylation, focusing on rRNAs and snRNAs.

4.1. rRNA pseudouridylation

It has long been known that rRNAs contain a large number of Ψs [66]; however, the function of these Ψs remained unknown until 1996, when a large number of box H/ACA RNAs were identified and the mechanism of box H/ACA RNA-guided rRNA pseudouridylation was uncovered [44,68,69]. It was soon realized that Ψs are often clustered in functionally important regions of rRNAs, such as the peptidyl transferase center (PTC), the decoding center, the A-site finger (ASF) region, and helix 69, which interacts with tRNAs at the A and P sites during translation [70]. At the 3-dimensional level, Ψs are also present at sites where ribosomal subunits interact [70]. These observations suggest strongly that Ψs are functionally important. Indeed, subsequent experiments carried out in *S. cerevisiae* demonstrated that ribosomal Ψs are important for ribosome biogenesis and protein translation, and that they act in a synergistic manner [71–75]. Further genetic and biochemical analyses have indicated that Ψs also play an important role in translational re-coding [76,77].

4.2. Spliceosomal snRNA pseudouridylation

Spliceosomal snRNAs are also enriched with Ψs. Vertebrate U2 snRNA is the most extensively modified example. Indeed, there are 13 constitutively modified Ψs in vertebrate U2—over 20% of U2 uridines are converted to Ψs [67]. Importantly, these Ψs are also clustered in functionally important regions [67]. Biochemical and genetic experiments conducted in *Xenopus* oocytes, Hela nuclear extracts, and *S. cerevisiae* cells have demonstrated the importance of Ψs in U2 snRNP biogenesis, spliceosome assembly and splicing [78–84]. Most recently, Wu et al. further dissected, using yeast genetics and biochemistry, the role of U2 Ψs in spliceosome assembly and splicing [9]. Their results revealed genetic cross-talk between the Ψs located in the U2 branch site recognition region and Prp5, an ATPase required for pre-splicing complex formation. Further biochemical analysis demonstrated that the Ψs in the branch site recognition region contribute to the physical interaction of U2 with Prp5, impacting Prp5's U2-dependent ATPase activity and hence affecting spliceosome assembly.

Interestingly, it was reported that pseudouridylation of yeast U2 snRNA can also be induced at novel sites under stress conditions (nutrient deprivation, heat-shock) and that these inducible Ψs appear to influence pre-mRNA splicing as well [25]. Nutrient deprivation-induced pseudouridylation has been shown to be regulated by the TOR (Target of Rapamycin) signaling pathway [85]. Both RNA-dependent and RNA-independent mechanisms are involved in inducible pseudouridylation [25].

4.3. Other noncoding RNA pseudouridylation

Pseudouridylation occurs not only in the most abundant noncoding RNAs, including rRNA, snRNA and tRNA, but also in other noncoding RNAs as well. The recent development and application of high-throughput pseudouridine-seq technologies has led to the identification of new Ψs in many more noncoding RNAs, and the list is expanding rapidly. In addition to rRNA, snRNA, and tRNA, a number of relatively low-abundance noncoding RNAs, including box H/ACA RNA, box C/D RNA, telomerase RNA, 7SK RNA, RNase MRP RNA, and Steroid Receptor RNA Activator (SRA) (a noncoding RNA), also contain Ψ [10–14]. It has been shown that tRNA pseudouridylation contributes to tRNA stability, decoding fidelity and recoding efficiency during translation [86–88]. It has also been reported that SRA pseudouridylation and 7SK pseudouridylation appear to play a role in transcriptional regulation [10,89–91]. However, the function of Ψs in most other noncoding RNAs remains unknown.

5. mRNA pseudouridylation

In the early phase of pseudouridylation research, noncoding RNA was the only focus of investigation. However, as ever-increasing amounts of RNA-guided pseudouridylation sites were identified in snRNAs and rRNAs, a suspicion began to arise that pseudouridylation might also occur in coding RNAs. This culminated in a shift in experimental attention towards mRNA pseudouridylation beginning around 2011. Now, after 7 years of research, these efforts have begun to bear fruit.

5.1. Targeted pseudouridylation promotes stop codon read-through

Because mRNA carries the genetic information that directs the synthesis of protein, one interesting question was whether mRNA

pseudouridylation would change coding specificity. As all three stop codons (UAA, UGA, and UAG) contain a uridine at the first position, Karijolich and Yu first tested the coding specificity of pseudouridylated stop codons [92]. To do this, they synthesized a short mRNA molecule containing, from 5′ to 3′, a 6XHIS tag, a stop codon (UAA), and a FLAG tag. By changing the uridine of the stop codon to C or Ψ, they created two additional mRNA molecules. These mRNAs were then translated into protein using cell lysate, and peptide formation was analyzed by dot blot using anti-HIS and anti-FLAG antibodies. Interestingly, while the HIS level was essentially the same regardless of which mRNA was used, FLAG levels varied depending on the experimental codon. Specifically, when mRNA with a normal stop codon UAA was used, no FLAG was detected (as expected); however, when the other two mRNAs were used, FLAG was detected, demonstrating that U-to-Ψ change in the stop codon, just like stop codon mutation (U-to-C change), promotes nonsense suppression. Taking advantage of box H/ACA RNA-guided RNA pseudouridylation, Karijolich and Yu then directly tested stop codon pseudouridylation in *S. cerevisiae*. They created an artificial box H/ACA RNA targeting a premature stop codon (either UAA, UAG or UGA) within a reporter gene. Upon co-transformation of yeast cells with both the reporter and box H/ACA RNA genes, they detected site-specific pseudouridylation, albeit at a low efficiency (7–10%), at the target stop codon. Importantly, the low level of U-to-Ψ conversion was sufficient for them to detect stop codon read-through. These results suggested that pseudouridylation of stop codons does indeed promote stop codon read-through both *in vitro* and *in vivo* (yeast cells) [92] (Fig. 4). Recently, Adachi et al. have demonstrated that targeted pseudouridylation of a premature termination codon (stop codon) also results in the suppression of nonsense mediated mRNA decay (NMD) (Fig. 4, and unpublished data).

To identify the amino acids incorporated at the pseudouridylated stop codons, Karijolich and Yu purified the read-through proteins and sequenced them using LC-MS/MS [92]. The results indicate that ΨAA and ΨAG both primarily code for Ser and Thr, while ΨGA codes for Tyr and Phe. Inspection of codons and anticodons (of tRNAs) led to the realization that it is virtually impossible to avoid purine-purine interactions during decoding. To directly visualize these interactions, Fernández et al. crystalized the ribosome bound with pseudouridylated stop codon ΨAG and its "cognate" anticodon of $tRNA^{Ser}$(IGA), and showed that unusual base pairing occurs between two purines during decoding, suggesting that the catalytic center of the ribosome is rather flexible [93].

