The Role of Noncoding RNA Pseudouridylation in Nuclear Gene Expression Events

Yang Zhao[1], William Dunker[1], Yi-Tao Yu[2]* and John Karijolich[1,3]*

[1] *Department of Pathology, Microbiology, and Immunology, School of Medicine, Vanderbilt University, Nashville, TN, United States,* [2] *Department of Biochemistry and Biophysics, Center for RNA Biology, School of Medicine and Dentistry, University of Rochester, Rochester, NY, United States,* [3] *Vanderbilt-Ingram Cancer Center, Nashville, TN, United States*

Pseudouridine is the most abundant internal RNA modification in stable noncoding RNAs (ncRNAs). It can be catalyzed by both RNA-dependent and RNA-independent mechanisms. Pseudouridylation impacts both the biochemical and biophysical properties of RNAs and thus influences RNA-mediated cellular processes. The investigation of nuclear-ncRNA pseudouridylation has demonstrated that it is critical for the proper control of multiple stages of gene expression regulation. Here, we review how nuclear-ncRNA pseudouridylation contributes to transcriptional regulation and pre-mRNA splicing.

Keywords: RNA pseudouridylation, steroid receptor RNA activator, 7SK RNA, spliceosomal small nuclear RNA, transcription, pre-mRNA splicing

INTRODUCTION

The proper control of gene expression in the nucleus is achieved by the actions of a diverse set of factors. In addition to proteins, noncoding RNAs (ncRNAs) participate in most, if not all, stages of nuclear gene expression, including both RNA polymerase II (Pol II) transcription and pre-mRNA splicing (Mercer et al., 2009). ncRNAs employ numerous mechanisms to accomplish their function, including binding to and modulating protein function, base pairing with complementary nucleic acids, or directly catalyzing biochemical reactions. Like proteins, proper modification and folding into higher-order structures are a prerequisite for their function.

In addition to the four canonical nucleosides, more than 140 chemically distinct modified RNA nucleosides have been identified in nature (Machnicka et al., 2013). Pseudouridine (ψ), first discovered over 60 years ago (Cohn and Elliot, 1951), is the most abundant internal RNA modification in stable RNAs. ψ, the C5-glycoside isomer of uridine, is formed through an internal transglycosylation reaction in which the N1–C1' bond between the uracil base and the ribose sugar is broken and a C5–C1' glycosidic bond is reformed (**Figure 1**). As a consequence of isomerization, an additional hydrogen bond donor is present at the non-Watson–Crick edge. The distinct structure of ψ increases both the rigidity of the phosphodiester backbone and the thermodynamic stability of ψ–A compared with U–A. This effect is mediated by water-coordinated hydrogen bonding and base stacking (Charette and Gray, 2000).

Initial evidence for a functional role of ψ partially came from the fact that ψ residues are clustered in functionally important and evolutionarily conserved regions of tRNA (Grosjean et al., 1995; Hopper and Phizicky, 2003), rRNA (Branlant et al., 1981; Maden, 1990), and small nuclear RNA (snRNA) (Reddy and Busch, 1988; Massenet et al., 1998; Narlikar et al., 2002; Karijolich and Yu, 2010). Indeed, experimental data have confirmed important roles for pseudouridylation in multiple aspects of gene expression regulation, including spliceosomal small nuclear ribonucleoprotein (snRNP) biogenesis, efficiency of pre-mRNA splicing, and translation fidelity (Karijolich et al., 2010). Here, we will provide an overview of the mechanisms of pseudouridylation and then highlight
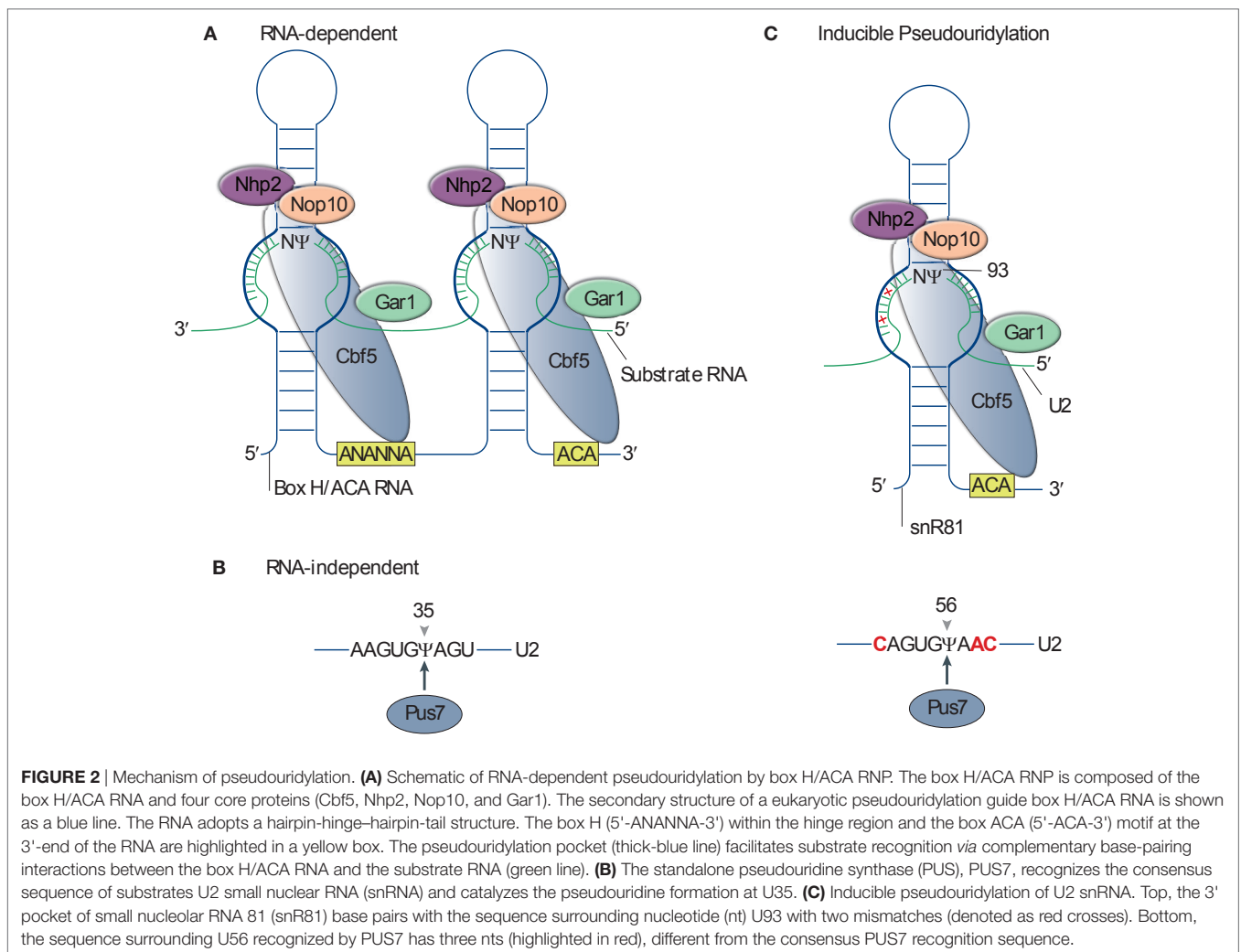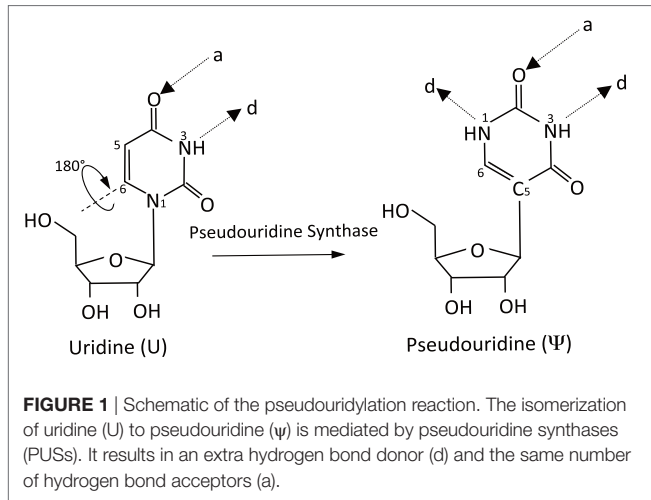
several pseudouridylated ncRNAs and their effects on nuclear gene expression events, specifically transcription and pre-mRNA splicing.

FIGURE 1 | Schematic of the pseudouridylation reaction. The isomerization of uridine (U) to pseudouridine (ψ) is mediated by pseudouridine synthases (PUSs). It results in an extra hydrogen bond donor (d) and the same number of hydrogen bond acceptors (a).

MECHANISMS OF PSEUDOURIDYLATION

The past decades have seen remarkable progress toward defining mechanisms by which pseudouridylation is catalyzed. Pseudouridylation of ncRNA is catalyzed by pseudouridine synthases (PUSs) through two distinct mechanisms, namely RNA-dependent pseudouridylation and RNA-independent pseudouridylation.

