

STRUCTURAL BIOLOGY

RNA structure served *in vivo*

Two transcriptome-scale sequencing methods provide a more complete view of RNA structure in its native context.

Once seen as a humble messenger of information, RNA's reprinted role as a diverse and busy regulator rests on its structural flexibility and ability to partner with proteins and nucleic acids. "A lot of RNA-based regulation is really centered on understanding RNA structure," says Howard Chang, a biologist at Stanford University. Methods to further this understanding have advanced rapidly, but important blind spots persist. Two approaches now bring us closer to the goal of studying full-length RNA in the molecular milieu of the cell. Both address secondary structure—the base-pairing interactions that RNA assumes as it folds on itself or hybridizes with other RNA molecules.

Jernej Ule and Nicholas Luscombe at the University College London and the Francis Crick Institute led an effort to target the RNA length 'ceiling' that available methods could probe. "When you reach the limit of 100 nucleotides," says Ule, "the number of computational predictions often becomes too large."

Their hybrid and individual-nucleotide-resolution ultraviolet cross-linking and immunoprecipitation (hiCLIP) approach is related to other methods that stabilize interactions by cross-linking and that then use an antibody to pull down RNA complexed to protein (Sugimoto *et al.*, 2015). For example, CLASH (cross-linking, ligation and sequencing of hybrids), from the lab of David Tollervey at the University of Edinburgh, includes a ligation step that joins RNA strands that participate in duplexes and can identify double-stranded regions near cross-link sites. hiCLIP modifies this by first ligating a mixture of two adaptors, then chemically deprotecting one of them before performing a second ligation that links this adaptor to the second strand in the duplex.

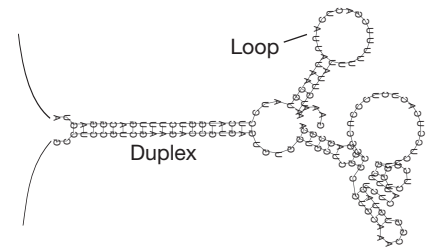
The additional adaptor solves important challenges; it makes it possible to identify the junction of the two strands unambiguously,

improve mapping of these short sequence fragments and distinguish aberrant sequencing products. The overhang also provides flexibility for short ends of the RNA duplex to circularize efficiently for ligation and recovery. Only about 2% of resulting sequence reads span duplex regions, but they were sufficient to identify over 4,000 mRNA duplexes.

By applying hiCLIP to Staufen 1, a double-stranded RNA-binding protein that regulates transcript localization, Ule, Luscombe and their teams made the unexpected discovery that long-range duplexes commonly connect regions separated by hundreds and even thousands of nucleotides. Many loops join untranslated regions near the stop codon and tail ends of the transcript, whereas few interactions occur in the coding sequence of highly translated mRNAs.

The groups of Howard Chang and his Stanford colleague Eric Kool took a different, protein-independent, approach to structure determination based on chemicals that preferentially react with open or single-stranded regions and interfere with reverse transcription at those sites (Spitale *et al.*, 2015). In cells, dimethyl sulfate sequencing can do this, but it reacts with two nucleotides, thereby limiting resolution, whereas RNA-selective 2'-hydroxyl acylation and primer extension (SHAPE) using a cell-permeable chemical probe can measure the structural reactivity of all four RNA nucleotides, but not across the transcriptome.

Going global with SHAPE required solving a noise problem. Because low chemical levels are needed to ensure that single molecules are not modified more than once, the vast majority of sequenced RNAs are actually unmodified. The researchers tackled this by adding an enrichment step, which required synthesizing a bifunctional molecule that was cell permeable and not overreactive, a feat that "really depended on [Kool's] chemical intuition," says Chang. The new molecule can react with any RNA nucleotide and includes an azide handle for biotin addition in an easy



Two methods probe RNA secondary structure *in vivo* by identifying double-stranded regions (duplexes) near a bound protein or single-stranded regions such as loops. Adapted from Sugimoto *et al.*, Nature Publishing Group.

'click-chemistry' reaction, allowing structurally informative molecules to be purified.

"You think about the SHAPE chemistry as spray-painting the transcriptome with a chemical tag," says Chang. "Any part that's single-stranded can get spray-painted." Protein binding produces a shadow in the paint that becomes apparent when *in vivo* RNA is compared to naked RNA *in vitro*. Using *in vivo* click SHAPE (icSHAPE) in mouse embryonic stem cells enabled the researchers to predict the binding of protein and microRNA regulators. From dynamic structural signatures between wild-type and *Mettl3* N⁶-methyladenosine (m⁶A) methyltransferase mutant cells, they identified a role for m⁶A, the most common chemical modification in RNA: "it's actually a structural switch for RNA," says Chang. The local regions surrounding an m⁶A moiety lose double-stranded structure when m⁶A is present.

These two new methods extend the range of structures that we can view in the cell to long-range loops, long noncoding RNAs and dynamic structures under any set of genetic and environmental perturbations.

Tal Nawy

RESEARCH PAPERS

Spitale, R.C. *et al.* Structural imprints *in vivo* decode RNA regulatory mechanisms. *Nature* **519**, 486–490 (2015).

Sugimoto, Y. *et al.* hiCLIP reveals the *in vivo* atlas of mRNA secondary structures recognized by Staufen 1. *Nature* **519**, 491–494 (2015).