Box H/ACA RNA-guided pseudouridylation does not appear to be limited to only mRNA in yeast. Using the same technique, Chen et al. showed that even a pre-mRNA can be site-specifically targeted for pseudouridylation in *Xenopus* oocytes [94]. These exciting results prompted a search for potential naturally occurring pseudouridines in

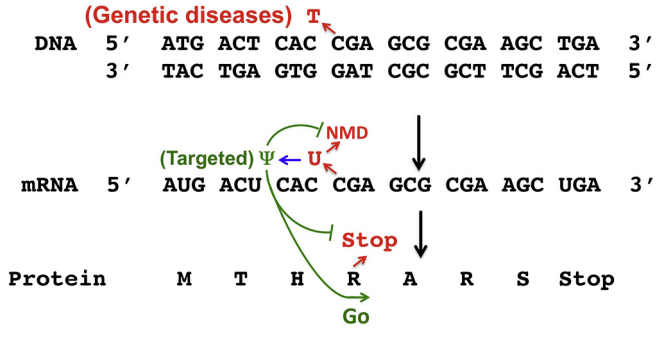

Fig. 4. Nonsense suppression induced by targeted pseudouridylation. Gene expression (from DNA to RNA to protein) is schematized. Black letters represent wild-type sequences; red letters and red arrows indicate mutations (present in disease genes) and the consequences of mutations, respectively; the blue arrow indicates U-to-Ψ conversion; the green letter and green lines/arrow depict the isomerized nucleotide (Ψ) and the consequences, respectively.

mRNAs. Using the guide sequences of known yeast box H/ACA RNAs, Karijolich and Yu conducted a BLAST search against the yeast genome, and identified a number of potential hits in coding genes. Thus, they predicted the existence of naturally occurring Ψs in mRNA [92].

5.2. Identification of naturally occurring Ψs in mRNA by high throughput pseudouridine-seq technologies

Although Ψ was predicted or believed to be widespread and present in mRNA, experimental evidence for this was still lacking. To address this problem, four groups independently developed four similar but distinct high throughput pseudouridine-seq methods (Pseudo-seq, Ψ-seq, PSI-seq and CeU-seq), all of which employed similar strategies that coupled CMC-modification and primer-extension (Fig. 2A) with deep sequencing [12–15]. These methods permitted a transcriptome-wide survey of pseudouridylation. Specifically, pseudo-seq and Ψ-seq were used to survey pseudouridylation in both yeast and mammalian cells; PSI-seq was used to identify Ψs only in yeast RNAs; and CeU-seq was used to map Ψs only in mammalian cells.

Together, the 4 studies identified a total of 1496 Ψs (including stress-inducible Ψs) in 867 yeast mRNAs, and a total of 3083 Ψs (including stress-inducible Ψs) in 2174 human mRNAs (5′ and 3′ UTRs and the coding region). However, Zaringhalam and Papavasiliou have, using comparison analysis, noted that only a small fraction of these Ψs and mRNAs were identified consistently by two or more of these methods [95]. Indeed, only 48 putative Ψs across 46 human mRNAs were independently identified by two or more of these methods (Table 1). Likewise, only 38 Ψs in 38 yeast mRNAs were consistently identified by at least two of these methods (Table 2). Although a large number of putative Ψs have been identified in mRNAs, only one of them has been independently verified by the TLC-based method (Fig. 2B) [14].

Why do the various methods of pseudouridine-seq generate somewhat inconsistent results? It should be noted that across the mammalian mRNAs analyzed by Pseudo-seq, Ψ-seq, and CeU-seq, some discrepancy is not totally surprising given the variations in cell types and culture conditions used in each lab. Despite this, the low overlaps suggest that the four novel high-throughput pseudouridine-seq methods are far from perfect. It has long been known that CMC-modification followed by primer-extension (Fig. 2A) has intrinsic problems. For example, as discussed earlier, each of these processes involves multiple steps, and small differences can accumulate at each one, potentially leading to large total variances. In addition, Ψ-specific CMC-modification may not be 100% efficient, leading to a low level of primer-extension stops and therefore impacting final detection of Ψs. Furthermore, CMC also modifies U and G (in addition to Ψ), and the reversal of U-CMC and G-CMC to unmodified U and G may not be complete in the bicarbonate buffer (Fig. 2A), resulting in false positives. However, it should be noted that the CMC-modification-primer-extension-based assay has been tested extensively in detecting Ψ sites in rRNA and snRNA, where pseudouridylation at a given site is more than 90%. The results indicate that the false positives are minimal. Hence the CMC-modification-primer-extension-based method itself may not be the main reason for the low overlap rate. On the other hand, studies have suggested that bioinformatic pipelines have significant effects on sites calling (especially for the sites with a low level of pseudouridylation). It is therefore possible that some potential limitations of the method, together with potential differences in parameter settings during bioinformatic data analysis, might have resulted in the low number of mutually identified Ψs. It is thus urgently important to use the same computational standard for the mapping of pseudouridylation sites.

Despite the discrepancy in their results, the four surveys have successfully indicated that Ψ is present not only in noncoding RNAs, but also in a number of mRNAs. In the latter case, Ψ appears to be distributed throughout the entire sequence, including the 5′ UTR, 3′ UTR and the coding region, although Ψ seems to be underrepresented in the

Table 1
Identical Ψ sites detected in human mRNAs by at least two of the three methods (Pseudo-seq, Ψ-seq, or CeU-seq).