RNA-DEPENDENT PSEUDOURIDYLATION

The RNA-dependent pseudouridylation machinery consists of one unique box H/ACA RNA and four core proteins. Box H/ACA RNAs are one of the most evolutionarily conserved families of small ncRNAs and are present in all eukaryotes. They function as guide RNAs to direct pseudouridylation in mRNA, rRNA, spliceosomal snRNAs, and various other types of ncRNAs. With a median length of 133 nucleotides (nts), eukaryotic box H/ACA RNAs adopt a hairpin-hinge–hairpin-tail secondary structure (**Figure 2A**). Within the internal hinge region, there is a H box motif (ANANNA), while an ACA box motif (ACA) is positioned near the 3' end. Within each hairpin structure,

FIGURE 2 | Mechanism of pseudouridylation. **(A)** Schematic of RNA-dependent pseudouridylation by box H/ACA RNP. The box H/ACA RNP is composed of the box H/ACA RNA and four core proteins (Cbf5, Nhp2, Nop10, and Gar1). The secondary structure of a eukaryotic pseudouridylation guide box H/ACA RNA is shown as a blue line. The RNA adopts a hairpin-hinge–hairpin-tail structure. The box H (5'-ANANNA-3') within the hinge region and the box ACA (5'-ACA-3') motif at the 3'-end of the RNA are highlighted in a yellow box. The pseudouridylation pocket (thick-blue line) facilitates substrate recognition *via* complementary base-pairing interactions between the box H/ACA RNA and the substrate RNA (green line). **(B)** The standalone pseudouridine synthase (PUS), PUS7, recognizes the consensus sequence of substrates U2 small nuclear RNA (snRNA) and catalyzes the pseudouridine formation at U35. **(C)** Inducible pseudouridylation of U2 snRNA. Top, the 3' pocket of small nucleolar RNA 81 (snR81) base pairs with the sequence surrounding nucleotide (nt) U93 with two mismatches (denoted as red crosses). Bottom, the sequence surrounding U56 recognized by PUS7 has three nts (highlighted in red), different from the consensus PUS7 recognition sequence.

an internal loop (pseudouridylation pocket) is present, which facilitates substrate recognition *via* complementary base-pairing interactions between the box H/ACA RNA (the loop sequence) and the substrate RNA. Both hairpins of H/ACA RNAs can carry functional pseudouridylation pockets and can thus independently direct pseudouridylation of uridines in separate RNAs or separate uridines located within the same substrate RNA. In the guide RNA–target RNA interaction, the guide sequences hybridize 5' and one nt 3' of the target uridine to the substrate RNA immediately, thereby framing it. The distance between the target uridine and the H or ACA box of the guide RNA is usually 14–15 nts (Kiss et al., 2010).

The four core proteins associated with box H/ACA RNAs are Cbf5 (dyskerin in human and NAP57 in rodents), Nhp2 (L7Ae in archaeal), Gar1, and Nop10 (Yu and Meier, 2014). Cbf5 is the enzymatic component of the RNP and catalyzes the U-to-ψ isomerization reaction (**Figure 2A**). Structures of the enzymes from various species show a high degree of evolutionary conservation, especially in the PUS and Archaeosine transglycosylase (PUA) domain (Hamma et al., 2005; Manival et al., 2006; Rashid et al., 2006). The other three core proteins are also essential, and the depletion of them in yeast, with the exception of GAR1, causes the loss of all H/ACA RNAs (Girard et al., 1992; Bousquet-Antonelli et al., 1997).

Facilitated by the development of *in vitro* systems for the reconstitution of enzymatically active RNP complexes, the crystal structure of the box H/ACA RNP was first solved using archaeal components (Li and Ye, 2006). These experiments demonstrated that the L7Ae, Nop10, and the catalytic domain of Cbf5 bound to the upper stem of the guide RNA, whereas the PUA domain of Cbf5 anchored the lower stem and ACA motif. The substrate RNA is recruited and the target uridine is precisely placed within the active site *via* complementary base-pairing interactions with the bipartite guides, while extensive protein interactions help to stabilize the interaction. In contrast to the other core proteins, Gar1 does not physically interact with the box H/ACA guide RNA or substrate RNA. Instead, Gar1 interacts directly with the thumb loop of Cbf5 and participates in regulating substrate turnover. Studies in yeast have revealed that the structure of the eukaryotic box H/ACA RNP is highly similar to the archaeal one, however, with exceptions. In particular is the independence of the RNP's activity from Nhp2 binding and a novel C-terminal extension in Gar1, which interacts with Cbf5. It is hypothesized that these functional and structural differences reflect the evolutionary adaptations of eukaryotic box H/ACA RNP to the variable RNA structure and moderate temperature range in which eukaryotes live in, respectively (Li et al., 2011).

RNA-INDEPENDENT PSEUDOURIDYLATION

RNA-independent pseudouridylation in eukaryotes acts through standalone enzymes called PUS enzymes. In contrast to the box H/ACA RNP-based mechanism, PUS enzymes carry out both substrate recognition and the internal transglycosylation reaction. Substrate recognition is achieved *via* consensus sequences and/or

secondary structure elements of the substrate RNA (**Figure 2B**). In eukaryotes, there are 10 different PUS enzymes, numbered PUS1 through PUS10. These are classified into five families (TruA, TruB, TruD, RluA, and PUS10) based on their bacterial counterparts (Hamma and Ferre-D'Amare, 2006). Although the primary sequences have diverged, all PUSs, including Cbf5, share a conserved catalytic domain and likely a conserved catalytic mechanism based on the solved crystal structure (Foster et al., 2000; Hoang and Ferre-D'Amare, 2001, 2004; Sivaraman et al., 2002, 2004; Del Campo et al., 2004; Ericsson et al., 2004; Kaya et al., 2004; Mizutani et al., 2004; Hoang et al., 2006; McCleverty et al., 2007). This domain structure is composed predominately of anti-parallel β-sheets, with one face decorated by two groups of α-helices and loops. A forefinger–thumb structure formed by these loops pinches the target RNA, while the strictly conserved catalytic aspartate residue participates in the enzymatic reaction.

Unlike the box H/ACA RNPs, which have been found only to reside within the nucleus, PUS enzymes have been found in the nucleus, cytoplasm, and mitochondria. Each PUS enzyme targets either one specific or multiple uridines in many RNA species, including snRNAs, rRNAs, and tRNAs in both cytoplasm and mitochondria.

PSEUDOURIDYLATION IS INDUCIBLE

Until relatively recently, RNA modifications, including pseudouridylation, were considered constitutive. In 2011, Wu et al. (2011) provided the first evidence that RNA modifications were inducible and demonstrated that yeast U2 snRNA was conditionally pseudouridylated when cells were subjected to nutrient deprivation or heat shock (Wu et al., 2011). In addition to the three constitutive pseudouridines (ψ35, ψ42, and ψ44) of yeast U2 snRNA, two novel pseudouridines (ψ56 and ψ93) were detected in stressed cells. Further detailed analyses revealed that both the RNA-independent PUS (PUS7 catalyzes ψ56 formation) and the box H/ACA RNP-dependent [small nucleolar RNA 81 (snR81)-guided box H/ACA RNP catalyzes ψ93 formation] modification machineries are involved in inducible pseudouridylation. Interestingly, both ψ56 and ψ93 are "imperfect" substrates for their respective modification machineries. The sequences flanking positions U56 and U93 in U2 snRNA are similar, but not identical to the sequences surrounding the constitutively pseudouridylated targets of PUS7 and snR81, respectively. For example, the 3' pocket of snR81 base pairs with the sequence surrounding nt U93 of U2 snRNA with two mismatches. In addition, ψ56 formation is mediated by PUS7 engaging a substrate whose sequence differs by three nts from the consensus PUS7 recognition sequence (**Figure 2C**). Inducible pseudouridylation is also functionally relevant, as demonstrated by the observation that the artificial introduction of ψ93 reduces the efficiency of pre-mRNA splicing. Interestingly, it was recently shown that the TOR-signaling pathway regulates ψ93 formation (Wu et al., 2016b).

Inducible pseudouridylation of other snRNAs has been reported subsequently. U6 snRNA is inducibly pseudouridylated at U28 by PUS1 during the yeast filamentous growth program (Basak and Query, 2014). Further analysis of mutants indicates that U6–ψ28 is functionally relevant, as all U6 snRNA mutations

that resulted in strong pseudouridylation at position U28 also exhibited a pseudohyphal growth phenotype, whereas blocking U6–ψ28 formation prevents filamentous growth.

