

Pseudouridine in a new era of RNA modifications

Cell Research (2015) 25:153-154. doi:10.1038/cr.2014.143; published online 4 November 2014

Two articles recently published in *Nature* and *Cell* report the first transcriptome-wide maps of pseudouridine (Ψ) at single-base resolution through selective chemical labeling, suggesting new mechanisms and functions of Ψ in mRNA and non-coding RNA molecules.

Pseudouridine (Ψ) was among the first post-transcriptional modifications discovered and is overall one of the most abundant [1]. It is present in a wide range of cellular RNAs and is highly conserved across species. Despite the extensive research of Ψ in tRNA and rRNA [2], its presence and potential function in mRNA are largely unexplored due to the lack of effective methods. In two recent papers published in *Nature* and *Cell*, Carlile *et al.* [3] and Schwartz *et al.* [4] performed transcriptome-wide mapping of Ψ using deep sequencing, pushing us one step further to the understanding of the molecular and evolutionary prevalence of Ψ in eukaryotic cells.

Ψ is derived from uridine (U) via base-specific isomerization catalyzed by Ψ synthases. The site-specific pseudouridylation goes through either snoRNA-dependent (requires H/ACA RNP) or -independent mechanism (requires pseudouridine synthase (PUS) family enzymes) [2]. It has an extra hydrogen-bond donor at its non-Watson-Crick edge. When incorporated into RNA, Ψ can alter RNA secondary structure by increasing base stacking, improving base pairing and rigidifying sugar-phosphate backbone [5]. The chemical and physical properties of RNA can be altered with the incorporation of Ψ , which could contribute to subsequent cellular functions. It was recently dis-

covered that pseudouridylation could be induced by stress [2]. The replacement of multiple U sites with Ψ in synthetic RNA molecules results in an increased protein expression level [6], while artificially incorporated Ψ in mRNAs mediates nonsense-to-sense codon conversion (recoding) by facilitating unusual base pairing in the ribosome decoding center, thus demonstrating a new means of generating protein diversity [7].

With the recent advances in the studies of dynamic RNA modifications in post-transcriptional gene expression regulation, pseudouridine comes back to the forefront. RNA modifications have been previously thought to be static, discrete, and utilized to fine-tune RNA structure and function. Yet the emerging studies showed that N^6 -methyladenosine (m^6A) in mRNA and certain non-coding RNA (ncRNA) is reversible [8]. This abundant and dynamic mRNA modification codes additional regulatory information on top of the primary sequence [9-11]. However, unlike m^6A which is reversible, the conversion from U to Ψ is almost certainly irreversible as the C-N glycosidic bond of U is isomerized to a much more inert C-C bond in Ψ (Figure 1A). The irreversibility of this modification suggests distinct roles of Ψ in response to stimuli or stresses.

In the recent two papers, Carlile *et al.* [3] and Schwartz *et al.* [4] each reported a transcriptome-wide sequencing approach for mapping Ψ at single-base resolution. They both took advantage of the known reaction between Ψ and *N*-cyclohexyl-*N'*-(2-morpholinoethyl)carbodiimide metho-*p*-toluenesulphonate (CMC) to form N3-CMC- Ψ (Figure 1B). CMC is known to selectively

modify Ψ and the formed N3-CMC- Ψ is known to block reverse transcription [12]. In the work of Carlile *et al.* [3], researchers developed Pseudo-seq by labeling Ψ in uniformly fragmented RNA by CMC, followed by reverse transcription and size selection for truncated cDNAs, which correspond to sequences from 3' end to one nucleotide downstream of the modified Ψ sites. Subsequent circularization, amplification and deep sequencing revealed Ψ sites in both mRNA and ncRNA molecules. About 260 Ψ sites in 238 protein-coding transcripts in yeast and 96 Ψ sites in 89 mRNAs in human were identified. They also investigated pseudouridylation under various growth states and found that a subset of Ψ sites in mRNA and ncRNA are differentially modified, suggesting a pseudouridylation response to environmental cues. Through genetic perturbation of PUS genes, they further revealed that PUS1, PUS2, PUS4 and PUS7 are involved in mRNA pseudouridylation. Schwartz *et al.* [4] used a similar strategy to develop Ψ -seq, but with a focus on quantitative measurement. Instead of performing size selection, they relied on computational method to minimize background and determine the high-confidence Ψ sites, while the addition of synthetic spike-in probes facilitates the quantification of the relative stoichiometry. Similarly, they reported 328 unique Ψ sites in yeast mRNAs and ncRNAs. 108 of these sites were found to be associated with their writer PUS and/or snoRNA through genetic perturbation experiments. They also determined the consensus sequence of each Ψ site recognized by cognate PUS, and showed the changes of pseudouridylation patterns under different

