Opening new doors for understanding eukaryotic RNA splicing

WANG Hong-Wei

School of Life Sciences, Tsinghua University, Beijing 100084, China

Received September 6, 2015; accepted September 10, 2015; published online September 15, 2015

Citation: Wang HW. Opening new doors for understanding eukaryotic RNA splicing. *Sci China Life Sci*, 2015, 58: 1171–1172, doi: 10.1007/s11427-015-4934-0

Since the central dogma of molecular biology was proposed by Francis Crick in 1956, many fundamental discoveries have been made to strengthen or expand the theory for gene expression. Nowadays, we know that genetic information flow in a cell is a far more complicated process than the original succinct but explicit version. We now know that besides DNA duplication, DNA to RNA transcription, RNA to protein translation, there are also RNA to DNA transcription (reverse transcription), RNA to RNA replication, and RNA splicing and processing between the transcription and translation steps. Modifications and other processes to regulate and fine-tune gene expression have also been discovered in most recent years. All of these events are executed by multiple supramolecular machineries such as DNA and RNA polymerases, ribosomes, and spliceosomes. These machineries are normally multi-functional complexes composed of dozens of protein and RNA molecules and their activities are regulated by many other co-factor proteins or RNA molecules.

In the past few decades, structural biology, as a tool to decipher the precise atomic coordinates within a molecule, has played essential roles in elucidating most of the critical steps in gene expression. The atomic models of DNA polymerases revealed how DNA is duplicated; the atomic model of RNA polymerase complex explained in great detail of how RNA is transcribed from its DNA gene; the structures of ribosome in its various states described vividly how proteins are translated from messenger RNAs. Not only were the structures of the major molecular machines involved in the aforementioned processes solved, but also their different

working states and regulatory co-factors have been revealed in great structural details. These structures brought major leaps for our understanding of central dogma. In contrast, a major step of the gene expression, RNA splicing, is still largely a jigsaw puzzle in fog, due to the lack of structural information of the major machinery, the spliceosome. More than 30 years after its first component was discovered, the highest resolution of spliceosome complex was only solved at about 20 Å using electron microscopy, far away for us to understand the detailed mechanism of how RNA splicing is performed. This embarrassing situation now gets to a turning point to catch up soon with other central dogma steps. The most recently solved spliceosome complex structure at near atomic resolution by professor Yigong Shi and his colleagues in Tsinghua University wipes off the fog and opens new doors for understanding RNA splicing in a new foreground [1,2].

Almost all RNA transcripts in eukaryotic cells need additional processing steps before their maturation. The most critical processing step for messenger RNAs (mRNA) is the splicing, which removes the non-coding region (intron) and ligates the coding regions (exon) from a newly transcribed RNA to generate a mature RNA for translation. In splicing, the two ends of an intron are brought to spatial proximity for two successive trans-esterification reactions to occur, producing a lariat intron structure and ligated exons. This highly regulated and complicated process is performed by the dynamic ribonucleoprotein complex, spliceosome, which comprises more than 100 protein components and five small nuclear RNAs [3]. During splicing, the spliceosome undergoes major compositional and conformational

email: hongweiwang@tsinghua.edu.cn

changes to fulfill its role in promoting the correct RNA processing. Biochemical studies in the past few decades have accumulated valuable knowledge on the splicing mechanism, and structural work of several individual protein components or sub-complexes of spliceosome provided glimpses of the complex's activity [3]. These are, however, far from enough for us to understand the mechanism of splicing within the spliceosome, such as how the intron of pre-RNA transcript is recognized, folded, cleaved, and removed and how protein components in the complex coordinate with RNAs to perform the reaction in high precision.

To achieve atomic resolution structure of the entire spliceosome complex is a necessity to answer many of the key questions of splicing mechanism. A major obstacle for this goal is due to the molecular machine's intrinsic property as a dynamic and flexible multi-megadalton complex. This makes it very difficult to be purified in abundant and homogenous form, to be crystallized, therefore to be solved structurally using X-ray crystallography. The most recent technique revolution in cryo-electron microscopy made it possible to solve macromolecules at atomic resolution without the crystallization step [4]. The even more powerful capability of the new technology is to separate compositional or conformational heterogeneous molecule structures by computation after taking images of them in frozen-hydrated specimen in an electron microscope. The new method allows much smaller amount of materials than X-ray crystallography to be used for structural studies at reasonably high resolution. The new technology allowed the revelation of the spliceosome U4/U6.U5 tri-snRNP at 5.6 Å resolution for the first time [5]. Shi and his colleagues applied this new technology of cryo-electron microscopy to the spliceosome complex purified from cells using a modified protocol and solved the structure of the complex at 3.6

Å resolution [1]. In this structure, for the first time, we can see how proteins and snRNAs tangle together in a highly coordinated way and how the splicing active center is spatially organized among different snRNAs and the mRNA [2]. The structure reveals clearly that spliceosome is also a ribozyme, like ribosome and telomerase, in other steps of central dogma. This further underscores the view that RNA, as the hub in the central dogma scheme, is likely to be the earliest form of life emerging on the ancient earth, i.e. the RNA world.

Just like the first structures of DNA polymerase, RNA polymerase, and ribosomes, the seminal first structure of the spliceosome sets the ground for further structural and mechanistic studies of RNA splicing. The technology and procedures developed in solving the structure will soon stimulate the pace of this exciting field moving forward and lead to many new discoveries ahead. We may expect optimistically in the near future revealing the full splicing cycle in atomic details, deciphering the mechanism of alternative splicing regulation, and possibly implementing new gene editing tools or biomedical treatment targeting RNA splicing.

- 1 Yan C, Hang J, Wan R, Huang M, Wong CC, Shi Y. Structure of a yeast spliceosome at 3.6-Å resolution. *Science*, 2015, 349: 1182–1191
- 2 Hang J, Wan R, Yan C, Shi Y. Structural basis of pre-mRNA splicing. *Science*, 2015, 349: 1191–1198
- 3 Will CL, Lührmann R. Spliceosome structure and function. *Cold Spring Harb Perspect Biol*, 2011, 3: pii: a003707
- 4 W. Kuhlbrandt, *Biochemistry*. The resolution revolution. *Science*, 2014, 343: 1443–1444
- 5 Nguyen TH, Galej WP, Bai XC, Savva CG, Newman AJ, Scheres SH, Nagai K. The architecture of the spliceosomal U4/U6.U5 tri-snRNP. *Nature*, 2015, 523: 47–52

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.