Recently, transcriptome-wide mapping of pseudouridines in yeast and human cells has revealed its presence in mRNA (Carlile et al., 2014; Lovejoy et al., 2014; Schwartz et al., 2014). While mRNA pseudouridylation appears to be primarily catalyzed by the standalone PUSs (PUS1–PUS4, PUS6, PUS7, and PUS9), several pseudouridine residues are catalyzed by box H/ACA RNPs. Remarkably, mRNA pseudouridylation was also found to be highly inducible and in a stress-specific manner. For example, by comparing mRNA ψ profiles of untreated cells to those of cells exposed to heat shock or H_2O_2, it was found that the inducible pseudouridylation profiles were largely nonoverlapping (Li et al., 2015).
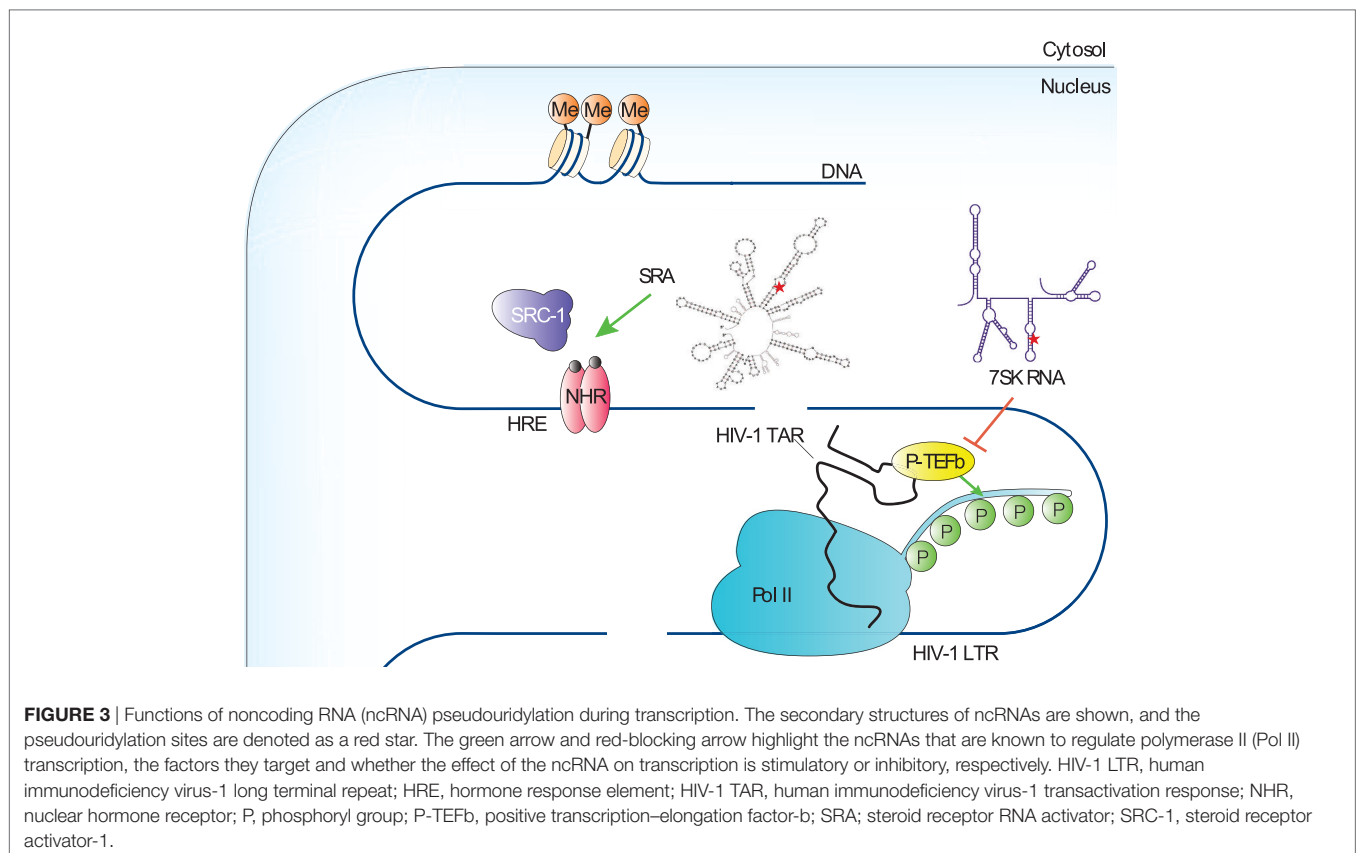
FUNCTION OF ncRNA PSEUDOURIDYLATION IN TRANSCRIPTION

Transcription is the primary control point for gene expression. It therefore determines cellular function and cell identity and is subjected to tight regulation to achieve a high degree of specificity and efficiency. The eukaryotic DNA template is packaged by histone proteins into a highly condensed structure called chromatin. The chromatin structure is dynamically regulated by both histone modifications and chromatin-remodeling factors (Narlikar et al.,

2002). Promoters contain elements that bind to transcriptional activators and repressors, as well as the transcription machinery (Smale and Kadonaga, 2003; Kadonaga, 2004). RNA Pol II is the enzyme to catalyze the transcription reaction of mRNA from DNA. Pol II is recruited to promoters by transcriptional activators in a holoenzyme form together with general transcription factors and a multiprotein complex called the Srb/Mediator (Bjorklund and Kim, 1996). Following transcription initiation, Pol II transits to a productive elongation status through interactions with multiple elongation factors (Zhou et al., 2012). Given the central role of transcription in gene expression, it is not surprising that transcription is subject to diverse steps of regulation. Here, we will discuss two pseudouridylated ncRNAs and their function in regulating RNA Pol II transcription (**Figure 3**).

STEROID RECEPTOR RNA ACTIVATOR (SRA) AND TRANSCRIPTION PREINITIATION

One layer of transcriptional control comes from the binding of activator and repressor proteins to the promoters of target genes in a sequence-specific manner (Smale and Kadonaga, 2003). Coactivators and corepressors, which interact with activators and repressors, are required to achieve optimal transcriptional regulation in cells (Kadonaga, 2004). The SRA was first identified as a transcriptional coactivator for several steroid-hormone receptors, including receptors for androgens (ARs), estrogens

FIGURE 3 | Functions of noncoding RNA (ncRNA) pseudouridylation during transcription. The secondary structures of ncRNAs are shown, and the pseudouridylation sites are denoted as a red star. The green arrow and red-blocking arrow highlight the ncRNAs that are known to regulate polymerase II (Pol II) transcription, the factors they target and whether the effect of the ncRNA on transcription is stimulatory or inhibitory, respectively. HIV-1 LTR, human immunodeficiency virus-1 long terminal repeat; HRE, hormone response element; HIV-1 TAR, human immunodeficiency virus-1 transactivation response; NHR, nuclear hormone receptor; P, phosphoryl group; P-TEFb, positive transcription–elongation factor-b; SRA; steroid receptor RNA activator; SRC-1, steroid receptor activator-1.

(ERs), glucocorticoids (GRs), and progestins (PRs) (Lanz et al., 1999). Interestingly, while it was initially presumed that a protein encoded by a specific 5'-spliced variant of the SRA gene was the functional factor, subsequent experiments demonstrated that the factor was an ncRNA.

Steroid receptor RNA activator operates as part of a ribonucleoprotein complex containing steroid receptor activator-1, which is an AF-2 coactivator (Lanz et al., 1999). Computer-assisted modeling suggests that SRA adopts a highly complex secondary structure containing 11 topological substructures (STRs). Mutagenesis of each STR indicated that 5 of 11 STRs are required for SRA to coactivate transcription, and STR7 is the most important one for SRA function (Lanz et al., 2002).

In a study to identify coactivators for retinoic acid receptor (RAR) in mouse S91 melanoma cells, mPUS1p was unexpectedly identified. In addition, SRA turned out to be a substrate of mPUS1p (Zhao et al., 2004). Using chromatin immunoprecipitation, RAR, PUS1p, and SRA were found to cooccupy the retinoic acid response promoter in a ligand-independent complex. PUS1-mediated pseudouridylation of SRA promotes the formation of an "active" structure and aids in establishing the transcription preinitiation complex upon ligand binding. Further supporting a role of SRA pseudouridylation in transcriptional regulation, mutations in PUS1p that disrupt its interaction with RAR or its pseudouridylation activity attenuate the activation of RAR-dependent transcription. mPUS1p also significantly augmented transactivation by other nuclear receptors (NRs) including thyroid hormone receptor (TR), GR, AR, PR, and ER, illustrating that this mechanism likely applies universally to the regulation of NR-dependent transcription.

In addition to mPUS1p, mPUS3p also modifies SRA and serves as an NR coactivator (Zhao et al., 2007). Unlike mPUS1p, mPUS3p does not enhance sex steroid receptor activity, suggesting that substrate-site specificity may have distinct roles. Indeed, *in vitro* mPUS1p and mPUS3p generally modify different positions in SRA, with a few positions commonly targeted. Intriguingly, the order of modification of SRA by mPUS1p and mPUS3p determines the positions within SRA that are required to be pseudouridylated. However, it is important to note that the only *in vivo*-pseudouridylated site identified in SRA is U206. Interestingly, a U206A mutation, which promotes hyperpseudouridylation of SRA *in vitro*, switches SRA from a coactivator to a molecule with dominant-negative activity *in vivo*.

Pseudouridylation of SRA by mPUS1p and mPUS3p is a highly complex posttranscriptional mechanism that controls a coactivator–corepressor switch in SRA with major consequences for NR signaling (Zhao et al., 2007). Unexpectedly, pseudouridylation of SRA occurs in a stem-loop structure STR5 (at position U206), whose secondary structure was not shown to be important for SRA function. Moreover, the thermodynamic and secondary structure differences between STR5 in hSRA-WT and hSRA-U206A are relatively minor, which further suggest that the secondary structure remodeling is unlikely to explain the large biochemical and functional effects observed. Instead, it is proposed that ψ206 stabilizes stems I and II of STR5 in a higher-order conformation through the base-stacking-enhancing properties of ψ, resulting in masking new sites and preventing hyperpseudouridylation by mPUS1p and mPUS3p. This in turn may interfere with the binding of SRA to other proteins that define its function as a scaffold for both repressors and activators.

The physiological importance of PUS1p-mediated SRA pseudouridylation is illustrated by a disorder known as mitochondrial myopathy and sideroblastic anemia (MLASA). It is caused by an inactivating mutation in human PUS1p (Bykhovskaya et al., 2004). Some abnormalities in these patients, such as facial dysmorphisms, are suggested to be the consequence of defective hSRA–NR signaling (Fernandez-Vizarra et al., 2007). In addition, ERs and ARs are important targets for cancer therapy. Given the importance of SRA in NR-transcriptional regulation, coupled with the functional significance of PUS1p-mediated SRA pseudouridylation, recent work has suggested that the disruption of SRA pseudouridylation could serve as a novel RNA-based cancer therapeutic (Ghosh et al., 2012).