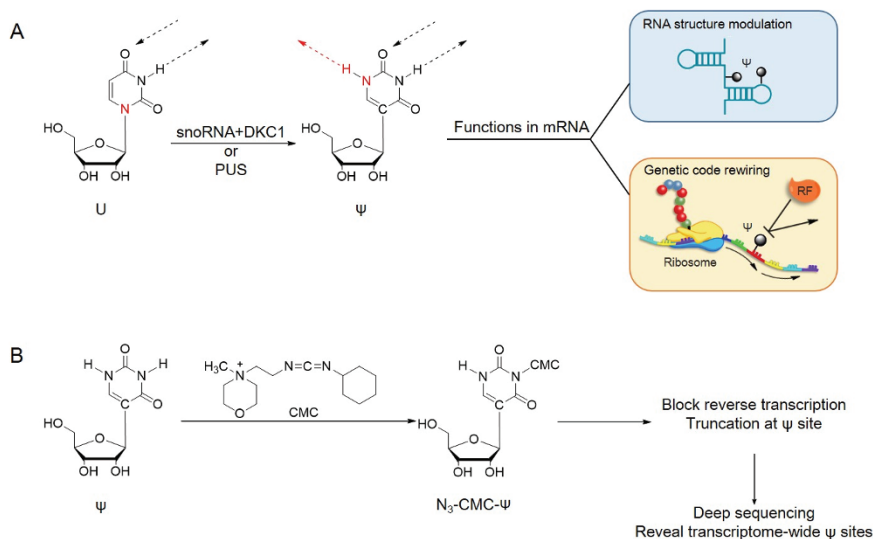


Figure 1 (A) The installation and functional roles of Ψ . Ψ is installed through either snoRNA-dependent (requires H/ACA RNP) or -independent mechanism (requires PUS family enzymes). Ψ can either change the RNA secondary structure or facilitate recoding during translation. Red codon: stop codon. RF: releasing factor. (B) A CMC labeling-based Ψ -sequencing strategy at single-base resolution. CMC selectively reacts with Ψ to form N₃-CMC- Ψ , which blocks reverse transcription, thus enabling the identification of transcriptome-wide Ψ positions through deep sequencing.

be enriched using specific antibodies or through chemical means. Besides reverse transcription, other enzymatic reactions could also be employed to achieve more sensitive detection.

Collectively, the papers by Carlile *et al.* [3] and Schwartz *et al.* [4] provide the first global detections of Ψ , and suggest potential regulatory functions of this modification in mRNA. Future research will focus on functional roles of Ψ , either through changing the RNA structure or recoding during translation. Identification and characterization of potential readers could be critical for further understanding of the biological functions of RNA pseudouridylation.

Boxuan Simen Zhao¹, Chuan He¹

¹Department of Chemistry and Institute for Biophysical Dynamics, Howard Hughes Medical Institute, 929 East 57th Street, The University of Chicago, Chicago, IL, 60637, USA

Correspondence: Chuan He

E-mail: chuanhe@uchicago.edu

References

- Cohn WE, Volkin W. *Nature* 1951; **167**:483-484.
- Ge J, Yu YT. *Trends Biochem Sci* 2013; **38**:210-218.
- Carlile TM, Rojas-Duran MF, Zinshteyn B, *et al.* *Nature* 2014; **515**:143-146.
- Schwartz S, Bernstein DA, Mumbach MR, *et al.* *Cell* 2014; **159**:148-162.
- Charette M, Gray MW. *IUBMB Life* 2000; **49**:341-351.
- Karikó K, Muramatsu H, Keller JM, *et al.* *Mol Ther* 2012; **20**:948-53.
- Karijohil J, Yu YT. *Nature* 2011; **474**:395-398.
- Jia G, Fu Y, Zhao X, *et al.* *Nat Chem Biol* 2011; **7**:885-887.
- Dominissini D, Moshitch-Moshkovitz S, Schwartz S, *et al.* *Nature* 2012; **485**:201-206.
- Meyer KD, Saletore Y, Zumbo P, *et al.* *Cell* 2012; **149**:1635-1646.
- Wang X, Lu Z, Gomez A, *et al.* *Nature* 2014; **505**:117-120.
- Ofengand J, Del Campo M, Kaya Y. *Methods* 2001; **25**:365-373.

growth conditions. The researchers convincingly demonstrated the dramatic induction of pseudouridylation in both mRNA and ncRNA upon heat shock in yeast, and suggested a mechanism involving relocalization of one of the pseudouridine synthases, Pus7p, upon heat shock stimuli. By performing Ψ -seq on patient samples with dyskeratosis congenita, a congenital disorder with mutations in the protein subunits (such as dyskerin) of box H/ACA ribonucleoprotein (RNP), they reported a subtle yet noticeable decrease in the Ψ level in rRNAs and at one highly conserved site on telomerase RNA component (TERC) in the patient cells; the TERC Ψ could be important for the stabilization of TERC. This result suggests a potential molecular mechanism of the disease whereby mutations of *DKC1*/dyskerin disrupt its functions as a protein subunit of H/ACA RNP and as a PUS itself.

These two works suggest new roles for RNA pseudouridylation: this modification not only exists ubiquitously in

both mRNAs and various ncRNAs but also is dynamically regulated by environment cues, adding a potential new mechanism for the regulation of mRNA fate. As RNA secondary structure is closely related to multiple aspects of mRNA metabolism and function, it is possible that pseudouridylation induces structural changes and thus affects functions. Besides its function in mediating nonsense-to-sense codon conversion by ribosomes [7], pseudouridylation could also introduce post-transcriptional genetic recoding, thus diversifying the proteome. These studies have also suggested connections between RNA pseudouridylation and multiple human diseases. A robust tool to evaluate global Ψ dynamics and changes can contribute to revealing the underlying mechanisms.

Both studies take advantage of the CMC-based modification chemistry that is selective to Ψ [12]; however, the modified CMC- Ψ could serve as a target for developing more effective sequencing methods. The modified CMC- Ψ could