What is your idea of perfect happiness?

To me there are two kinds of perfect happiness. One is with my family, especially my daughter. I never knew the joy of having a kid until I had her! Another kind of perfect happiness would be the same as many others who work as scientists and perhaps other fields. This is the rewarding feeling when I have confirmation of an answer to a challenging and long-standing unanswered question.

Why did you start working on your current research topic? (What first drew you to the question?)

In 2009, perhaps one of the most exciting discoveries in Molecular Biology was the widespread expression of the 'mRNA-like' lncRNAs that originate from intergenic regions. I worked on the lncRNA NEAT1 at that time and the long isoform of NEAT1 is nonpolyadenylated. So I asked myself the question: 'Do all lncRNAs look similar to mRNAs?' This idea motived me to search for additional novel types of lncRNAs in the nonpolyadenylated transcriptomes. At that time, I had just completed my PhD and had successfully applied for independent funding (a State of Connecticut Stem Cell Grant). I therefore began to develop methods to visualize and characterize nonpolyadenylated RNAs. This work, initiated in the USA but completed in China, has led to the discovery of several classes of RNA species in my lab located in Shanghai since 2011.

What was the most exciting question in your field right now?

Studies on lncRNAs are advancing at a rapid pace, but some characteristics of each lncRNA are still underappreciated. The recent advent of cutting-edge tools in Molecular Biology such as cryoelectron microscopy and gene-editing technologies has equipped RNA biologists with new ways to understand lncRNA molecules in greater detail. How different lncRNAs are processed, structured, and act as well as their distinctions and commonalities in gene regulation will be exciting to address in the future.

[1]State Key Laboratory of Molecular Biology, CAS Center for Excellence in Molecular Cell Science, Shanghai Key Laboratory of Molecular Andrology, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China
[2]School of Life Science and Technology, ShanghaiTech University, 100 Haike Road, Shanghai, China

*Correspondence: linglingchen@sibcb.ac.cn (L.-L. Chen).
http://dx.doi.org/10.1016/j.tibs.2016.07.006

Spotlight

Direct Duplex Detection: An Emerging Tool in the RNA Structure Analysis Toolbox

Chase A. Weidmann,[1,2] Anthony M. Mustoe,[1,2] and Kevin M. Weeks[1,*]

While a variety of powerful tools exists for analyzing RNA structure, identifying long-range and intermolecular base-pairing interactions has remained challenging. Recently, three groups introduced a high-throughput strategy that uses psoralen-mediated crosslinking to directly identify RNA–RNA duplexes in cells. Initial application of these methods highlights the preponderance of long-range structures within and between RNA molecules and their widespread structural dynamics.

The role of RNA as a central functional molecule in biology has become increasingly clear. Long noncoding RNAs (lncRNAs), small nuclear and nucleolar RNAs, miRNAs, and the untranslated regions of mRNAs regulate integral cellular processes including transcription, splicing, localization, ribosome assembly, and translation [1]. Even within coding regions of mRNAs, regulatory information is layered on top of the protein message via differential usage of synonymous codons. Correspondingly, the list of human diseases associated with mutations in and misregulation of RNA molecules is growing at an accelerating pace [2].

RNA molecules mediate diverse functions in part by folding into unique structures comprising complementary base-pairing and higher-order tertiary interactions. A classical example is the L-shaped structure of tRNA, which is essential to its role in peptide synthesis. Because underlying structure is inextricably linked to the complex activities of most RNAs, the ability to accurately identify intracellular RNA structure is paramount and multiple technologies have been developed for RNA structure analysis (Figure 1). Two long-used methods are sequence covariation analysis and thermodynamic free-energy minimization. Covariation analysis is limited by the need for a diverse multiple sequence alignment and thermodynamic approaches, which work well for short RNAs, struggle when used to explore biological hypotheses for large and complex RNAs. For most RNAs, structure analysis driven by experimental information is therefore essential. NMR, X-ray crystallography, and cryoelectron microscopy (cryo-EM) provide unparalleled resolution but are low throughput and limited to 'well-behaved' molecules. The most widely useful class of methods is single-nucleotide-resolution structure probing experiments, which use a chemical or enzymatic reagent that react with RNA in a quantitative and structurally informative way [3].

