

## PRIZE ESSAY



FINALIST:  
CELL AND  
MOLECULAR  
BIOLOGY

**Ruixue Wan**  
Ruixue Wan received her undergraduate degree from Sun Yat-sen Univer-

sity in Guangzhou, China, and her Ph.D. from Tsinghua University in Beijing. She is currently a postdoctoral fellow at Tsinghua, where she is conducting research on structural and biochemical investigations of the spliceosome to elucidate the mechanism of pre-mRNA splicing and the regulation of the spliceosome.

[www.sciencemag.org/content/362/6417/904.2](http://www.sciencemag.org/content/362/6417/904.2)

## CELL AND MOLECULAR BIOLOGY

# A key component of gene expression, revealed

## High-resolution microscopy sheds light on the molecular mechanisms of the spliceosome

By **Ruixue Wan**

In eukaryotes, genetic information stored in DNA is first transcribed into precursor messenger RNAs (pre-mRNAs), in which noncoding (introns) and coding regions (exons) appear alternatively. Introns must be removed, and exons are ligated to generate a mature mRNA that contains continuous protein-coding sequences through a sophisticated process called pre-mRNA splicing, which was discovered 40 years ago (1, 2). Splicing comprises two sequential transesterification reactions, branching and exon ligation, executed by a multi-megadalton, ribonucleoprotein (RNP) complex known as the spliceosome (3).

The spliceosome is composed of five uridine-rich, small nuclear RNPs (U1, U2, U4, U5, and U6 snRNPs—each of which contains one U snRNA and associated proteins), the NineTeen complex (NTC), the NTC-related complex (NTR), and various protein factors. To catalyze splicing, the spliceosome must recognize the three conserved splice sites (5' and 3' splice sites at exon-intron junctions and the branch point sequence) and undergo precise stepwise assembly and remodeling (Fig. 1A).

In the three decades between the discovery of the spliceosome (4–6) and the beginning of this century, scientists sought to identify its constitutional components and biochemically analyze the various spliceosomal complexes by immunoprecipitation, cross-linking mass spectrometry, in vitro splicing assay, and so on. However, our understanding of the molecular mechanism was impeded owing to the lack of high-resolution structures of the fully assembled spliceosome.

The spliceosome exhibits exceptional conformational flexibility and compositional dynamics. Compared to the relatively stable structures of RNA polymerase and the ribosome, an atomic structure of the spliceosome was long thought to be insurmountable. Electron microscopy (EM) studies achieved only moderate resolutions, with the highest

at 20 to 29 Å, which only unveiled the overall shape and global features (7). For my Ph.D. thesis research, I chose to confront this challenging target, investigating the molecular mechanism of the pre-mRNA splicing process by capturing the atomic structures of the spliceosome using the rapidly developing cryo-EM technology.

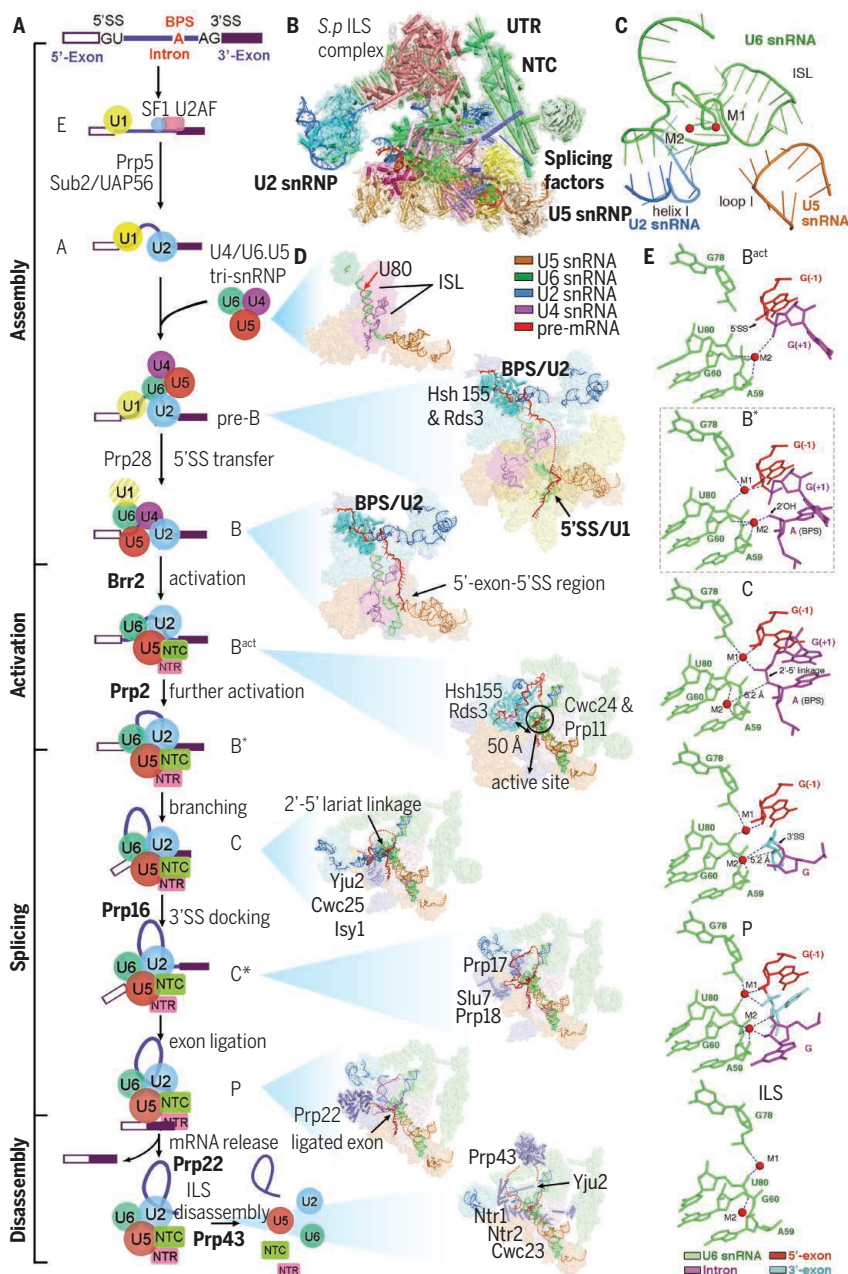
By tagging Cdc5 with TAP-tag in *Schizosaccharomyces pombe*, my co-workers and I were able to purify the intact spliceosome, which eventually yielded a structure at the overall resolution of 3.6 Å (8, 9) (Fig. 1B). This structure, containing 37 proteins and four RNA molecules, reveals the organizational principles of a functional spliceosome and the RNA-based splicing active site arrangement for the first time (Fig. 1B, C).