Methods	Coordinate	Gene	Position	Amino acid	Codon	Dependency (shown in Ψ-seq)
Ψ-seq, CeU-seq (Ψ-seq, position 550 was also detected)	chr1:33480153	AK2	551	F156	UΨC	Pus4
Ψ-seq, CeU-seq	chr11:94607530	AMOTL1	6611	3′UTR	3′UTR	Pus4
Ψ-seq, CeU-seq	chr19:19765470	ATP13A1	1723	V565	GΨA	Pus7
Ψ-seq, CeU-seq[a]	chr20:17581686	DSTN	562[a]	F102	UUΨ	unknown
Ψ-seq, CeU-seq	chr10:120832492	EIF3A	597	V150	GUΨ	unknown
Ψ-seq, CeU-seq (Ψ-seq, position 654 was also detected)	chr14:69846986	ERH	655	3′UTR	3′UTR	Pus4
Ψ-seq, CeU-seq	chr8:119122857	EXT1	1202	F143	UΨC	Pus4
Ψ-seq, CeU-seq	chr2:131807076	FAM168B	3779	3′UTR	3′UTR	unknown
Ψ-seq, CeU-seq	chr6:153296738	FBXO5	231	S41	ΨCU	Pus4
Ψ-seq, CeU-seq	chr12:2913076	FKBP4	2230	3′UTR	3′UTR	Pus4
Ψ-seq, CeU-seq	chr15:41059798	GCHFR	600	3′UTR	3′UTR	Pus7
Ψ-seq, CeU-seq	chr15:72636387	HEXA	1828	3′UTR	3′UTR	Pus7
Ψ-seq, CeU-seq	chr1:6282924	ICMT	3144	3′UTR	3′UTR	unknown
Ψ-seq, CeU-seq	chr10:1090038	IDI1	332	V71	GUΨ	Pus4
Ψ-seq, CeU-seq	chr7:1516290	INTS1	5168	F1689	UΨC	Pus4
Ψ-seq, CeU-seq	chr16:47345317	ITFG1	963	V301	GUΨ	Pus4
Ψ-seq, CeU-seq	chr2:20237324	LAPTM4A	792	S95	ΨCA	Pus4
Ψ-seq, CeU-seq	chr19:34685385	LSM14A	320	V41	GUΨ	Pus4
Ψ-seq, CeU-seq	chr22:35820401	MCM5	2452	3′UTR	3′UTR	Pus7
Ψ-seq, CeU-seq	chr17:37564158	MED1	4528	S1439	ΨCC	Pus4
Ψ-seq, CeU-seq	chr10:64960744	MSN	2921	3′UTR	3′UTR	unknown
Ψ-seq, CeU-seq	chr9:127074876	NEK6	542	Y94	ΨAC	unknown
Ψ-seq, CeU-seq	chr16:69376048	NIP7	1354	3′UTR	3′UTR	Pus4
Ψ-seq, CeU-seq (Ψ-seq, position 454 were also detected)	chr4:48850446	OCIAD1	455	S80	ΨCC	Pus4
Ψ-seq, CeU-seq	chr10:128726344	OCRL	4968	3′UTR	3′UTR	Pus4
Ψ-seq, CeU-seq	chr1:40030820	PABPC4	2101	F401	UΨC	Pus4
Ψ-seq, CeU-seq	chr4:129191455	PGRMC2	2727	3′UTR	3′UTR	unknown
Ψ-seq, CeU-seq	chr1:6683668	PHF13	3245	3′UTR	3′UTR	Pus4
Ψ-seq, CeU-seq	chr5:108714952	PJA2	456	S79	ΨCC	Pus4
Ψ-seq, CeU-seq (Ψ-seq, position 1277 was also detected)	chr7:102952258	PMPCB	1278	S415	ΨCA	Pus4
Ψ-seq, CeU-seq	chr1:166823415	POGK	3558	3′UTR	3′UTR	Pus4
Ψ-seq, CeU-seq	chr12:53837226	PRR13	235	S24	ΨCC	Pus4
Ψ-seq, CeU-seq	chr10:73576420	PSAP	2457	3′UTR	3′UTR	unknown
Ψ-seq, CeU-seq	chr1:36068933	PSMB2	953	V63	GUΨ	Pus4
Ψ-seq, CeU-seq	chr7:75518106	RHBDD2	1618	3′UTR	3′UTR	Pus7
Ψ-seq, CeU-seq	chr17:4848196	RNF167	1387	S313	ΨCU	Pus4
Ψ-seq, CeU-seq	chr3:23959343	RPL15	116, 373, and 386	5′UTR	5′UTR	DKC1
Ψ-seq, CeU-seq	chr1:53493672	SCP2	1435	V18	GUΨ	Pus4
Ψ-seq, CeU-seq	chr17:2278920	SGSM2	2412	F700	UΨC	Pus4
Ψ-seq, CeU-seq	chr16:28883911	SH2B1	2054	F594	UΨC	Pus4
Ψ-seq, CeU-seq (Ψ-seq, position 2318 were also detected)	chr1:43392397	SLC2A1	2319	3′UTR	3′UTR	DKC1
Ψ-seq, CeU-seq	chr5:68419196	SLC30A5	2249	V647	GUΨ	Pus4
Ψ-seq, CeU-seq	chr10:77385465	TAF9B	2469	3′UTR	3′UTR	Pus4
Ψ-seq, CeU-seq	chr8:141297807	TRAPPC9	2190	F627	UΨC	Pus4
Ψ-seq, CeU-seq	chr7:100470818	TRIP6	1494	S441	UCΨ	unknown
Ψ-seq, CeU-seq	chr12:124497242	ZNF664	1528	S184	ΨCG	Pus4

[a] While Ψ-seq detected a Ψ at position 562, CeU-seq identified the nucleotide at position 561 as Ψ.

5′ UTRs of mammalian mRNAs and in the 3′ UTRs of *S. cerevisiae* mRNAs [12,14,95]. By comparing Ψ maps generated before and after knocking out or knocking down Pus enzymes or box H/ACA RNAs, the researchers were also able to show that some pseudouridylation enzymes (both Pus enzymes and box H/ACA RNPs) are responsible for mRNA pseudouridylation at specific sites. Interestingly, upon exposing cells to stress (nutrient deprivation, heat shock, oxidation, etc.), pseudouridylation is induced in mRNA [12–15], a phenomenon first observed in yeast U2 snRNA [25]. It appears that, like in inducible U2 snRNA pseudouridylation [25], inducible mRNA pseudouridylation can be catalyzed by both stand-alone Pus enzymes and box H/ACA RNPs.

Most recently, Nakamoto et al. have used a similar PSI-seq to conduct a transcriptome-wide survey of Ψ formation in the parasite Toxoplasma gondii, and generated a great deal of information about TgPUS1-dependent and TgPUS1-independent pseudouridylation [21]. Of note, the authors previously identified TgPUS1 as a Pus enzyme that is necessary for differentiation of the parasite from active to chronic infection [96]. In particular, they identified TgPUS1-dependent and TgPUS1-independent Ψs in several different types of RNAs, including mRNA. Like in yeast and mammals, Ψs were found in all regions of Toxoplasma mRNA. However, their data suggest that Ψs are underrepresented in the 3′-UTRs, just as they are in *S. cerevisiae* mRNAs.

5.3. Possible functions of mRNA pseudouridylation

Even though there are relatively few commonly identified Ψs across the 4 pseudouridine-seq methods, the evidence for natural mRNA pseudouridylation (either constitutive or inducible) is overwhelming and undeniable. Yet, to date, little is known about the biological function of mRNA pseudouridylation. Why is Ψ distributed in the 5′ UTR, the coding region and the 3′ UTR? What potential role(s) do these Ψs play during mRNA processing and gene expression?

Given the known mRNA processing pathway, it is tempting to speculate on the potential roles of Ψ (Fig. 5). For instance, it is possible that through altering mRNA structure and recruiting regulatory protein factors, Ψ could contribute to pre-mRNA splicing, 3′ end formation/polyadenylation, and even 5′ capping. However, the latter seems less likely given that capping occurs early and co-transcriptionally, perhaps even before Ψ is introduced into pre-mRNA/mRNA. On the other hand, Ψs could constitute a localization signal, regulating RNA sub-cellular localization. This is particularly relevant when taking into account that mRNA pseudouridylation is dynamic and inducible in response to

Table 2
Identical Ψ sites detected in yeast mRNAs by at least two of the three methods (Pseudo-seq, Ψ-seq, or PSI-seq[a]).