7SK snRNA AND TRANSCRIPTION ELONGATION

For the past three decades, most of the attention has been put on the early stages of the transcription cycle involving the recruitment of Pol II to gene promoters and assembly of active preinitiation complexes, which were thought to be the principal points where transcription was controlled (Kuras and Struhl, 1999; Ptashne, 2005). In recent years, however, accumulating evidences indicate that the subsequent stages of the transcription cycle are also highly regulated (Guenther et al., 2007; Muse et al., 2007). Notably, the promoter-proximal pausing and release of Pol II has been identified as a major rate-limiting step for controlling the expression of many metazoan genes and plays a critical role in cell growth, renewal, and differentiation (Levine, 2011; Zhou et al., 2012).

Shortly after initiation, Pol II is paused at a promoter-proximal region by negative elongation factors NELF and DSIF, resulting in a short nascent transcript ~12 nts in length. Promoter-proximately paused Pol II is the major form of Pol II found on metazoan chromosomes and is poised for entry into productive elongation. The positive transcription elongation factor-b (P-TEFb) is the major factor required to overcome this restriction. P-TEFb, composed of cyclin-dependent kinase 9 (CDK9) and cyclin T1, phosphorylates and thereby antagonizes the inhibitory actions of NELF and DSIF, triggering the release of Pol II from promoter-proximal pausing. In addition to this, P-TEFb also phosphorylates the C-terminal domain of the largest subunit of Pol II, which then serves as a platform for assembling key transcription and RNA-processing factors that promote transcriptional elongation, cotranscriptional processing of pre-mRNA, and lastly termination (Zhou et al., 2012).

Under normal growth conditions, up to 90% of cellular P-TEFb is sequestered in an inactive complex called the 7SK RNP (Yang et al., 2001). Within this complex, a noncoding RNA, namely 7SK snRNA, functions as a scaffold and mediates the interaction of P-TEFb with HEXIM1 or -2 and thus inhibits CDK9's kinase activity (Yik et al., 2003). 7SK snRNA, transcribed by RNA Pol III, is an abundant noncoding RNA and highly conserved in

higher eukaryotes (Wassarman and Steitz, 1991). Although its levels remain relatively constant, several genome-wide studies suggested that 7SK was pseudouridylated (Kishore et al., 2013; Carlile et al., 2014). Indeed, site-specific and quantitative pseudouridylation assays demonstrated that up to 94% of 7SK snRNA in HeLa cells is pseudouridylated at residue U250. Although the guide RNA has not been identified, it is clear that the box H/ACA RNP machinery catalyzes this modification as the depletion of DKC1 significantly reduces 7SK snRNA pseudouridylation. In addition, 7SK snRNA pseudouridylation was demonstrated to play a critical role in regulating the formation of the 7SK-P-TEFb snRNP, as mutation of U250, or depletion of DKC1, reduced the binding of CDK9 and HEXIM to 7SK snRNA (Zhao et al., 2016).

The identification of a key role for pseudouridylation in 7SK snRNP stability has potential clinical relevance. For instance, one emerging strategy in curing human immunodeficiency virus (HIV) infection is to "shock," or reactivate, the latent HIV reservoirs for their subsequent "kill" by Highly Active Anti-Retroviral Therapy (Richman et al., 2009; Deeks, 2012). Along this line, Zhao et al. demonstrated that reducing 7SK snRNA pseudouridylation

destabilized the 7SK snRNP and activated Tat-dependent HIV-1 transcription. Furthermore, reduction in 7SK snRNA pseudouridylation in combination with latency reversal agents (LRAs) significantly increased the reversal of HIV latency, implicating the DKC1-box H/ACA RNP as a promising new target to eradicate latent viral reservoirs (Zhao et al., 2016). It will be interesting to determine whether the widely used chemotherapeutic agent 5-fluorouracil, which is an inhibitor of PUS enzymes, can act synergistically with current LRAs to further enhance the reversal of HIV latency.

FUNCTION OF SPLICEOSOMAL snRNA PSEUDOURIDYLATION IN PRE-mRNA SPLICING

Most genes in eukaryotes are not transcribed in mature but rather as pre-mRNA, containing coding exons as well as noncoding introns. Therefore, further pre-mRNA splicing is needed to remove intronic sequences and assemble exons into mature
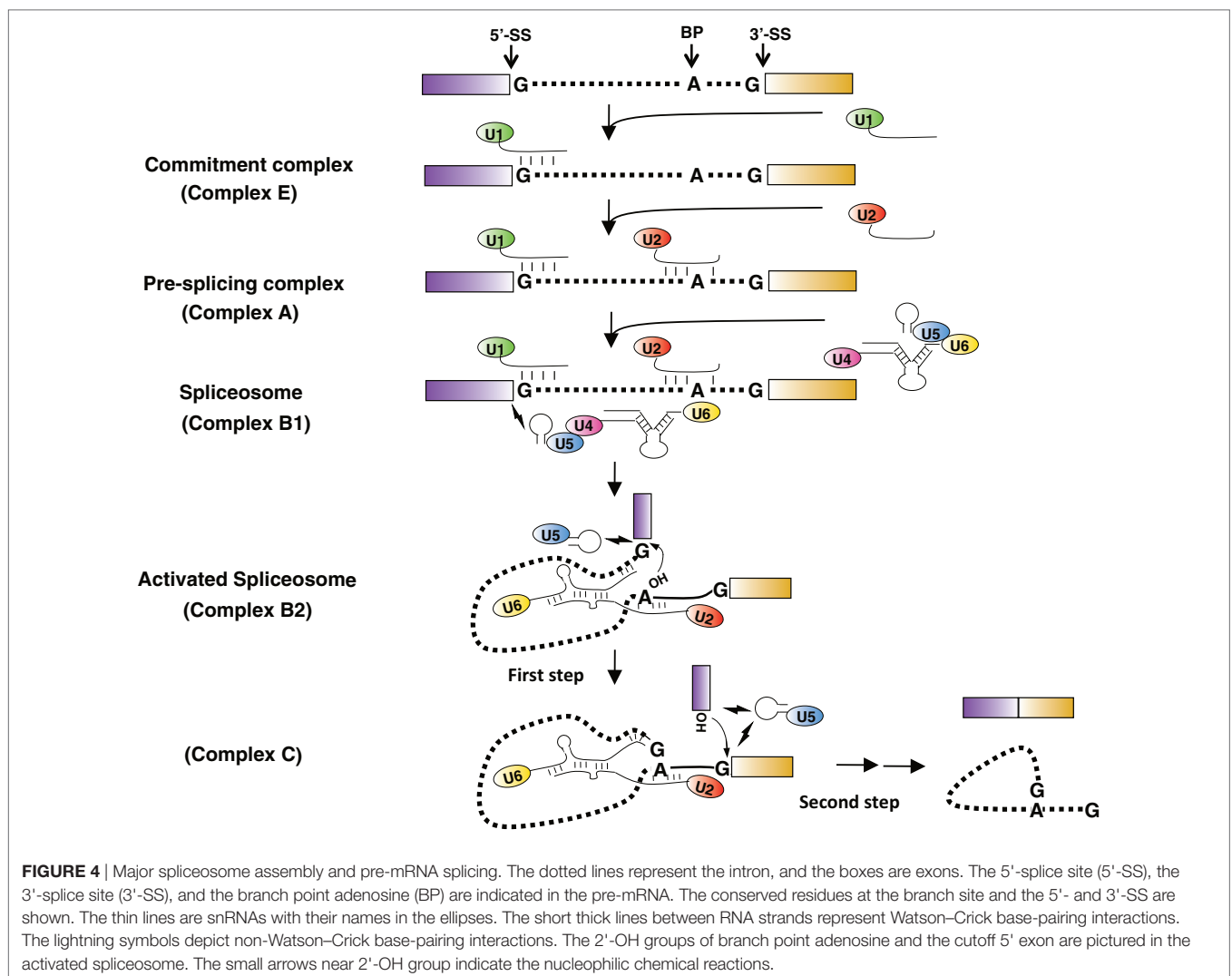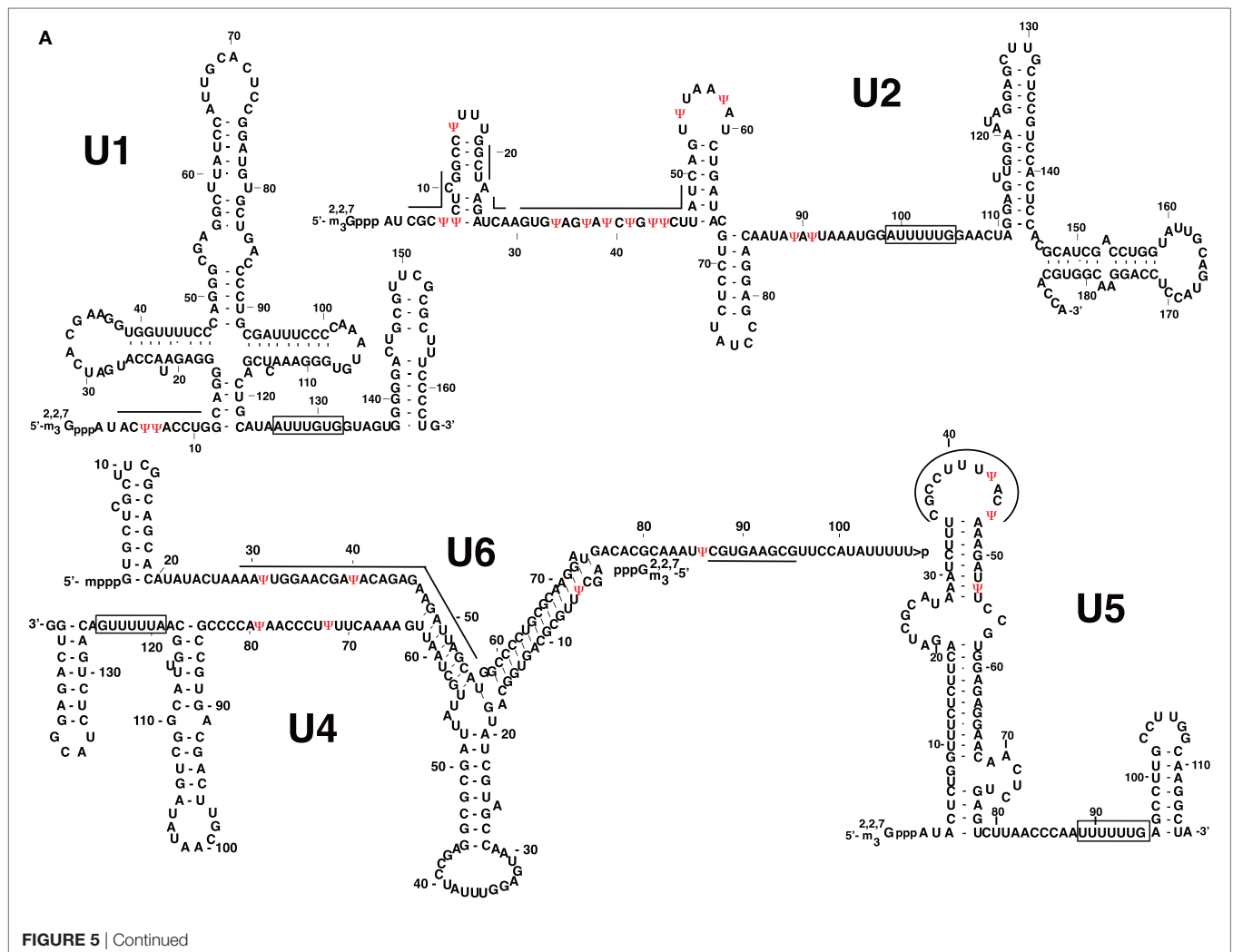