The active site is located in the heart of the spliceosome, consisting of the intramolecular stem-loop (ISL) of U6 snRNA, the associated Mg<sup>2+</sup> ions, helix I of the U2/U6 duplex, and loop I of U5 snRNA. Two catalytic metal ions, M1 and M2, are coordinated by zigzagged U6 snRNA in the active site. Thus, the spliceosome is in essence a protein-directed metalloribozyme. Analysis suggests that it represents a late-stage spliceosome known as the intronariat spliceosome or ILS complex because of the absence of exon sequences and the 20 Å distance separated from the lariat junction away to the catalytic ions.

To further investigate the mechanism of spliceosome assembly and catalysis, we sought to capture the structures of the spliceosome at other working states by tagging different protein components and overexpressing adenosine triphosphatase (ATPase)-defective mutants. Prior to my thesis defense, we solved the atomic or near-atomic structures of nearly all key intermediates of the spliceosome, including pre-B and B (10), B<sup>act</sup> (11), C (12), C<sup>\*</sup> (13), P (14), and ILS complexes (15) from *S. cerevisiae* (Fig. 1D). We also determined the structure of the preassembled U4/U6.U5 tri-snRNP from *S. cerevisiae* (16). Local resolution of all the structures in the core regions reaches at least 3.0 Å, allowing us to accurately identify the catalytic metal coordination during the splicing reaction (Fig. 1E).



Beijing Advanced Innovation Center for Structural Biology,  
School of Medicine, Tsinghua University, Beijing 100084, China  
Email: wxruisnow@163.com



### Molecular mechanism of pre-mRNA splicing revealed by structures of the spliceosome at different states.

(A) A scheme of the assembly and catalysis cycle of the spliceosome. This process includes four phases: assembly based on the pre-mRNA substrate, activation to prepare for the reaction, two-step splicing reaction, and disassembly of the spliceosome. State transition of the spliceosome is mainly driven by eight DEXD/H ATPases/helicases (colored red). (B) The cryo-EM structure of the ILS complex from *S. pombe* (PDB code 3JB9). (C) A close-up view of the active site of *S. pombe* ILS complex. (D) Cryo-EM structures of the U4/U6.U5 tri-snRNP, pre-B, B, B<sup>act</sup>, B<sup>\*</sup>, C, C<sup>\*</sup>, P, and ILS complex from *S. cerevisiae*. Structures are color-coded by subcomplexes related to (A). RNA components are highlighted in cartoon style. The orientation and scale are presented the same for comparison. (E) Coordination of catalytic metal ions during splicing reaction (from B<sup>act</sup> to ILS complex, B<sup>\*</sup> is predicted). The RNA elements are shown in the same orientations.