Methods (HS = heat shock induced)	Coordinate	Gene	Position	Amino acid	Codon
Pseudo-seq, Ψ-seq (HS)	chrXII:200345	ADE16	971	V324	GΨA
Pseudo-seq, Ψ-seq	chrIV:331025	BDF2	2	M1	AΨG
Pseudo-seq, PSI-seq (Log and HS)	chrXV:60739	CDC33	286	F96	ΨUC
Pseudo-seq, Ψ-seq (HS)	chrXV:296080	CMK2	41	V14	GΨA
Pseudo-seq, Ψ-seq	chrI:32596	GDH3	1030	S344	YCT
Pseudo-seq, Ψ-seq (HS)	chrIV:766454	GIR2	749	V250	GΨA
Pseudo-seq, Ψ-seq	chrIII:51028	GLK1	191	M64	AΨG
Pseudo-seq, Ψ-seq (HS)	chrXII:661272	GSY2	557	V186	GΨA
Pseudo-seq, Ψ-seq	chrIX:52395	GUT2	1314	T438	ACΨ
Pseudo-seq, PSI-seq (HS)	chrXIII:632426	HSC82	72	Y24	UAΨ
Pseudo-seq, Ψ-seq (HS)	chrIII:156194	HSP30	914	V305	GΨU
Pseudo-seq, Ψ-seq	chrXVI:568435	ICL2	562	S188	ΨCU
Pseudo-seq, Ψ-seq (Mid-log and HS)	chrX:383242	KAR2	1916	F639	UΨC
Pseudo-seq, Ψ-seq	chrX:314164	MPM1	709	S237	ΨCC
Pseudo-seq, Ψ-seq (HS)	chrIII:49581	PDI1	641	V214	GΨA
Pseudo-seq, Ψ-seq, PSI-seq	chrXVI:731681	RPL11A	68	V23	GΨU
Pseudo-seq, PSI-seq (HS)	chrXV:93771	RPL18A	185	V62	GΨU
Ψ-seq (HS), PSI-seq (HS)	chrXIV:331941	RPL42A	106	F36	ΨUU
Pseudo-seq, Ψ-seq (HS)	chrVIII:382275	RPL42B	36	C12	UGΨ
Pseudo-seq, PSI-seq (Log and HS)	chrII:300266	RPL4A	101	I34	AΨU
Pseudo-seq, Ψ-seq	chrVIII:499441	RPN10	363	S121	UCΨ
Pseudo-seq, Ψ-seq (HS)	chrVII:921633	RPS0A	604	Y202	ΨAC
Pseudo-seq, PSI-seq (HS)	chrXII:673244	RPS28B	114	R38	CGΨ
Pseudo-seq, Ψ-seq (HS)	chrIV:930301	RTN1	57	C19	UGΨ
Pseudo-seq, Ψ-seq (HS)	chrVII:292922	SNF4	890	V297	GΨA
Pseudo-seq, Ψ-seq (Mid-log and HS)	chrV:468452	SPI1	83	V28	GΨA
Pseudo-seq, Ψ-seq (HS)	chrXII:836860	TAL1	498	C166	UGΨ
Pseudo-seq, PSI-seq (Log and HS)	chrXVI:700832	TEF1	239	F80	UΨC
Pseudo-seq, Ψ-seq	chrII:477909	TEF2	239	F80	UΨC
Pseudo-seq, Ψ-seq (HS)	chrIV:1112118	TIM11	176	V59	GΨA
Pseudo-seq, PSI-seq (HS)	chrIII:111180	YCL002C	420	C140	UGΨ
Ψ-seq, PSI-seq	chrXII:637854	YEF3	1074	F358	UUΨ
Pseudo-seq, Ψ-seq	chrV:173300	YER010C	39	C13	UGΨ
Pseudo-seq, Ψ-seq	chrVII:623051	YGR067C	1736	F579	UΨC
Pseudo-seq, Ψ-seq (HS)	chrVII:960865	YHB1	962	V321	GΨA
Pseudo-seq, Ψ-seq (HS)	chrVIII:380666	YHR140W	95	V32	GΨA
Pseudo-seq, PSI-seq	chrXIII:449299	YMR090W	55	L19	ΨUG
Pseudo-seq, Ψ-seq	chrXVI:126070	YPL225W	65	V22	GΨU

[a] The position numbers generated by PSI-seq are often not consistent with those presented by Pseudo-seq and Ψ-seq. When the differences are within ± 2 bases, the PSI-seq numbers are also listed (considered "identical").

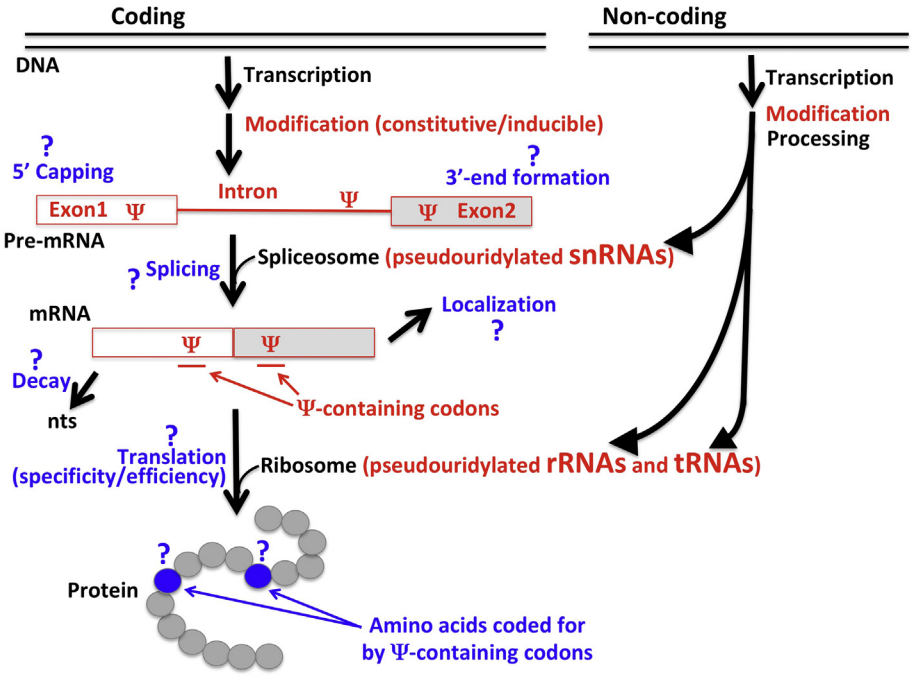

Fig. 5. Ψs in gene expression. Pseudouridylated coding RNAs (pre-mRNA and mRNA) and non-coding RNAs (snRNA, tRNA and rRNA) are shown in red. Ψs of snRNA contribute to pre-mRNA splicing; Ψs of tRNA aid in RNA stability and translation fidelity; Ψs of rRNA are involved in ribosome biogenesis and translation. Although pseudouridylation of stop codons promotes stop codon read-through and suppresses NMD, the function of Ψs in other sites/regions of mRNA is not clear. The blue question marks and keywords indicate speculative functions.

various stress conditions [12–15].