State-of-the-art strategies leverage high-throughput sequencing (-seq) and integrate the unique advantages of multiple tools. For example, in the selective 2′-hydroxyl acylation analyzed by primer extension and mutational profiling

(SHAPE-MaP) approach, highly accurate structure models can be obtained by using nucleotide-resolution structure-probing data to restrain thermodynamic RNA modeling algorithms [4]. Typically, mutational analysis or comparative genomics is then used to validate *de novo* structure models. Nevertheless, it is challenging to identify long-range intramolecular pairing interactions and new strategies are needed to detect intermolecular interactions.

Several groups have recently introduced high-throughput crosslinking-based structure analysis experiments that can detect RNA duplexes in cells. These approaches are psoralen analysis of RNA interactions and structures (PARIS), ligation of interacting RNAs followed by high-throughput sequencing (LIGR-seq), and sequencing of psoralen-crosslinked, ligated, and selected hybrids (SPLASH) [5–7]. These three duplex-detection experiments are similar and employ five major steps: (i) a derivative of the chemical reagent psoralen is used to selectively crosslink duplexed RNA strands; (ii) crosslinked duplexes are enriched; (iii) the two strands of each duplex are ligated together; (iv) the crosslink is reversed; and (v) the ligated junctions are sequenced to identify the interacting RNA sequences (Figure 1). The trapping of RNA duplexes through strand-to-strand crosslinks distinguishes PARIS, LIGR-seq, and SPLASH from prior methods. For example, crosslinking, ligation, and sequencing of hybrids (CLASH) and RNA hybrid and individual-nucleotide-resolution UV crosslinking and immunoprecipitation (hiCLIP) can be used to detect RNA duplexes but rely on crosslinking between RNA duplexes and duplex-specific RNA-binding proteins, thus excluding non-protein-bound duplexes [8,9]. RNA proximity ligation is another promising method but, since it forgoes crosslinking, only the subset of duplexes that persists throughout downstream biochemical processing steps is detected [10].