These structures provide unprecedented clarity toward a mechanistic understanding of pre-mRNA splicing. It appears that the spliceosome catalyzes the two steps of splicing reaction using a single active site, which remains almost unchanged once formed in the B<sup>act</sup> complex. The U6 snRNA coordinates the

catalytic metal ions and plays a notable role in catalyzing the two steps of splicing. U1, U2, U5, and U6 snRNAs serve to recognize and transfer the RNA elements on the pre-mRNA, including 5' and 3' splice sites, branch point sequence (BPS) and 5' exon (Fig. 1D). U4 snRNA plays an important role in sequestering

U6 snRNA before the reaction.

Throughout the splicing reaction, at least 20 components (mainly from U6 snRNA, U5 snRNP, NTR, and a small portion of U2 snRNA and NTC) act as a rigid scaffold, maintaining the conformation of the active site. The 3' end of U2 snRNA and ~13 proteins (part of U2 snRNP and NTC) are compositionally unchanged, but mobile, facilitating the delivery of critical reaction groups on the RNA molecules into the active site for the two-step reactions. The compositional change and remodeling of the spliceosome are mainly driven by ATPase/helicases and its cofactors. Thus, a near-complete cycle of pre-mRNA splicing can be recapitulated in atomic details (Fig. 1).

As an essential step in eukaryotic gene expression, splicing is relevant to many diseases (17). For example, Hsh155 is frequently mutated in hematological malignancies (18). Based on the structure of the B<sup>act</sup> complex, we mapped 36 cancer-derived mutations to Hsh155. Most of these residues bind upstream of BPS, thus compromising the binding ability and adversely affect splicing.

Forty years after the discovery of pre-mRNA splicing, we have finally begun to understand its underlying structural mechanism. This work will enable the understanding of the intricacies of an essential biological process. It will also aid drug discovery for numerous diseases that are linked to splicing malfunction, including many cancers, blindness, and other genetic diseases. The next challenge is to investigate the more complex regulation mechanism of the spliceosome, such as alternative splicing. ■

### REFERENCES AND NOTES

1. S. M. Berget, C. Moore, P. A. Sharp, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 3171 (1977).
2. L. T. Chow, R. E. Gelinas, T. R. Broker, R. J. Roberts, *Cell* **12**, 1 (1977).
3. M. J. Moore, C. C. Query, P. A. Sharp, *Splicing of Precursors to mRNA by the Spliceosome*. The RNA World (Cold Spring Harbor Monograph Archive, 1993).
4. E. Brody, J. Abelson, *Science* **228**, 963 (1985).
5. D. Frendewey, W. Keller, *Cell* **42**, 355 (1985).
6. P. J. Grabowski, S. R. Seiler, P. A. Sharp, *Cell* **42**, 345 (1985).
7. R. Lührmann, H. Stark, *Curr. Opin. Struct. Biol.* **19**, 96 (2009).
8. C. Yan *et al.*, *Science* **349**, 1182 (2015).
9. J. Hang, R. Wan, C. Yan, Y. Shi, *Science* **349**, 1191 (2015).
10. R. Bai, R. Wan, C. Yan, J. Lei, Y. Shi, *Science* **360**, 1423 (2018).
11. C. Yan, R. Wan, R. Bai, G. Huang, Y. Shi, *Science* **353**, 904 (2016).
12. R. Wan, C. Yan, R. Bai, G. Huang, Y. Shi, *Science* **353**, 895 (2016).
13. C. Yan, R. Wan, R. Bai, G. Huang, Y. Shi, *Science* **355**, 149 (2016).
14. R. Bai, C. Yan, R. Wan, J. Lei, Y. Shi, *Cell* **171**, 1589 (2017).
15. R. Wan, C. Yan, R. Bai, J. Lei, Y. Shi, *Cell* **171**, 120 (2017).
16. R. Wan *et al.*, *Science* **351**, 466 (2016).
17. M. M. Scotti, M. S. Swanson, *Nat. Rev. Genet.* **17**, 19 (2016).
18. C. N. Hahn, H. S. Scott, *Nat. Genet.* **44**, 9 (2011).

### ACKNOWLEDGMENTS

I am extremely grateful to my Ph.D. adviser, Y. Shi, for his mentorship. I also thank all of the members of the Shi lab, both past and present, especially L. Zhou and R. Bai.

10.1126/science.aav6875

## A key component of gene expression, revealed

Ruixue Wan

*Science* **362** (6417), 904.

DOI: 10.1126/science.aav6875

### ARTICLE TOOLS

<http://science.sciencemag.org/content/362/6417/904.2>

### REFERENCES

This article cites 17 articles, 8 of which you can access for free  
<http://science.sciencemag.org/content/362/6417/904.2#BIBL>

### PERMISSIONS

<http://www.sciencemag.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of Service](#)

---

*Science* (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. 2017 © The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. The title *Science* is a registered trademark of AAAS.