Ψ could also play a role in mRNA stability. Indeed, there are several lines of experimental evidence that support this notion. For instance, using Ψ-seq, Schwartz et al. measured the levels of 142 mRNAs in wild-type and pus7-deletion *S. cerevisiae* strains. These mRNAs, which normally harbor Pus7-dependent Ψ, were present in higher levels in the wild-type strain (where they were Ψ normally) than they were in the pus7-deletion strain (in which no Ψ was present), suggesting that mRNA pseudouridylation contributes to mRNA stabilization [13]. This is consistent with a previous report by Kariko et al. indicating that when transfecting synthetic mRNA into mammalian cells, pseudouridylated versions displayed increased stability compared to mRNA containing no Ψ [97]. However, the recent work of Nakamoto et al. (discussed above) found that mRNAs containing TgPUS1-dependent Ψ are modestly (10%) more stable in TgPUS1-mutant parasite (where they lack Ψ) than they are in wild-type parasite (where Ψ is present) [21]. This result suggests that pseudouridylation helps to destabilize mRNA. Taken together, it appears that Ψ can act in two opposite directions, enhancing mRNA stability in some instances and reducing mRNA stability in others, depending on the organisms, genes, and conditions used.

Ψs in the coding region might also affect coding specificity during translation. This is supported by the previous study by Karijolich and Yu demonstrating that targeted pseudouridylation can convert stop codons into sense codons coding for specific amino acids (discussed above) [92]. However, whether pseudouridylation of U-containing sense codons can alter coding specificity remains an open question. The recent study by Nakamoto et al. has shown that Ψ is disproportionally present at the first position of modified sense codons [21]. Given the importance of the first position of a codon in decoding, it is tempting to speculate that Ψs might be able to either change the coding specificity of sense codons or affect translation efficiency. Likewise, Ψs in the untranslated regions, especially in the 3′UTR, could also play a role in translation regulation.

6. Concluding remarks

Pseudouridylation is a post-transcriptional RNA modification that has been known for decades. Over the years, most research into pseudouridylation has been confined to the major noncoding RNAs, including rRNAs and snRNAs. A large amount of data concerning pseudouridylation of these RNAs has been accumulated. We now know that Ψs within rRNAs play an important role in ribosome biogenesis and protein translation. Likewise, Ψs within the spliceosomal snRNAs contribute to pre-mRNA splicing. Two independent pseudouridylation mechanisms, one RNA-independent and the other RNA-dependent, have also been identified. In the meantime, several pseudouridylation assays have been developed, making it easier to conduct RNA pseudouridylation research. Although progress has been remarkable, there are still a number of questions that remain to be answered. Most notably, understanding the molecular details of the function of Ψs, in other words, how they contribute to the function of the RNAs in which they reside, remains a challenge. However, given the recent growing interest in RNA modification, it is expected that issues regarding the function of noncoding RNA pseudouridylation will be clarified.

The recent development of pseudouridine-seq techniques has led to the identification of a number of new Ψs in various types of RNA, including mRNA. However, as discussed earlier, the new methods appear to be far from perfect. This offers a great opportunity to develop new, better methods, or to further perfect existing technologies. The identification of new Ψs in mRNA has opened the door to further experiments directed towards understanding the function of mRNA pseudouridylation. Why are Ψs present in different regions of mRNA, and what role might mRNA pseudouridylation play? Although this is a new territory in pseudouridylation research, given the immense success of investigations into pseudouridylation of noncoding RNA, we are confident that a picture of the function of mRNA pseudouridylation will soon emerge.

Transparency document

The Transparency document associated with this article can be found, in online version.

Acknowledgments

We thank the members of the Yu laboratory for valuable discussions. The work carried out in the Yu laboratory was supported by grants GM104077 and AG39559 (to Y.-T.Y.) from the National Institutes of Health.