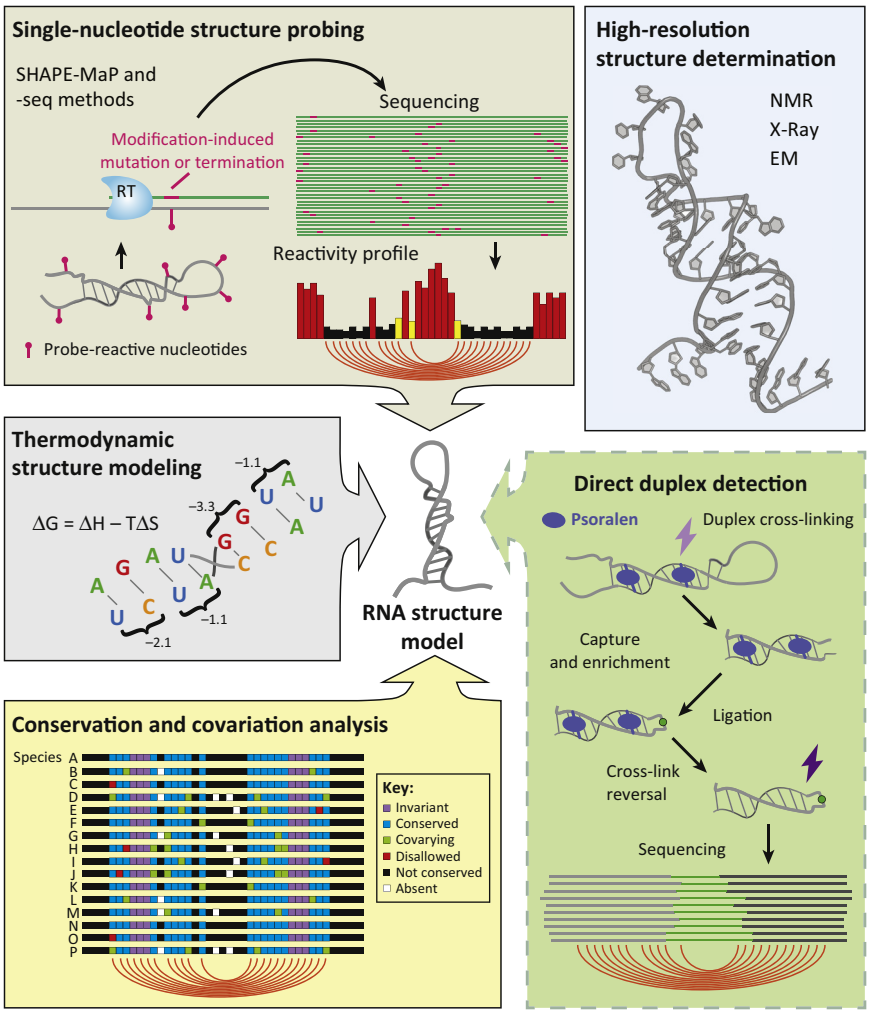

Figure 1. Direct Duplex-Detection Methods Join a Host of Powerful Tools for RNA Structure Analysis. NMR, X-ray crystallography, and cryoelectron microscopy (cryo-EM) can be used to determine atomic-resolution structures for 'well-behaved' RNAs (blue box). For most other RNAs, a suite of complementary methods facilitates secondary structure determination. RNA structure can be probed with chemical or enzymatic reagents that, when read out by sequencing, provide a nucleotide-resolution profile of RNA flexibility [selective 2′-hydroxyl acylation analyzed by primer extension and mutational profiling (SHAPE-MaP) and -seq methods; brown box]. Thermodynamic stability parameters can be used to estimate the minimum energy structure for an RNA sequence (gray box). Base-paired nucleotides can be inferred from conservation patterns observed in sequence alignments of homologous RNAs (yellow box). This existing toolset is complemented by new, direct duplex-detection methods in which paired RNA strands are crosslinked with psoralen and ligated together followed by detection via sequencing (green box).

These new psoralen-mediated duplex-detection methods feature important advances in RNA structural analysis. First, unlike structure probing, duplex detection provides direct evidence of specific pairing interactions. Lu *et al.* report that up to 40% of PARIS-identified duplexes in mammalian cells are separated by distances of greater than 200 nucleotides and 5–10% span distances greater than 1000 nucleotides. The ability to detect very long-range interactions suggests that it may soon be possible to examine potential functional roles of higher-order structure in mRNAs and lncRNAs. The additional ability to detect intermolecular duplexes offers a much-needed method for studying higher-order structure in multi-RNA complexes. A

second advantage is that duplex-detection methods are single-molecule experiments, with each ligated sequence read representing a single occurrence of an inter- or intra-RNA duplex. Thus, it is possible to directly observe RNA regions that participate in mutually exclusive duplex interactions, suggesting that an RNA adopts multiple structures.

PARIS, LIGR-seq, and SPLASH represent a novel class of tools for understanding RNA structure–function relationships, but it is important to recognize several current limitations. First, the methods are complex and require multiple biochemical processing steps that are each likely to introduce significant biases. Further, ligated junctions represent only a small fraction of obtained sequencing reads, with most data discarded as waste. Third, the resolution of these approaches is on the scale of tens of nucleotides rather than individual nucleotides. Finally, these methods primarily provide qualitative confirmation that a duplex exists, and the absence of a cross-link does not indicate an absence of structure. More quantitative data are needed for *de novo* RNA structure determination, and understanding dynamic RNAs requires quantitative information about the relative populations of different states. Thus, as currently implemented, duplex-detection methods are most useful for identifying interesting structures for subsequent in-depth follow up and for validating structures established by complementary approaches.

RNA structure analysis has matured to a state beyond what was imaginable just several years ago, with researchers now attempting to parse structure–function relationships *in vivo* for highly complex and dynamic RNA molecules. Direct duplex-detection experiments are a welcome addition to the RNA structure analysis toolbox. As we push forward on the frontiers of RNA structural and mechanistic biology, we will need to leverage and integrate multiple innovative tools.

