Discovery of cap-specific adenosine-N^6-methyltransferase (CAPAM)

RNA molecules are enzymatically modified after transcription. To date, about 150 types of chemical modifications have been identified in various RNAs from all domains of life. N^6-methyladenosine (m^6A), one of the most abundant modifications in eukaryotic mRNA, plays critical roles in RNA metabolism and function. Recent advances in deep sequencing technology enable us to map m^6A in a transcriptome-wide manner, highlighting the biological importance of this modification and establishing the concept of "epitranscriptomics" [1]. Most m^6A sites in internal regions of mRNAs are introduced by an m^6A writer, METTL3 methyltransferase, working in conjunction with accessory proteins. m^6A is a reversible modification that can be removed by the eraser proteins ALKBH5 and FTO. m^6A is recognized by several reader proteins, including the YTH family proteins, which have diverse influences on their target mRNAs, affecting processes such as translational regulation and mRNA decay. m^6A is also present at the transcription start site of mRNAs.

The 7-methylguanosine (m^7G) cap, a characteristic 5′-terminal structure of eukaryotic mRNAs, plays extensive roles in mRNA processing, stability, nuclear export, and translation. In addition to the m^7G cap structure, the 5′-terminal regions of the mRNAs of higher eukaryotes are frequently methylated [2]. The first and second nucleotides of mRNAs are 2′-*O*-methylated by CMTR1 and CMTR2, respectively. These modifications serve as molecular markers that distinguish between self and non-self mRNAs by antagonizing RNA sensors, allowing escape from the innate immune system. In vertebrate mRNAs, if the 5′ terminal nucleoside is adenosine, the N^6 position of the A is methylated by an unidentified enzyme to form N^6,2′-*O*-dimethyladenosine (m^6Am) (Fig. 1(a)). Recent studies indicated that m^6Am is involved in mRNA stability and metabolism. However, the biogenesis and functional role of this modification remained elusive.

To identify the enzyme responsible for N^6 methylation of m^6Am, we first established a highly sensitive method for analyzing the 5′-terminal chemical structures of capped mRNAs using RNA mass spectrometry (RNA-MS). We then analyzed the modification status of the 5' cap structure of mRNAs obtained from several strains harboring knockouts of uncharacterized methyltransferases that are conserved in vertebrates but not in yeast. When we knocked out the candidate gene *PCIF1*, m^6Am disappeared completely and was replaced by Am at the first nucleotide, indicating that PCIF1 is responsible for m^6Am formation. We then confirmed that recombinant PCIF1 could form m^6Am on the capped mRNA *in vitro* in the presence of *S*-adenosylmethionine (SAM), demonstrating that PCIF1 is a cap-specific adenosine-N^6-methyltransferase; hence, we renamed the protein CAPAM [3]. Biochemical characterization revealed that CAPAM methylates mRNA containing $m^7GpppAm$ more efficiently than mRNA containing m^7GpppA.

CAPAM/PCIF1 was originally identified as a protein that interacts with the phosphorylated CTD of RNA polymerase II (RNAPII) via its N-terminal WW domain [4]. We also showed that the WW domain of CAPAM specifically interacts with Ser5-phosphorylated CTD, indicating that CAPAM is recruited to the early elongation complex of RNAPII and co-transcriptionally introduces N^6-methylation in nascent transcripts. Our results implied hierarchical formation of

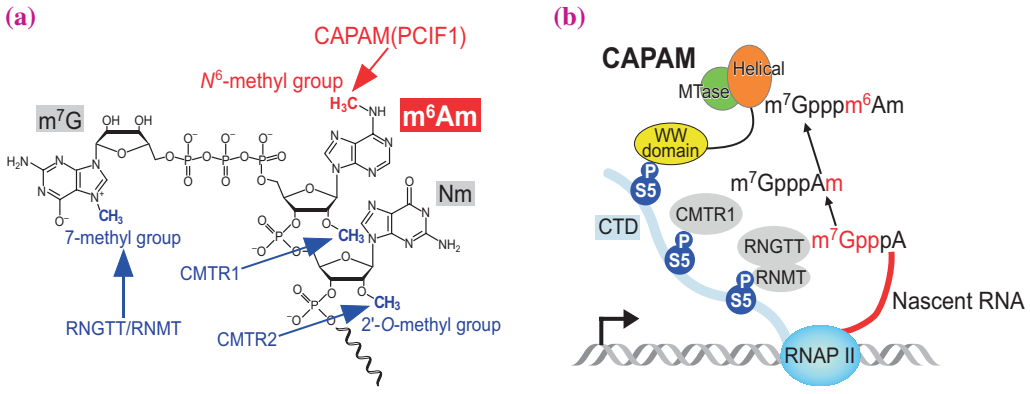

Fig. 1. **(a)** Chemical structure of the 5'-terminus of vertebrate mRNA bearing adenosine at the transcription start site. **(b)** Hierarchical formation of m^7Gpppm^6Am in the early elongation complex of RNA polymerase II.

m^7Gpppm^6Am–pppA, GpppA, m^7GpppA, $m^7GpppAm$, and m^7Gpppm^6Am (Fig. 1(b)).