References

[1] F.F. Davis, F.W. Allen, Ribonucleic acids from yeast which contain a fifth nucleotide, J. Biol. Chem. 227 (1957) 907–915.
[2] W.E. Cohn, 5-Ribosyl uracil, a carbon-carbon ribofuranosyl nucleoside in ribonucleic acids, Biochim. Biophys. Acta 32 (1959) 569–571.
[3] M. Charette, M.W. Gray, Pseudouridine in RNA: what, where, how, and why, IUBMB Life 49 (2000) 341–351.
[4] P.F. Agris, The importance of being modified: roles of modified nucleosides and Mg^{2+} in RNA structure and function, Prog. Nucleic Acid Res. Mol. Biol. 53 (1996) 79–129.
[5] D.R. Davis, Stabilization of RNA stacking by pseudouridine, Nucleic Acids Res. 23 (1995) 5020–5026.
[6] J.G. Arnez, T.A. Steitz, Crystal structure of unmodified tRNA(Gln) complexed with glutaminyl-tRNA synthetase and ATP suggests a possible role for pseudo-uridines in stabilization of RNA structure, Biochemistry 33 (1994) 7560–7567.
[7] M.I. Newby, N.L. Greenbaum, Sculpting of the spliceosomal branch site recognition motif by a conserved pseudouridine, Nat. Struct. Biol. 9 (2002) 958–965.
[8] M.I. Newby, N.L. Greenbaum, Investigation of Overhauser effects between pseudouridine and water protons in RNA helices, Proc. Natl. Acad. Sci. U. S. A. 99 (2002) 12697–12702.
[9] G. Wu, H. Adachi, J. Ge, D. Stephenson, C.C. Query, Y.T. Yu, Pseudouridines in U2 snRNA stimulate the ATPase activity of Prp5 during spliceosome assembly, EMBO J. 35 (2016) 654–667.
[10] Y. Zhao, W. Dunker, Y.T. Yu, J. Karijolich, The role of noncoding RNA pseudouridylation in nuclear gene expression events, Front. Bioeng. Biotechnol. 6 (2018) 8.
[11] J. Karijolich, C. Yi, Y.T. Yu, Transcriptome-wide dynamics of RNA pseudouridylation, Nat. Rev. Mol. Cell Biol. 16 (2015) 581–585.
[12] T.M. Carlile, M.F. Rojas-Duran, B. Zinshteyn, H. Shin, K.M. Bartoli, W.V. Gilbert, Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells, Nature 515 (2014) 143–146.
[13] S. Schwartz, D.A. Bernstein, M.R. Mumbach, M. Jovanovic, R.H. Herbst, B.X. Leon-Ricardo, J.M. Engreitz, M. Guttman, R. Satija, E.S. Lander, G. Fink, A. Regev, Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA, Cell 159 (2014) 148–162.
[14] X. Li, P. Zhu, S. Ma, J. Song, J. Bai, F. Sun, C. Yi, Chemical pulldown reveals dynamic pseudouridylation of the mammalian transcriptome, Nat. Chem. Biol. 11 (2015) 592–597.
[15] A.F. Lovejoy, D.P. Riordan, P.O. Brown, Transcriptome-wide mapping of pseudouridines: pseudouridine synthases modify specific mRNAs in *S. cerevisiae*, PLoS One 9 (2014) e110799.
[16] R.C. Gupta, B.A. Roe, K. Randerath, The nucleotide sequence of human tRNAGly (anticodon GCC), Nucleic Acids Res. 7 (1979) 959–970.
[17] R.W. Holley, G.A. Everett, J.T. Madison, A. Zamir, Nucleotide sequences in the yeast alanine transfer ribonucleic acid, J. Biol. Chem. 240 (1965) 2122–2128.
[18] A. Bakin, J. Ofengand, Four newly located pseudouridylate residues in *Escherichia coli* 23S ribosomal RNA are all at the peptidyltransferase center: analysis by the application of a new sequencing technique, Biochemistry 32 (1993) 9754–9762.
[19] X. Zhao, Y.T. Yu, Detection and quantitation of RNA base modifications, RNA 10 (2004) 996–1002.
[20] N. Liu, M. Parisien, Q. Dai, G. Zheng, C. He, T. Pan, Probing N6-methyladenosine RNA modification status at single nucleotide resolution in mRNA and long noncoding RNA, RNA 19 (2013) 1848–1856.
[21] M.A. Nakamoto, A.F. Lovejoy, A.M. Cygan, J.C. Boothroyd, mRNA pseudouridylation affects RNA metabolism in the parasite *Toxoplasma gondii*, RNA 23 (2017) 1834–1849.
[22] Y.T. Yu, M.D. Shu, J.A. Steitz, A new method for detecting sites of 2′-O-methylation in RNA molecules, RNA 3 (1997) 324–331.
[23] J. Lapham, Y.T. Yu, M.D. Shu, J.A. Steitz, D.M. Crothers, The position of site-directed cleavage of RNA using RNase H and 2′-O-methyl oligonucleotides is dependent on the enzyme source, RNA 3 (1997) 950–951.
[24] M. Hengesbach, M. Meusburger, F. Lyko, M. Helm, Use of DNAzymes for site-specific analysis of ribonucleotide modifications, RNA 14 (2008) 180–187.
[25] G. Wu, M. Xiao, C. Yang, Y.T. Yu, U2 snRNA is inducibly pseudouridylated at novel sites by Pus7p and snR81 RNP, EMBO J. 30 (2011) 79–89.
[26] M.D. De Zoysa, Y.T. Yu, Posttranscriptional RNA pseudouridylation, Enzymes 41 (2017) 151–167.

