外 国 語 要 旨

Structural bioinformatics analyses for quaternary structure and functional sites prediction of mRNA cleavage and polyadenylation specificity factor (CPSF)

Saki Aoto

During the oocyte maturation and early embryonic development in higher eukaryotes, gene expressions are controlled not only by transcription, but also by translation. In immature oocytes, transcription activity of gene is low, but transcripts stored in the cytosol are utilized for high protein expression. The stored mRNAs named maternal mRNAs are transcribed before meiosis and wait for the time that they should be translated in the oocyte cytoplasm. In the oocyte maturation process, new proteins are translated solely from the maternal mRNAs. This phenomenon is also observed in neurons.

Messenger RNAs generally undergo a number of maturation processes in nucleus, such as splicing, 5'-capping, and poly(A)-tail elongation. The matured mRNAs are then translated immediately after they are transported to cytoplasm. Maternal mRNAs are, however, transported to cytoplasm only with a short poly(A) tail and elongation of the poly(A) tail is suppressed until the timing of translation. The maternal mRNAs are known to be associated with several different proteins and form mRNPs. The complex molecule protects themselves from degradation and regulates the timing of poly(A)-tail elongation.

All maternal mRNAs have cytoplasmic polyadenylation element (CPE) and a highly conserved AAUAAA hexanucleotide sequence (HEX) in their 3’-UTR as the targets for proteins forming the mRNP. CPE is the binding site of CPE-binding protein (CPEB), one of the core components of the maternal mRNP that recruits poly(A) deadenylase and is involved in translation repression. CPEB also binds cleavage and polyadenylation specificity factor complex (CPSF). CPSF binds to HEX and plays a key role in translation activation. CPSF cleaves immature mRNA 3’-UTR prior to polyadenylation by Gld2, one of the poly(A) polymerases (PAPs). Phosphorylation of CPEB is considered to trigger a structural change in the maternal mRNP, which in turn changes the level of translation in maternal mRNAs. The mechanisms of translation regulation by maternal mRNPs have been studied focusing on the function of each component. However, integration of these functions for the understanding of the overall mechanisms of translation repression in maternal mRNPs has not been tried. For the integration, three-dimensional structure of maternal mRNPs complex is required. An effort has been made for determining the structure of maternal mRNPs, but it remains to be elucidated. To understand the mechanisms of translational regulation of the maternal mRNPs, I predicted a three-dimensional

structure of CPSF, the core component of maternal mRNPs, that consists of CPSF160, CPSF100, CPSF73 and CPSF30. Based on the predicted structure, the interactions of core components to maternal mRNPs and mRNA were predicted as well.

The 3D structures of the CPSF subunits were built by comparative modeling method. To assemble CPSF subunits, the in-house protein-interface prediction tool was improved and utilized. For predicting an assembly of three or more complex, the biggest problem was to figure out direct interactions (partners) among the subunits. To address this problem in CPSF, information from the molecular phylogenetic analyses of each CPSF subunit were used. IntS9, a close homolog of CPSF100, and IntS11, a close homolog of CPSF73 are known to make a direct interaction, hence CPSF100 and CPSF73 likely interact with each other directly. From this observation, a docking calculation between CPSF100 and CPSF73 was performed. The 3D structure of CPSF160-CPSF100-CPSF73 complex was then built by docking CPSF160 and CPSF100-CPSF73 heterodimer. CPSF160-CPSF100-CPSF73 complex turned out to have a U-shape and CPSF160 bound CPSF73 through CPSF100. The molecular function of CPSF100 has been unclear, but the association with CPSF100 was known to be required for efficient endonuclease activity of CPSF73. Based on the model, CPSF100 seems important for bringing CPSF73 close to CPSF160 that binds mRNAs. The prediction of the binding site of CPEB and Gld2 on CPSF resulted in finding both binding sites on the same domain of CPSF160, and this enabled direct interactions with each other. This modeling structure is consistent with the previous experiment showing that CPEB on CPSF bound to Gld2. Gld2 also binds to CPSF73, bridging the opening of the U-shape with its catalytic site facing away from CPSF. Several RNA-binding sites were predicted on CPSF including a one on CPSF160 located next to the binding site of CPEB. The RNA-binding site was about 80 Å apart from the catalytic site of CPSF73, and this distance matched the distance between HEX and the site cleaved by CPSF73. As a result, the catalytic site of Gld2 was located far from poly(A)-binding site on mRNA and this configuration was seemingly sustained by the interaction of CPEB. The binding pose of Gld2 prevents itself from cleaving the maternal mRNA, hence the translation of the maternal mRNAs should be suppressed. At the initiation of the translation of maternal mRNA, CPEB is reported to be phosphorylated. Once CPEB is phosphorylated, it is speculated that CPEB could not sustain all the interactions and Gld2 should be released from both CPEB and CPSF, then Gld2 recognized the 3'-terminus of the mRNAs and Gld2 should initiate polyadenylation.

In this study, the 3D structure of the core components of maternal mRNPs in translational repression state was built and predicted the bindings sites of Gld2 and CPEB. The structure of CPSF helped to build a hypothesis for the mechanism of mRNPs translational repression, and could be used to obtain the model that the binding pose of Gld2 is altered after phosphorylation of CPEB. The predicted CPSF complex structure will pave the way for the understanding of the mechanism of the translational regulation on maternal mRNPs.