[1]Department of Chemistry, University of North Carolina, Chapel Hill, NC 27599-3290, USA
[2]These authors contributed equally to this work.

*Correspondence: weeks@unc.edu (K.M. Weeks).

http://dx.doi.org/10.1016/j.tibs.2016.07.001

References

1. Cech, T.R. and Steitz, J.A. (2014) The noncoding RNA revolution – trashing old rules to forge new ones. *Cell* 157, 77–94
2. Huarte, M. (2015) The emerging role of lncRNAs in cancer. *Nat. Med.* 21, 1253–1261
3. Kwok, C.K. *et al.* (2015) The RNA structurome: transcriptome-wide structure probing with next-generation sequencing. *Trends Biochem. Sci.* 40, 221–232
4. Siegfried, N.A. *et al.* (2014) RNA motif discovery by SHAPE and mutational profiling (SHAPE-MaP). *Nat. Methods* 11, 959–965
5. Lu, Z. *et al.* (2016) RNA duplex map in living cells reveals higher-order transcriptome structure. *Cell* 165, 1267–1279
6. Aw, J.G.A. *et al.* (2016) *In vivo* mapping of eukaryotic RNA interactomes reveals principles of higher-order organization and regulation. *Mol. Cell* 62, 603–617
7. Sharma, E. *et al.* (2016) Global mapping of human RNA–RNA interactions. *Mol. Cell* 62, 618–626
8. Sugimoto, Y. *et al.* (2015) hiCLIP reveals the *in vivo* atlas of mRNA secondary structures recognized by Staufen 1. *Nature* 519, 491–494
9. Kudla, G. *et al.* (2011) Cross-linking, ligation, and sequencing of hybrids reveals RNA–RNA interactions in yeast. *Proc. Natl. Acad. Sci. U.S.A.* 108, 10010–10015
10. Ramani, V. *et al.* (2015) High-throughput determination of RNA structure by proximity ligation. *Nat. Biotechnol.* 33, 980–984

Spotlight

CENPs and Sweet Nucleosomes Face the FACT

Magdalena Murawska[1] and Andreas G. Ladurner[1,2,3,*]

Chaperones mediate vital interactions between histones and DNA during chromatin assembly and reorganization. Two recent studies reveal novel substrates for the essential and conserved histone chaperone FAcilitates Chromatin Transcription (FACT). Prendergast *et al.* show that FACT helps deposit important histone-fold proteins on centromeres. Raj *et al.* find that FACT preferentially binds *O*-GlcNAcylated nucleosomes, suggesting that FACT may contribute to nutrient-regulated cellular programs.

FACT is a pivotal histone chaperone that facilitates the turnover and dynamics of histones during DNA transcription, replication, and repair. It is also involved in heterochromatic silencing [1]; however, its function in this context is less clear. In human cells FACT plays a role in depositing the centromeric histone CENP-A [2] while in yeast it prevents spurious CENP-A deposition outside centromeres through interaction with a complex that degrades the centromeric histone [3]. There has been progress in deciphering how FACT interacts with histones, including through high-resolution structures [4,5], but questions remain about how FACT gains specificity towards nucleosomes and what role it plays at centromeres.

Two recent studies reveal exciting insights into FACT biology. A paper from Genevieve Almouzni's team reports that FACT interacts with and contributes to the deposition of the histone-fold protein CENP-T/W, a heterodimer essential for kinetochore function [6]. In a second report, Benjamin Davis's team demonstrates that *O*-GlcNAcylation of histone H2B Ser112 strongly enhances the interaction of FACT with nucleosomes [7]. These novel interactions broaden the biological scope for this essential chaperone.

The kinetochore is a large macromolecular assembly comprising >100 proteins. It orchestrates mitotic chromosome segregation by linking microtubules with centromeric chromatin. Although much progress has been made at the cell biological level and we better understand its biochemical composition, less is known about how kinetochores are assembled and how they interact with chromatin. The histone-fold proteins CENP-T, -W, -S, and -X are thought to form a nucleosome-like, heterotetrameric complex that is crucial for DNA–kinetochore contacts [8], so recent interest in these proteins has centered on their chromatin interaction.