[27] A.C. Rintala-Dempsey, U. Kothe, Eukaryotic stand-alone pseudouridine synthases — RNA modifying enzymes and emerging regulators of gene expression? RNA Biol. 14 (2017) 1185–1196.
[28] T. Hamma, A.R. Ferre-D'Amare, Pseudouridine synthases, Chem. Biol. 13 (2006) 1125–1135.
[29] P.G. Foster, L. Huang, D.V. Santi, R.M. Stroud, The structural basis for tRNA recognition and pseudouridine formation by pseudouridine synthase I, Nat. Struct. Biol. 7 (2000) 23–27.
[30] C. Hoang, A.R. Ferre-D'Amare, Cocrystal structure of a tRNA Psi55 pseudouridine synthase: nucleotide flipping by an RNA-modifying enzyme, Cell 107 (2001) 929–939.
[31] C. Hoang, A.R. Ferre-D'Amare, Crystal structure of the highly divergent pseudouridine synthase TruD reveals a circular permutation of a conserved fold, RNA 10 (2004) 1026–1033.
[32] J. Sivaraman, V. Sauve, R. Larocque, E.A. Stura, J.D. Schrag, M. Cygler, A. Matte, Structure of the 16S rRNA pseudouridine synthase RsuA bound to uracil and UMP, Nat. Struct. Biol. 9 (2002) 353–358.
[33] J. Sivaraman, P. Iannuzzi, M. Cygler, A. Matte, Crystal structure of the RluD pseudouridine synthase catalytic module, an enzyme that modifies 23S rRNA and is essential for normal cell growth of *Escherichia coli*, J. Mol. Biol. 335 (2004) 87–101.
[34] M. Del Campo, J. Ofengand, A. Malhotra, Crystal structure of the catalytic domain of RluD, the only rRNA pseudouridine synthase required for normal growth of *Escherichia coli*, RNA 10 (2004) 231–239.
[35] U.B. Ericsson, P. Nordlund, B.M. Hallberg, X-ray structure of tRNA pseudouridine synthase TruD reveals an inserted domain with a novel fold, FEBS Lett. 565 (2004) 59–64.
[36] Y. Kaya, M. Del Campo, J. Ofengand, A. Malhotra, Crystal structure of TruD, a novel pseudouridine synthase with a new protein fold, J. Biol. Chem. 279 (2004) 18107–18110.
[37] K. Mizutani, Y. Machida, S. Unzai, S.Y. Park, J.R. Tame, Crystal structures of the catalytic domains of pseudouridine synthases RluC and RluD from *Escherichia coli*, Biochemistry 43 (2004) 4454–4463.
[38] C. Hoang, J. Chen, C.A. Vizthum, J.M. Kandel, C.S. Hamilton, E.G. Mueller, A.R. Ferre-D'Amare, Crystal structure of pseudouridine synthase RluA: indirect sequence readout through protein-induced RNA structure, Mol. Cell 24 (2006) 535–545.
[39] C.J. McCleverty, M. Hornsby, G. Spraggon, A. Kreusch, Crystal structure of human Pus10, a novel pseudouridine synthase, J. Mol. Biol. 373 (2007) 1243–1254.
[40] P. Ganot, M. Caizergues-Ferrer, T. Kiss, The family of box ACA small nucleolar RNAs is defined by an evolutionarily conserved secondary structure and ubiquitous sequence elements essential for RNA accumulation, Genes Dev. 11 (1997) 941–956.
[41] Y.T. Yu, U.T. Meier, RNA-guided isomerization of uridine to pseudouridine—pseudouridylation, RNA Biol. 11 (2014) 1483–1494.
[42] Y.T. Yu, R.M. Terns, M.P. Terns, Mechanisms and functions of RNA-guided RNA modification, in: H. Grosjean (Ed.), Fine-tuning of RNA Functions by Modification and Editing, Springer-Verlag Press, 2005, pp. 223–262 Place Published.
[43] A.M. Kiss, B.E. Jady, E. Bertrand, T. Kiss, Human box H/ACA pseudouridylation guide RNA machinery, Mol. Cell. Biol. 24 (2004) 5797–5807.
[44] J. Ni, A.L. Tien, M.J. Fournier, Small nucleolar RNAs direct site-specific synthesis of pseudouridine in ribosomal RNA, Cell 89 (1997) 565–573.
[45] G. Badis, M. Fromont-Racine, A. Jacquier, A snoRNA that guides the two most conserved pseudouridine modifications within rRNA confers a growth advantage in yeast, RNA 9 (2003) 771–779.
[46] P. Schattner, W.A. Decatur, C.A. Davis, M. Ares Jr., M.J. Fournier, T.M. Lowe, Genome-wide searching for pseudouridylation guide snoRNAs: analysis of the *Saccharomyces cerevisiae* genome, Nucleic Acids Res. 32 (2004) 4281–4296.
[47] P. Schattner, S. Barberan-Soler, T.M. Lowe, A computational screen for mammalian pseudouridylation guide H/ACA RNAs, RNA 12 (2006) 15–25.
[48] A. Huttenhofer, M. Kiefmann, S. Meier-Ewert, J. O'Brien, H. Lehrach, J.P. Bachellerie, J. Brosius, RNomics: an experimental approach that identifies 201 candidates for novel, small, non-messenger RNAs in mouse, EMBO J. 20 (2001) 2943–2953.
[49] H. Jin, J.P. Loria, P.B. Moore, Solution structure of an rRNA substrate bound to the pseudouridylation pocket of a box H/ACA snoRNA, Mol. Cell 26 (2007) 205–215.
[50] H. Wu, J. Feigon, H/ACA small nucleolar RNA pseudouridylation pockets bind substrate RNA to form three-way junctions that position the target U for modification, Proc. Natl. Acad. Sci. U. S. A. 104 (2007) 6655–6660.
[51] L. Li, K. Ye, Crystal structure of an H/ACA box ribonucleoprotein particle, Nature 443 (2006) 302–307.
[52] B. Liang, S. Xue, R.M. Terns, M.P. Terns, H. Li, Substrate RNA positioning in the archaeal H/ACA ribonucleoprotein complex, Nat. Struct. Mol. Biol. 14 (2007) 1189–1195.
[53] J. Duan, L. Li, J. Lu, W. Wang, K. Ye, Structural mechanism of substrate RNA recruitment in H/ACA RNA-guided pseudouridine synthase, Mol. Cell 34 (2009) 427–439.
[54] R. Rashid, B. Liang, D.L. Baker, O.A. Youssef, Y. He, K. Phipps, R.M. Terns, M.P. Terns, H. Li, Crystal structure of a Cbf5-Nop10-Gar1 complex and implications in RNA-guided pseudouridylation and dyskeratosis congenita, Mol. Cell 21 (2006) 249–260.
[55] T. Hamma, S.L. Reichow, G. Varani, A.R. Ferre-D'Amare, The Cbf5-Nop10 complex is a molecular bracket that organizes box H/ACA RNPs, Nat. Struct. Mol. Biol. 12 (2005) 1101–1107.
[56] X. Manival, C. Charron, J.B. Fourmann, F. Godard, B. Charpentier, C. Branlant, Crystal structure determination and site-directed mutagenesis of the *Pyrococcus abyssi* aCBF5-aNOP10 complex reveal crucial roles of the C-terminal domains of both proteins in H/ACA sRNP activity, Nucleic Acids Res. 34 (2006) 826–839.
[57] K. Ye, H/ACA guide RNAs, proteins and complexes, Curr. Opin. Struct. Biol. 17 (2007) 287–292.
[58] S.L. Reichow, T. Hamma, A.R. Ferre-D'Amare, G. Varani, The structure and function of small nucleolar ribonucleoproteins, Nucleic Acids Res. 35 (2007) 1452–1464.
[59] S. Li, J. Duan, D. Li, B. Yang, M. Dong, K. Ye, Reconstitution and structural analysis of the yeast box H/ACA RNA-guided pseudouridine synthase, Genes Dev. 25 (2011) 2409–2421.
[60] T. Kiss, Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions, Cell 109 (2002) 145–148.
[61] A. Narayanan, A. Lukowiak, B.E. Jady, F. Dragon, T. Kiss, R.M. Terns, M.P. Terns, Nucleolar localization signals of box H/ACA small nucleolar RNAs, EMBO J. 18 (1999) 5120–5130.
[62] P. Richard, X. Darzacq, E. Bertrand, B.E. Jady, C. Verheggen, T. Kiss, A common sequence motif determines the Cajal body-specific localization of box H/ACA scaRNAs, EMBO J. 22 (2003) 4283–4293.
[63] K.T. Tycowski, M.D. Shu, A. Kukoyi, J.A. Steitz, A conserved WD40 protein binds the Cajal body localization signal of scaRNP particles, Mol. Cell 34 (2009) 47–57.
[64] X. Darzacq, B.E. Jady, C. Verheggen, A.M. Kiss, E. Bertrand, T. Kiss, Cajal body-specific small nuclear RNAs: a novel class of 2′-O-methylation and pseudouridylation guide RNAs, EMBO J. 21 (2002) 2746–2756.
[65] U.T. Meier, RNA modification in Cajal bodies, RNA Biol. 14 (2017) 693–700.
[66] J. Ofengand, M. Fournier, The pseudouridine residues of rRNA: number, location, biosynthesis, and function, in: H. Grosjean, R. Benne (Eds.), Modification and Editing of RNA, ASM Press, 1998, pp. 229–253 Place Published.
[67] R. Reddy, H. Busch, Small nuclear RNAs: RNA sequences, structure, and modifications, in: M.L. Birnstiel (Ed.), Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles, Springer-Verlag Press, 1988, pp. 1–37 Place Published.
[68] A.G. Balakin, L. Smith, M.J. Fournier, The RNA world of the nucleolus: two major families of small RNAs defined by different box elements with related functions, Cell 86 (1996) 823–834.
[69] P. Ganot, M.L. Bortolin, T. Kiss, Site-specific pseudouridine formation in preribosomal RNA is guided by small nucleolar RNAs, Cell 89 (1997) 799–809.
[70] W.A. Decatur, M.J. Fournier, rRNA modifications and ribosome function, Trends Biochem. Sci. 27 (2002) 344–351.
[71] T.H. King, B. Liu, R.R. McCully, M.J. Fournier, Ribosome structure and activity are altered in cells lacking snoRNPs that form pseudouridines in the peptidyl transferase center, Mol. Cell 11 (2003) 425–435.
[72] A. Baudin-Baillieu, C. Fabret, X.H. Liang, D. Piekna-Przybylska, M.J. Fournier, J.P. Rousset, Nucleotide modifications in three functionally important regions of the *Saccharomyces cerevisiae* ribosome affect translation accuracy, Nucleic Acids Res. 37 (2009) 7665–7677.
[73] X.H. Liang, Q. Liu, M.J. Fournier, Loss of rRNA modifications in the decoding center of the ribosome impairs translation and strongly delays pre-rRNA processing, RNA 15 (2009) 1716–1728.
[74] X.H. Liang, Q. Liu, M.J. Fournier, rRNA modifications in an intersubunit bridge of the ribosome strongly affect both ribosome biogenesis and activity, Mol. Cell 28 (2007) 965–977.
[75] D. Piekna-Przybylska, P. Przybylski, A. Baudin-Baillieu, J.P. Rousset, M.J. Fournier, Ribosome performance is enhanced by a rich cluster of pseudouridines in the A-site finger region of the large subunit, J. Biol. Chem. 283 (2008) 26026–26036.
[76] K. Jack, C. Bellodi, D.M. Landry, R.O. Niederer, A. Meskauskas, S. Musalgaonkar, N. Kopmar, O. Krasnykh, A.M. Dean, S.R. Thompson, D. Ruggero, J.D. Dinman, rRNA pseudouridylation defects affect ribosomal ligand binding and translational fidelity from yeast to human cells, Mol. Cell 44 (2011) 660–666.
[77] J. Ge, J. Karijolich, Y. Zhai, J. Zheng, Y.T. Yu, 5-Fluorouracil treatment alters the efficiency of translational recoding, Genes (Basel) 8 (2017).
[78] Y.T. Yu, M.D. Shu, J.A. Steitz, Modifications of U2 snRNA are required for snRNP assembly and pre-mRNA splicing, EMBO J. 17 (1998) 5783–5795.
[79] G. Donmez, K. Hartmuth, R. Luhrmann, Modified nucleotides at the 5′ end of human U2 snRNA are required for spliceosomal E-complex formation, RNA 10 (2004) 1925–1933.
[80] X. Zhao, Y.T. Yu, Pseudouridines in and near the branch site recognition region of U2 snRNA are required for snRNP biogenesis and pre-mRNA splicing in *Xenopus* oocytes, RNA 10 (2004) 681–690.
[81] C. Yang, D.S. McPheeters, Y.T. Yu, Psi35 in the branch site recognition region of U2 small nuclear RNA is important for pre-mRNA splicing in Saccharomyces cerevisiae, J. Biol. Chem. 280 (2005) 6655–6662.
[82] X. Zhao, Y.T. Yu, Incorporation of 5-fluorouracil into U2 snRNA blocks pseudouridylation and pre-mRNA splicing in vivo, Nucleic Acids Res. 35 (2007) 550–558.
[83] J.R. Patton, Multiple pseudouridine synthase activities for small nuclear RNAs, Biochem. J. 290 (Pt 2) (1993) 595–600.
[84] J.R. Patton, Ribonucleoprotein particle assembly and modification of U2 small nuclear RNA containing 5-fluorouridine, Biochemistry 32 (1993) 8939–8944.
[85] G. Wu, M.K. Radwan, M. Xiao, H. Adachi, J. Fan, Y.T. Yu, The TOR signaling pathway regulates starvation-induced pseudouridylation of yeast U2 snRNA, RNA 22 (2016) 1146–1152.
[86] F. Lecointe, O. Namy, I. Hatin, G. Simos, J.P. Rousset, H. Grosjean, Lack of pseudouridine 38/39 in the anticodon arm of yeast cytoplasmic tRNA decreases in vivo recoding efficiency, J. Biol. Chem. 277 (2002) 30445–30453.
[87] A. Alexandrov, I. Chernyakov, W. Gu, S.L. Hiley, T.R. Hughes, E.J. Grayhack, E.M. Phizicky, Rapid tRNA decay can result from lack of nonessential modifications, Mol. Cell 21 (2006) 87–96.
[88] K. Tomita, T. Ueda, K. Watanabe, The presence of pseudouridine in the anticodon alters the genetic code: a possible mechanism for assignment of the AAA lysine codon as asparagine in echinoderm mitochondria, Nucleic Acids Res. 27 (1999)