FIGURE 4 | Major spliceosome assembly and pre-mRNA splicing. The dotted lines represent the intron, and the boxes are exons. The 5'-splice site (5'-SS), the 3'-splice site (3'-SS), and the branch point adenosine (BP) are indicated in the pre-mRNA. The conserved residues at the branch site and the 5'- and 3'-SS are shown. The thin lines are snRNAs with their names in the ellipses. The short thick lines between RNA strands represent Watson–Crick base-pairing interactions. The lightning symbols depict non-Watson–Crick base-pairing interactions. The 2'-OH groups of branch point adenosine and the cutoff 5' exon are pictured in the activated spliceosome. The small arrows near 2'-OH group indicate the nucleophilic chemical reactions.
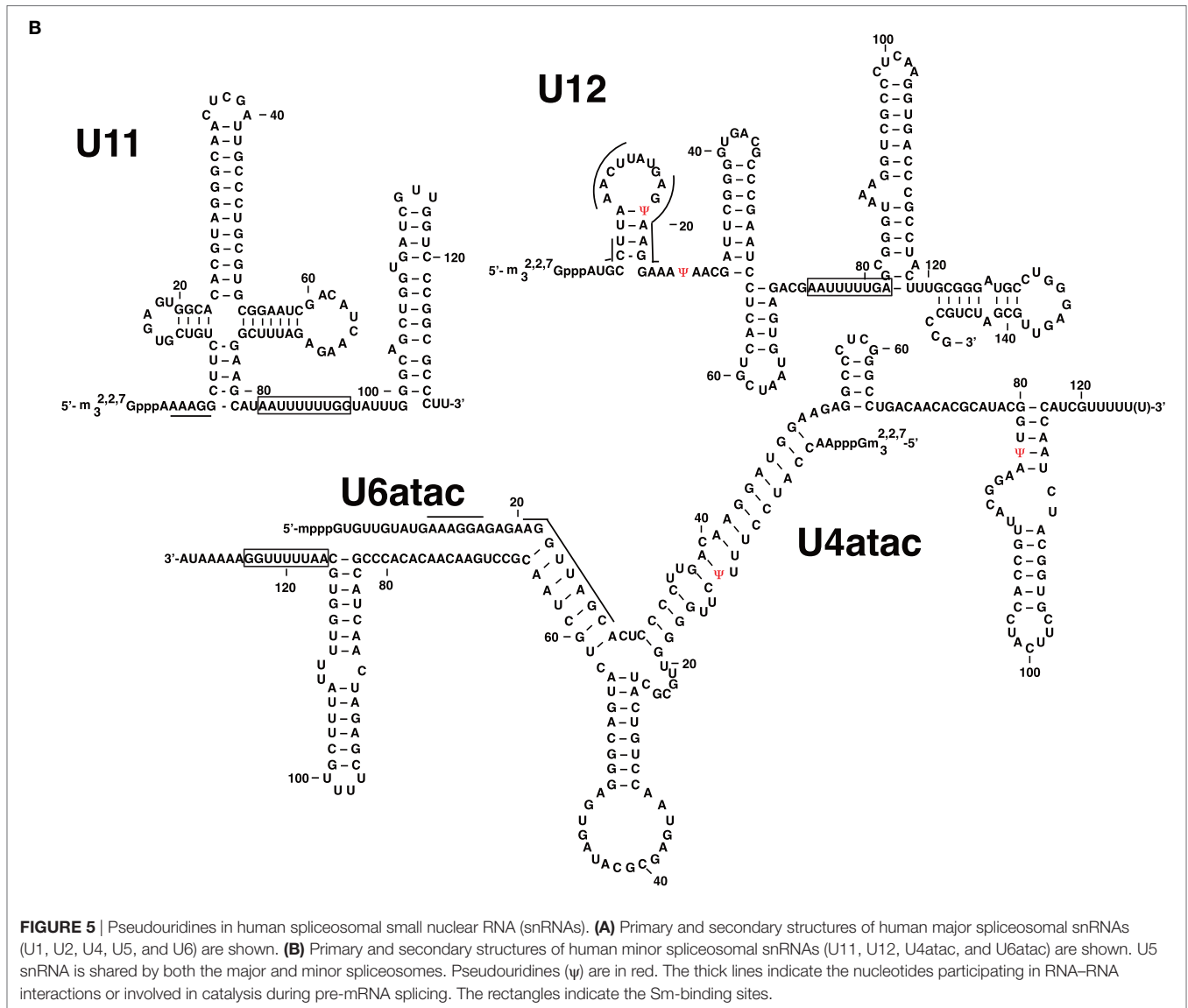
mRNA (Green, 1986; Newman, 1994). Splicing is catalyzed by the spliceosome, a massively large complex consisting of snRNAs and numerous protein components (Wahl et al., 2009; Matera and Wang, 2014). There are five snRNAs within the major spliceosome—U1, U2, U4, U5, and U6 that participate in the splicing reaction as snRNP complexes.

Pre-mRNA splicing is initiated by the recognition of the 5'-splice site (5'-SS) by the U1 snRNP *via* complementary base-pairing interactions (Kramer et al., 1984; Bindereif and Green, 1987; Ruby and Abelson, 1988; Seraphin and Rosbash, 1989). The branch-site sequence is then engaged by the U2 snRNP *via* complementary base-pairing interactions, resulting in the bulging out of the branch point adenosine of the pre-mRNA and the formation of the spliceosomal complex A (Zhuang and Weiner, 1989; Michaud and Reed, 1991; Wassarman and Steitz, 1992). Subsequently, the tri-snRNP, a complex of U4 snRNP, U6 snRNP, and U5 snRNP, is recruited, creating a fully assembled spliceosome (complex B1). Following a series of RNA–RNA rearrangements, U1 and U4 snRNPs are destabilized and released, resulting in the formation of an active spliceosome (complex B2) (Sawa and Abelson, 1992; Lesser and Guthrie,

1993). This complex catalyzes the first step of pre-mRNA splicing in which the 2'-OH group of the bulged-out branch point adenosine nucleophilically attacks the phosphate at the 5'-SS. The result of the first-step reaction is the generation of a lariat 2/3 intermediate and a cutoff 5' exon intermediate. After additional conformational changes, complex B2 is converted to complex C, and the second step of splicing is catalyzed, resulting in the production of mature mRNA and lariat intron products. The U2, U5, and U6 snRNPs are recycled for new rounds of pre-mRNA splicing (**Figure 4**) (Burge et al., 1999).

Interestingly, all of the major spliceosomal snRNAs are post-transcriptionally pseudouridylated (**Figure 5A**). U2 snRNA is the most extensively pseudouridylated snRNA, and unsurprisingly investigations into snRNA pseudouridylation have primarily focused on U2 snRNA. The functional study of U2 snRNA pseudouridylation was initiated in the early 1990s by Patton, who prevented U2 snRNA pseudouridylation in HeLa cell S100 and nuclear extracts by the incorporation of 5-fluorouridine (5-FU) (Patton, 1993a,b). Although 5-FU-substituted U2 snRNA was able to form a U2 snRNP, the snRNP was more vulnerable to dissociation by salt.