In its C-terminal region, CAPAM has a putative methyltransferase domain that has a topology similar to that of DNA m^6A methyltransferase [5]. To elucidate the molecular mechanism of N^6-methylation, we solved the crystal structures of human and zebrafish CAPAMs with or without ligands. Crystals of CAPAMs in complex with m^7G cap analog and *S*-adenosylhomocysteine (SAH) were obtained by the soaking method. X-ray diffraction data were collected at 100 K at SPring-8 **BL41XU** and PXI at the Swiss Light Source. We determined the crystal structures of CAPAMs at 1.8–2.9 Å resolution. The core region of CAPAM consists of the methyltransferase (MTase) and helical domains (Fig. 2(a)). The MTase domain adopts a canonical Rossmann fold that contains the conserved catalytic NPPF motif. The helical domain contains three-helix bundles, a four-helix bundle, and β sheets, and exhibits no structural similarity with any known protein. SAH is bound to a catalytic site of the MTase domain in a manner similar to what has been observed in class I MTases. The m^7G cap is bound to a pocket formed by the MTase and helical domains, and interacts tightly with the surrounding residues. In mutation studies, we confirmed the importance of these residues for N^6-methylation. The target adenosine residue adjacent to the m^7G cap is not visible in the electron density map. Hence, we modeled the target Am at the catalytic site of the MTase domain based on the structure of M.TaqI DNA m^6A methyltransferase (Fig. 2(b)). Our model suggested that the adenine of Am is recognized by the NPPF motif, while the ribose moiety of Am interacts with the histidine residue, contributing to the strong interaction with the $m^7GpppAm$ cap structure. Notably, the helical domain forms a positively charged groove, suggesting that it serves as a binding surface for the RNA chain (Fig. 2(c)).

Ribosome profiling and RNA-seq experiments suggested that CAPAM-mediated m^6Am formation promotes translation of mRNAs starting with A, rather than mRNA stabilization [3]. The molecular mechanism underlying m^6Am-mediated translational regulation remains to be determined. Regarding the physiological role of CAPAM, we observed a significant growth defect in *CAPAM*-knockout cells under oxidative stress, suggesting that CAPAM plays a key role in gene expression in response to such conditions. Given that *CAPAM* has been reported as a putative tumor suppressor gene in bladder cancer cells, CAPAM-mediated m^6Am modification might have a substantial impact on physiology and pathology in mammals.

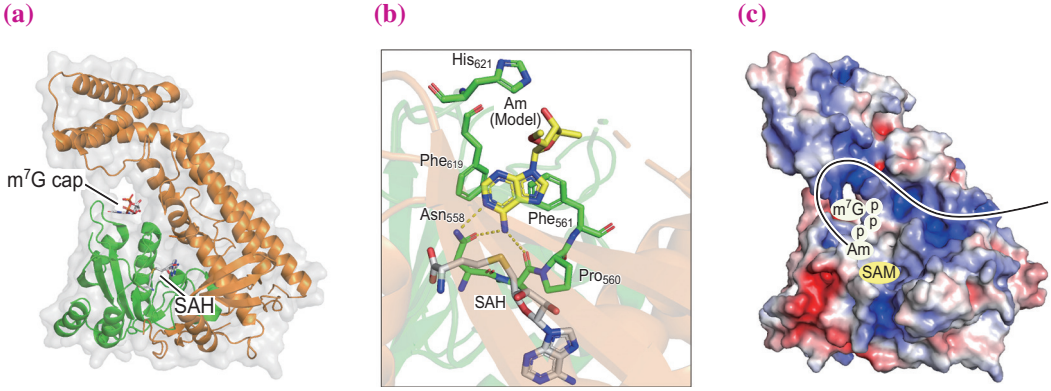

Fig. 2. **(a)** Crystal structure of zebrafish CAPAM in complex with m^7GpppA and SAH (PDB: 6IRZ). **(b)** Model structure of the catalytic site with the target Am residue. **(c)** Electrostatic surface potential of zebrafish CAPAM with a model RNA substrate.

S. Akichika[a], S. Hirano[b], O. Nureki[b] and T. Suzuki[a,*]

[a] Department of Chemistry and Biotechnology, The University of Tokyo
[b] Department of Biological Sciences, The University of Tokyo

*Email: ts@chembio.t.u-tokyo.ac.jp

References

[1] H. Shi *et al.*: Mol. Cell **74** (2019) 640.
[2] A.K. Banerjee: Microbiol. Rev. **44** (1980) 175.
[3] *S. Akichika, S. Hirano, Y. Shichino, Ta. Suzuki, H. Nishimasu, R. Ishitani, A. Sugita, Y. Hirose, S. Iwasaki, O. Nureki and Ts. Suzuki: Science* ***363*** *(2019) eaav0080.*
[4] H. Fan *et al.*: Biochem. Biophys. Res. Commun. **301** (2003) 378.
[5] L.M. Iyer *et al.*: BioEssays **38** (2016) 27.