1683–1689.
[89] X. Zhao, J.R. Patton, S.K. Ghosh, N. Fischel-Ghodsian, L. Shen, R.A. Spanjaard, Pus3p- and Pus1p-dependent pseudouridylation of steroid receptor RNA activator controls a functional switch that regulates nuclear receptor signaling, Mol. Endocrinol. 21 (2007) 686–699.
[90] X. Zhao, J.R. Patton, S.L. Davis, B. Florence, S.J. Ames, R.A. Spanjaard, Regulation of nuclear receptor activity by a pseudouridine synthase through posttranscriptional modification of steroid receptor RNA activator, Mol. Cell 15 (2004) 549–558.
[91] Y. Zhao, J. Karijolich, B. Glaunsinger, Q. Zhou, Pseudouridylation of 7SK snRNA promotes 7SK snRNP formation to suppress HIV-1 transcription and escape from latency, EMBO Rep. 17 (2016) 1441–1451.
[92] J. Karijolich, Y.T. Yu, Converting nonsense codons into sense codons by targeted pseudouridylation, Nature 474 (2011) 395–398.
[93] I.S. Fernandez, C.L. Ng, A.C. Kelley, G. Wu, Y.T. Yu, V. Ramakrishnan, Unusual base pairing during the decoding of a stop codon by the ribosome, Nature 500 (2013) 107–110.
[94] C. Chen, X. Zhao, R. Kierzek, Y.T. Yu, A flexible RNA backbone within the polypyrimidine tract is required for U2AF65 binding and pre-mRNA splicing in vivo, Mol. Cell. Biol. 30 (2010) 4108–4119.
[95] M. Zaringhalam, F.N. Papavasiliou, Pseudouridylation meets next-generation sequencing, Methods 107 (2016) 63–72.
[96] M.Z. Anderson, J. Brewer, U. Singh, J.C. Boothroyd, A pseudouridine synthase homologue is critical to cellular differentiation in *Toxoplasma gondii*, Eukaryot. Cell 8 (2009) 398–409.
[97] K. Kariko, H. Muramatsu, F.A. Welsh, J. Ludwig, H. Kato, S. Akira, D. Weissman, Incorporation of pseudouridine into mRNA yields superior nonimmunogenic vector with increased translational capacity and biological stability, Mol. Ther. 16 (2008) 1833–1840.