FIGURE 5 | Continued

FIGURE 5 | Pseudouridines in human spliceosomal small nuclear RNA (snRNAs). **(A)** Primary and secondary structures of human major spliceosomal snRNAs (U1, U2, U4, U5, and U6) are shown. **(B)** Primary and secondary structures of human minor spliceosomal snRNAs (U11, U12, U4atac, and U6atac) are shown. U5 snRNA is shared by both the major and minor spliceosomes. Pseudouridines (ψ) are in red. The thick lines indicate the nucleotides participating in RNA–RNA interactions or involved in catalysis during pre-mRNA splicing. The rectangles indicate the Sm-binding sites.

As experimental systems and assays developed, a more detailed description of the effects of U2 snRNA pseudouridylation on pre-mRNA splicing emerged. For example, Yu et al. demonstrated that while *in vitro*-transcribed U2 snRNA, which lacks modification, was unable to rescue a splicing defect in U2 snRNA-depleted oocytes, following prolonged reconstitution periods, U2 snRNA was pseudouridylated and able to reconstitute splicing activity. In addition, anti-snRNP immunoprecipitation coupled with glycerol-gradient sedimentation demonstrated that U2 snRNA lacking pseudouridine was unable to form functional 17 S U2 snRNP. Thus, a good correlation between modification status, U2 snRNP biogenesis, and pre-mRNA splicing was established. In addition, by creating chimeric U2 snRNAs between cellular-derived and *in vitro*-transcribed U2, Yu et al. demonstrated that the functionally important modifications primarily resided within the first 27 nts of U2 snRNA (Yu et al., 1998).

Pseudouridylation of residues within the branch-site recognition region is also functionally important for pre-mRNA splicing. Zhao and Yu demonstrated that pseudouridine residues within the branch-site recognition region of Xenopus U2 snRNA are required for U2 snRNP assembly and spliceosome assembly (Zhao and Yu, 2007). In addition, an NMR structure of yeast U2 snRNA:pre-mRNA branch-site helix demonstrated that ψ35 induces a dramatic structural alternation, which is required for the bulging out of the branch point adenosine and nucleophilic attack on the 5'-SS (Newby and Greenbaum, 2002). Consistent with this, a yeast knockout of PUS7, which catalyzes ψ35 of U2 snRNA, exhibited reduced fitness under conditions of high-salt media, or when in competition with a wild-type strain (Ma et al., 2003). The other pseudouridines within this region, ψ42 and ψ44, are also functionally relevant (Wu et al., 2016a). The deletion of both SNR81, responsible for ψ42, and PUS1, which catalyzes ψ44 formation, reduces the efficiency of pre-mRNA splicing, leading to growth defect. Further genetic and biochemical analyses demonstrated that U2 snRNA ψ42 and ψ44 facilitate the interaction with Prp5, a U2-dependent ATPase known to play

an important role in monitoring U2-intron branch-site interactions at early stages of spliceosome assembly (Xu and Query, 2007). Furthermore, Prp5 has reduced ATPase activity on U2 snRNA lacking ψ42 and ψ44, suggesting that these modifications regulate Prp5 enzymatic activity. Collectively, these data indicate that pseudouridylation within the branch-site recognition region plays a role not only in the biogenesis of functional snRNP but also in spliceosome assembly by influencing the enzymatic activity of Prp5.

In contrast to U2 snRNA, pseudouridylations within the other spliceosomal U snRNAs have received much less attention. However, this is not to say that they have not been investigated. For instance, the functional importance of two ψs within the 5' end of U1 snRNA has been investigated. Adopting an *in vitro*-splicing system in which two 5'-SS are in competition with each other, Roca et al. suggested that U1 snRNA pseudouridylation participates in 5'-SS discrimination (Roca et al., 2005). In addition, Freund et al. demonstrated that a ψ–G base pair, between the U1 snRNA and the substrate pre-mRNA, respectively, stabilized the U1 snRNA interaction with the 5'-SS of HIV-1 RNA (Freund et al., 2003). Lastly, as described earlier, the inducible pseudouridylation of U6 snRNA at U28 during the yeast filamentous growth program is also functionally relevant, as shown by differential growth phenotypes that are dependent on its pseudouridylation (Basak and Query, 2014).

CONCLUSION

Remarkable progress has been made toward elucidating the mechanism and function of RNA pseudouridylation in various cellular processes. However, a function for nc RNA pseudouridylation in transcriptional regulation has only just begun to emerge. Our limited understanding of the impact of pseudouridylation on transcription is partially due to the limited number of ncRNAs that were known to be pseudouridylated. However, recent efforts of transcriptome-wide mapping of RNA pseudouridylation have greatly expanded the catalog of known pseudouridylated RNAs and have identified novel modification sites within ncRNAs participating in transcriptional regulation.

For instance, metastasis-associated lung adenocarcinoma transcript 1 (MALAT1), which is involved in the epigenetic modulation of gene expression, as well as alternative splicing, is pseudouridylated at two distinct sites (Carlile et al., 2014). How these modifications contribute to MALAT1 function remains unclear and certainly worth investigating. As more ψs in ncRNA transcriptional regulators are identified, a better understanding of how these ncRNAs impact eukaryotic mRNA transcription will be achieved. Since many of these ncRNAs are associated with diseases, determining how the functional impact of their pseudouridylation may open the door to novel therapeutic strategies.

Although the mechanism and function of snRNA pseudouridylation in splicing is considerably well studied, there are still many unanswered questions. For example, in contrast to U2 snRNA, the functional significance of pseudouridylation within U1, U4, U5, and U6 is not clear. In addition, within higher eukaryotes, there exists a second spliceosome, the minor spliceosome, which consists of U5, in addition to four distinct U snRNAs, U11, U12, U4atac, and U6atac, and all of them are pseudouridylated (**Figure 5B**). Interestingly, the positions of minor spliceosomal snRNA pseudouridylation are homologous to those within major spliceosomal snRNAs, suggesting their importance in minor intron splicing (Massenet and Branlant, 1999). Detailed functional analysis of these pseudouridylations is required if we seek to understand the mechanism of minor spliceosome biogenesis and minor intron splicing. In addition, these studies may provide a better understanding of how introns are selectively recognized by the two distinct spliceosomes.

AUTHOR CONTRIBUTIONS

All authors contributed to writing the manuscript.

FUNDING

REFERENCES

Basak, A., and Query, C. C. (2014). A pseudouridine residue in the spliceosome core is part of the filamentous growth program in yeast. *Cell Rep.* 8, 966–973. doi:10.1016/j.celrep.2014.07.004

Bindereif, A., and Green, M. R. (1987). An ordered pathway of snRNP binding during mammalian pre-mRNA splicing complex assembly. *EMBO J.* 6, 2415–2424.

Bjorklund, S., and Kim, Y. J. (1996). Mediator of transcriptional regulation. *Trends Biochem. Sci.* 21, 335–337. doi:10.1016/S0968-0004(96)10051-7

Bousquet-Antonelli, C., Henry, Y., G'Elugne, J. P., Caizergues-Ferrer, M., and Kiss, T. (1997). A small nucleolar RNP protein is required for pseudouridylation of eukaryotic ribosomal RNAs. *EMBO J.* 16, 4770–4776. doi:10.1093/emboj/16.15.4770

Branlant, C., Krol, A., Machatt, M. A., Pouyet, J., Ebel, J. P., Edwards, K., et al. (1981). Primary and secondary structures of *Escherichia coli* MRE 600 23S ribosomal RNA. Comparison with models of secondary structure for maize chloroplast 23S rRNA and for large portions of mouse and human 16S mitochondrial rRNAs. *Nucleic Acids Res.* 9, 4303–4324. doi:10.1093/nar/9.17.4303

Burge, C. B., Tuschl, T., and Sharp, P. A. (1999). "Splicing of precursors to mRNAs by the spliceosomes," in *The RNA World*, 2nd Edn (Cold Spring Harbor Laboratory Press), 525–560.

Bykhovskaya, Y., Casas, K., Mengesha, E., Inbal, A., and Fischel-Ghodsian, N. (2004). Missense mutation in pseudouridine synthase 1 (PUS1) causes mitochondrial myopathy and sideroblastic anemia (MLASA). *Am. J. Hum. Genet.* 74, 1303–1308. doi:10.1086/421530

Carlile, T. M., Rojas-Duran, M. F., Zinshteyn, B., Shin, H., Bartoli, K. M., and Gilbert, W. V. (2014). Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. *Nature* 515, 143–146. doi:10.1038/nature13802

Charette, M., and Gray, M. W. (2000). Pseudouridine in RNA: what, where, how, and why. *IUBMB Life* 49, 341–351. doi:10.1080/152165400410182

Cohn, W. E. V., and Elliot, V. (1951). Nucleoside-5'-phosphates from ribonucleic acid. *Nature* 167, 483–484. doi:10.1038/167483a0

Deeks, S. G. (2012). HIV: shock and kill. *Nature* 487, 439–440. doi:10.1038/487439a

Del Campo, M., Ofengand, J., and Malhotra, A. (2004). Crystal structure of the catalytic domain of RluD, the only rRNA pseudouridine synthase required for normal growth of *Escherichia coli*. *RNA* 10, 231–239. doi:10.1261/rna.5187404

Ericsson, U. B., Nordlund, P., and Hallberg, B. M. (2004). X-ray structure of tRNA pseudouridine synthase TruD reveals an inserted domain with a novel fold. *FEBS Lett.* 565, 59–64. doi:10.1016/j.febslet.2004.03.085

Fernandez-Vizarra, E., Berardinelli, A., Valente, L., Tiranti, V., and Zeviani, M. (2007). Nonsense mutation in pseudouridylate synthase 1 (PUS1) in two brothers affected by myopathy, lactic acidosis and sideroblastic anaemia (MLASA). *J. Med. Genet.* 44, 173–180. doi:10.1136/jmg.2006.045252

Foster, P. G., Huang, L., Santi, D. V., and Stroud, R. M. (2000). The structural basis for tRNA recognition and pseudouridine formation by pseudouridine synthase I. *Nat. Struct. Biol.* 7, 23–27. doi:10.1038/71219

Freund, M., Asang, C., Kammler, S., Konermann, C., Krummheuer, J., Hipp, M., et al. (2003). A novel approach to describe a U1 snRNA binding site. *Nucleic Acids Res.* 31, 6963–6975. doi:10.1093/nar/gkg901

Ghosh, S. K., Patton, J. R., and Spanjaard, R. A. (2012). A small RNA derived from RNA coactivator SRA blocks steroid receptor signaling *via* inhibition of PUS1p-mediated pseudouridylation of SRA: evidence of a novel RNA binding domain in the N-terminus of steroid receptors. *Biochemistry* 51, 8163–8172. doi:10.1021/bi300602r

Girard, J. P., Lehtonen, H., Caizergues-Ferrer, M., Amalric, F., Tollervey, D., and Lapeyre, B. (1992). GAR1 is an essential small nucleolar RNP protein required for pre-rRNA processing in yeast. *EMBO J.* 11, 673–682.

Green, M. R. (1986). Pre-mRNA splicing. *Annu. Rev. Genet.* 20, 671–708. doi:10.1146/annurev.ge.20.120186.003323

Grosjean, H., Sprinzl, M., and Steinberg, S. (1995). Posttranscriptionally modified nucleosides in transfer RNA: their locations and frequencies. *Biochimie* 77, 139–141. doi:10.1016/0300-9084(96)88117-X

Guenther, M. G., Levine, S. S., Boyer, L. A., Jaenisch, R., and Young, R. A. (2007). A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* 130, 77–88. doi:10.1016/j.cell.2007.05.042

Hamma, T., and Ferre-D'Amare, A. R. (2006). Pseudouridine synthases. *Chem. Biol.* 13, 1125–1135. doi:10.1016/j.chembiol.2006.09.009

Hamma, T., Reichow, S. L., Varani, G., and Ferre-D'Amare, A. R. (2005). The Cbf5-Nop10 complex is a molecular bracket that organizes box H/ACA RNPs. *Nat. Struct. Mol. Biol.* 12, 1101–1107. doi:10.1038/nsmb1036

Hoang, C., Chen, J., Vizthum, C. A., Kandel, J. M., Hamilton, C. S., Mueller, E. G., et al. (2006). Crystal structure of pseudouridine synthase RluA: indirect sequence readout through protein-induced RNA structure. *Mol. Cell* 24, 535–545. doi:10.1016/j.molcel.2006.09.017

Hoang, C., and Ferre-D'Amare, A. R. (2001). Cocrystal structure of a tRNA Psi55 pseudouridine synthase: nucleotide flipping by an RNA-modifying enzyme. *Cell* 107, 929–939. doi:10.1016/S0092-8674(01)00618-3

Hoang, C., and Ferre-D'Amare, A. R. (2004). Crystal structure of the highly divergent pseudouridine synthase TruD reveals a circular permutation of a conserved fold. *RNA* 10, 1026–1033. doi:10.1261/rna.7240504

Hopper, A. K., and Phizicky, E. M. (2003). tRNA transfers to the limelight. *Genes Dev.* 17, 162–180. doi:10.1101/gad.1049103

Kadonaga, J. T. (2004). Regulation of RNA polymerase II transcription by sequence-specific DNA binding factors. *Cell* 116, 247–257. doi:10.1016/S0092-8674(03)01078-X

Karijolich, J., Kantartzis, A., and Yu, Y. T. (2010). RNA modifications: a mechanism that modulates gene expression. *Methods Mol. Biol.* 629, 1–19. doi:10.1007/978-1-60761-657-3_1

Karijolich, J., and Yu, Y. T. (2010). Spliceosomal snRNA modifications and their function. *RNA Biol.* 7, 192–204. doi:10.4161/rna.7.2.11207

Kaya, Y., Del Campo, M., Ofengand, J., and Malhotra, A. (2004). Crystal structure of TruD, a novel pseudouridine synthase with a new protein fold. *J. Biol. Chem.* 279, 18107–18110. doi:10.1074/jbc.C400072200

Kishore, S., Gruber, A. R., Jedlinski, D. J., Syed, A. P., Jorjani, H., and Zavolan, M. (2013). Insights into snoRNA biogenesis and processing from PAR-CLIP of snoRNA core proteins and small RNA sequencing. *Genome Biol.* 14, R45. doi:10.1186/gb-2013-14-5-r45

Kiss, T., Fayet-Lebaron, E., and Jady, B. E. (2010). Box H/ACA small ribonucleoproteins. *Mol. Cell* 37, 597–606. doi:10.1016/j.molcel.2010.01.032

Kramer, A., Keller, W., Appel, B., and Luhrmann, R. (1984). The 5' terminus of the RNA moiety of U1 small nuclear ribonucleoprotein particles is required for the splicing of messenger RNA precursors. *Cell* 38, 299–307. doi:10.1016/0092-8674(84)90551-8

Kuras, L., and Struhl, K. (1999). Binding of TBP to promoters *in vivo* is stimulated by activators and requires Pol II holoenzyme. *Nature* 399, 609–613. doi:10.1038/21239

Lanz, R. B., McKenna, N. J., Onate, S. A., Albrecht, U., Wong, J., Tsai, S. Y., et al. (1999). A steroid receptor coactivator, SRA, functions as an RNA and is present in an SRC-1 complex. *Cell* 97, 17–27. doi:10.1016/S0092-8674(00)80711-4

Lanz, R. B., Razani, B., Goldberg, A. D., and O'Malley, B. W. (2002). Distinct RNA motifs are important for coactivation of steroid hormone receptors by steroid receptor RNA activator (SRA). *Proc. Natl. Acad. Sci. U.S.A.* 99, 16081–16086. doi:10.1073/pnas.192571399

Lesser, C. F., and Guthrie, C. (1993). Mutations in U6 snRNA that alter splice site specificity: implications for the active site. *Science* 262, 1982–1988. doi:10.1126/science.8266093

Levine, M. (2011). Paused RNA polymerase II as a developmental checkpoint. *Cell* 145, 502–511. doi:10.1016/j.cell.2011.04.021

Li, L., and Ye, K. (2006). Crystal structure of an H/ACA box ribonucleoprotein particle. *Nature* 443, 302–307. doi:10.1038/nature05151

Li, S., Duan, J., Li, D., Yang, B., Dong, M., and Ye, K. (2011). Reconstitution and structural analysis of the yeast box H/ACA RNA-guided pseudouridine synthase. *Genes Dev.* 25, 2409–2421. doi:10.1101/gad.175299.111

Li, X., Zhu, P., Ma, S., Song, J., Bai, J., Sun, F., et al. (2015). Chemical pulldown reveals dynamic pseudouridylation of the mammalian transcriptome. *Nat. Chem. Biol.* 11, 592–597. doi:10.1038/nchembio.1836

Lovejoy, A. F., Riordan, D. P., and Brown, P. O. (2014). Transcriptome-wide mapping of pseudouridines: pseudouridine synthases modify specific mRNAs in *S. cerevisiae*. *PLoS ONE* 9:e110799. doi:10.1371/journal.pone.0110799

Ma, X., Zhao, X., and Yu, Y. T. (2003). Pseudouridylation (Psi) of U2 snRNA in *S. cerevisiae* is catalyzed by an RNA-independent mechanism. *EMBO J.* 22, 1889–1897. doi:10.1093/emboj/cdg191

Machnicka, M. A., Milanowska, K., Osman Oglou, O., Purta, E., Kurkowska, M., Olchowik, A., et al. (2013). MODOMICS: a database of RNA modification pathways—2013 update. *Nucleic Acids Res.* 41, D262–D267. doi:10.1093/nar/gks1007

Maden, B. E. (1990). The numerous modified nucleotides in eukaryotic ribosomal RNA. *Prog. Nucleic Acid Res. Mol. Biol.* 39, 241–303. doi:10.1016/S0079-6603(08)60629-7

Manival, X., Charron, C., Fourmann, J. B., Godard, F., Charpentier, B., and Branlant, C. (2006). 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, 826–839. doi:10.1093/nar/gkj482

Massenet, S., and Branlant, C. (1999). A limited number of pseudouridine residues in the human atac spliceosomal UsnRNAs as compared to human major spliceosomal UsnRNAs. *RNA* 5, 1495–1503. doi:10.1017/S1355838299991537

Massenet, S., Mougin, A., and Branlant, C. (1998). "Posttranscriptional modifications in the U small nuclear RNAs," in *Modification and Editing of RNA*, ed. H. Grosjean (Washington, DC: ASM Press), 201–228.

Matera, A. G., and Wang, Z. (2014). A day in the life of the spliceosome. *Nat. Rev. Mol. Cell Biol.* 15, 108–121. doi:10.1038/nrm3742

McCleverty, C. J., Hornsby, M., Spraggon, G., and Kreusch, A. (2007). Crystal structure of human PUS10, a novel pseudouridine synthase. *J. Mol. Biol.* 373, 1243–1254. doi:10.1016/j.jmb.2007.08.053

Mercer, T. R., Dinger, M. E., and Mattick, J. S. (2009). Long non-coding RNAs: insights into functions. *Nat. Rev. Genet.* 10, 155–159. doi:10.1038/nrg2521

Michaud, S., and Reed, R. (1991). An ATP-independent complex commits pre-mRNA to the mammalian spliceosome assembly pathway. *Genes Dev.* 5, 2534–2546. doi:10.1101/gad.5.12b.2534

Mizutani, K., Machida, Y., Unzai, S., Park, S. Y., and Tame, J. R. (2004). Crystal structures of the catalytic domains of pseudouridine synthases RluC and RluD from *Escherichia coli*. *Biochemistry* 43, 4454–4463. doi:10.1021/bi036079c

Muse, G. W., Gilchrist, D. A., Nechaev, S., Shah, R., Parker, J. S., Grissom, S. F., et al. (2007). RNA polymerase is poised for activation across the genome. *Nat. Genet.* 39, 1507–1511. doi:10.1038/ng.2007.21

Narlikar, G. J., Fan, H. Y., and Kingston, R. E. (2002). Cooperation between complexes that regulate chromatin structure and transcription. *Cell* 108, 475–487. doi:10.1016/S0092-8674(02)00654-2

Newby, M. I., and Greenbaum, N. L. (2002). Sculpting of the spliceosomal branch site recognition motif by a conserved pseudouridine. *Nat. Struct. Biol.* 9, 958–965. doi:10.1038/nsb873

Newman, A. J. (1994). Pre-mRNA splicing. *Curr. Opin. Genet. Dev.* 4, 298–304. doi:10.1016/S0959-437X(05)80057-7

Patton, J. R. (1993a). Multiple pseudouridine synthase activities for small nuclear RNAs. *Biochem. J.* 290(Pt 2), 595–600. doi:10.1042/bj2900595

Patton, J. R. (1993b). Ribonucleoprotein particle assembly and modification of U2 small nuclear RNA containing 5-fluorouridine. *Biochemistry* 32, 8939–8944. doi:10.1021/bi00085a027

Ptashne, M. (2005). Regulation of transcription: from lambda to eukaryotes. *Trends Biochem. Sci.* 30, 275–279. doi:10.1016/j.tibs.2005.04.003

Rashid, R., Liang, B., Baker, D. L., Youssef, O. A., He, Y., Phipps, K., et al. (2006). Crystal structure of a Cbf5-Nop10-Gar1 complex and implications in RNA-guided pseudouridylation and dyskeratosis congenita. *Mol. Cell* 21, 249–260. doi:10.1016/j.molcel.2005.11.017

Reddy, R., and Busch, H. (1988). "Small nuclear RNAs: RNA sequences, structure, and modifications," in *Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles*, ed. M. Birnsteil (Heidelberg: Springer-Verlag Press), 1–37.

Richman, D. D., Margolis, D. M., Delaney, M., Greene, W. C., Hazuda, D., and Pomerantz, R. J. (2009). The challenge of finding a cure for HIV infection. *Science* 323, 1304–1307. doi:10.1126/science.1165706

Roca, X., Sachidanandam, R., and Krainer, A. R. (2005). Determinants of the inherent strength of human 5' splice sites. *RNA* 11, 683–698. doi:10.1261/rna.2040605

Ruby, S. W., and Abelson, J. (1988). An early hierarchic role of U1 small nuclear ribonucleoprotein in spliceosome assembly. *Science* 242, 1028–1035. doi:10.1126/science.2973660

Sawa, H., and Abelson, J. (1992). Evidence for a base-pairing interaction between U6 small nuclear RNA and 5' splice site during the splicing reaction in yeast. *Proc. Natl. Acad. Sci. U.S.A.* 89, 11269–11273. doi:10.1073/pnas.89.23.11269

Schwartz, S., Bernstein, D. A., Mumbach, M. R., Jovanovic, M., Herbst, R. H., Leon-Ricardo, B. X., et al. (2014). Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA. *Cell* 159, 148–162. doi:10.1016/j.cell.2014.08.028

Seraphin, B., and Rosbash, M. (1989). Identification of functional U1 snRNA-pre-mRNA complexes committed to spliceosome assembly and splicing. *Cell* 59, 349–358. doi:10.1016/0092-8674(89)90296-1

Sivaraman, J., Iannuzzi, P., Cygler, M., and Matte, A. (2004). 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, 87–101. doi:10.1016/j.jmb.2003.10.003

Sivaraman, J., Sauve, V., Larocque, R., Stura, E. A., Schrag, J. D., Cygler, M., et al. (2002). Structure of the 16S rRNA pseudouridine synthase RsuA bound to uracil and UMP. *Nat. Struct. Biol.* 9, 353–358. doi:10.1038/nsb788

Smale, S. T., and Kadonaga, J. T. (2003). The RNA polymerase II core promoter. *Annu. Rev. Biochem.* 72, 449–479. doi:10.1146/annurev.biochem.72.121801.161520

Wahl, M. C., Will, C. L., and Luhrmann, R. (2009). The spliceosome: design principles of a dynamic RNP machine. *Cell* 136, 701–718. doi:10.1016/j.cell.2009.02.009

Wassarman, D. A., and Steitz, J. A. (1991). Structural analyses of the 7SK ribonucleoprotein (RNP), the most abundant human small RNP of unknown function. *Mol. Cell. Biol.* 11, 3432–3445. doi:10.1128/MCB.11.7.3432

Wassarman, D. A., and Steitz, J. A. (1992). Interactions of small nuclear RNA's with precursor messenger RNA during *in vitro* splicing. *Science* 257, 1918–1925. doi:10.1126/science.1411506

Wu, G., Adachi, H., Ge, J., Stephenson, D., Query, C. C., and Yu, Y. T. (2016a). Pseudouridines in U2 snRNA stimulate the ATPase activity of Prp5 during spliceosome assembly. *EMBO J.* 35, 654–667. doi:10.15252/embj.201593113

Wu, G., Radwan, M. K., Xiao, M., Adachi, H., Fan, J., and Yu, Y. T. (2016b). The TOR signaling pathway regulates starvation-induced pseudouridylation of yeast U2 snRNA. *RNA* 22, 1146–1152. doi:10.1261/rna.056796.116

Wu, G., Xiao, M., Yang, C., and Yu, Y. T. (2011). U2 snRNA is inducibly pseudouridylated at novel sites by PUS7p and snR81 RNP. *EMBO J.* 30, 79–89. doi:10.1038/emboj.2010.316

Xu, Y. Z., and Query, C. C. (2007). Competition between the ATPase Prp5 and branch region-U2 snRNA pairing modulates the fidelity of spliceosome assembly. *Mol. Cell* 28, 838–849. doi:10.1016/j.molcel.2007.09.022

Yang, Z., Zhu, Q., Luo, K., and Zhou, Q. (2001). The 7SK small nuclear RNA inhibits the CDK9/cyclin T1 kinase to control transcription. *Nature* 414, 317–322. doi:10.1038/35104575

Yik, J. H., Chen, R., Nishimura, R., Jennings, J. L., Link, A. J., and Zhou, Q. (2003). Inhibition of P-TEFb (CDK9/Cyclin T) kinase and RNA polymerase II transcription by the coordinated actions of HEXIM1 and 7SK snRNA. *Mol. Cell* 12, 971–982. doi:10.1016/S1097-2765(03)00388-5

Yu, Y. T., and Meier, U. T. (2014). RNA-guided isomerization of uridine to pseudouridine—pseudouridylation. *RNA Biol.* 11, 1483–1494. doi:10.4161/15476286.2014.972855

Yu, Y. T., Shu, M. D., and Steitz, J. A. (1998). Modifications of U2 snRNA are required for snRNP assembly and pre-mRNA splicing. *EMBO J.* 17, 5783–5795. doi:10.1093/emboj/17.19.5783

Zhao, X., Patton, J. R., Davis, S. L., Florence, B., Ames, S. J., and Spanjaard, R. A. (2004). Regulation of nuclear receptor activity by a pseudouridine synthase through posttranscriptional modification of steroid receptor RNA activator. *Mol. Cell* 15, 549–558. doi:10.1016/j.molcel.2004.06.044

Zhao, X., Patton, J. R., Ghosh, S. K., Fischel-Ghodsian, N., Shen, L., and Spanjaard, R. A. (2007). PUS3p- and PUS1p-dependent pseudouridylation of steroid receptor RNA activator controls a functional switch that regulates nuclear receptor signaling. *Mol. Endocrinol.* 21, 686–699. doi:10.1210/me.2006-0414

Zhao, X., and Yu, Y. T. (2007). Incorporation of 5-fluorouracil into U2 snRNA blocks pseudouridylation and pre-mRNA splicing *in vivo*. *Nucleic Acids Res.* 35, 550–558. doi:10.1093/nar/gkl1084

Zhao, Y., Karijolich, J., Glaunsinger, B., and Zhou, Q. (2016). Pseudouridylation of 7SK snRNA promotes 7SK snRNP formation to suppress HIV-1 transcription and escape from latency. *EMBO Rep.* 17, 1441–1451. doi:10.15252/embr.201642682

Zhou, Q., Li, T., and Price, D. H. (2012). RNA polymerase II elongation control. *Annu. Rev. Biochem.* 81, 119–143. doi:10.1146/annurev-biochem-052610-095910

Zhuang, Y., and Weiner, A. M. (1989). A compensatory base change in human U2 snRNA can suppress a branch site mutation. *Genes Dev.* 3, 1545–1552. doi:10.1101/gad.3.10.1545

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.