Published in final edited form as:
Nature. 2014 January 30; 505(7485): 621–622. doi:10.1038/505621a.

Molecular biology: A second layer of information in RNA

Silvia B. V. Ramos and **Alain Laederach**
Obstetrics and Gynecology and Biology Departments, University of North Carolina, Chapel Hill, North Carolina 27599-3280, USA

Silvia B. V. Ramos: silvia_ramos@med.unc.edu; Alain Laederach: alain@unc.edu

Abstract

Three studies have characterized the full complement of RNA folding in cells. They find large numbers of secondary structures in RNA, some of which may have functional consequences for the cell.

Generally, RNA is understood as a messenger of genetic information in the cell: it is transcribed from DNA and then translated into proteins[1]. Stretches of RNA that are complimentary in sequence have a propensity to pair, forming elements of secondary structure within RNA molecules. But the prevalence of secondary structure in messenger RNAs, and its role in RNA regulation, is not fully understood. In this issue, three reports[2,3,4] describe transcriptome-wide analysis of mRNA using structure-probing techniques, which begin to reveal the extent of secondary structure in the *transcriptomes* (all the RNA molecules present in a population of cells) of plants, humans, and yeast.

The chemical structure of RNA is analogous to DNA. It is comprised of a sugar-phosphate backbone and four distinct bases: adenine (A), cytosine (C), guanine (G) and uracil (U). As with DNA, these bases interact by forming hydrogen bonds, resulting in aptly named Watson–Crick pairs (G–C and A–U). However, unlike DNA, complementary bases from two RNA molecules do not pair up to form a double helix, a formation that prevents secondary structures from arising in DNA. Instead, the nucleotides of RNA are free to interact with one another within each molecule, resulting in folding of the RNA chain into secondary structures.

The functional consequences of secondary structural elements in RNA depend on their molecular context. Some specific structural elements have well-established post-transcriptional regulatory roles, but these are restricted to small subsets of mRNAs[5,6]. In some cases, for example ribosomal RNA (which is a part of the cellular machinery responsible for protein synthesis), secondary structural elements fold further into compact three-dimensional conformations capable of catalyzing reactions[7].

The three new studies combine well-established chemical and enzymatic structure-probing techniques for determining RNA secondary structure with next-generation sequencing, a method that allows sequencing of millions of stretches of nucleotides simultaneously. All three report unprecedented coverage of the transcriptome[8]. In doing so, they demonstrate unequivocally that most mRNAs have a propensity to form secondary structures *in vitro*, in the absence of any other cellular components.

Each group reported that some of the RNA structures they observed *in vitro* were altered *in vivo*. In fact, Weissman *et al*.[2] found evidence in yeast to suggest that RNA structures in the cell are actively unfolded by proteins. Nonetheless, the papers show that structural patterns are evolutionarily conserved at several functional sites within RNA molecules. These results provide the first *in vivo* data to suggest that, if given the opportunity, RNA will fold. This is consistent with many previous *in vitro* studies[9] of RNA structure and folding. As mRNA must be unfolded to successfully act as a messenger, the cell must find ways to get around the folding problem.

In addition to their structural characterization of the human transcriptome, Chang *et al*.[3] performed comparative structure probing in cell lines derived from a family trio (mother, father and child). In so doing, they were able to assess the structural consequences of natural human inter-generational genetic variation on the transcriptome. They discovered over 1,900 single-nucleotide mutations that alter RNA structure. These experiments therefore yielded thousands of new putative 'ribosnitches'[7,10]. A ribosnitch is broadly defined as an RNA in which a specific single-nucleotide mutation alters structure[7]; they are analogous to bacterial riboswitches, which change structure on binding of a small molecule and regulate transcription or translation[11].

RNA structure has the potential to influence post-transcriptional processes in the. Therefore, a subset of the putative ribosnitches could be functional. In fact, mutations that disrupt important RNA secondary structure elements can cause human disease[10]. Although the structural changes identified in Chang and colleagues' work are not by themselves indicators of malfunction, as the three individuals studied are presumably healthy, the newly identified putative ribosnitches have the potential to help identify mechanisms by which structural changes can give rise to disease, an exciting step forward.

The application of next-generation sequencing to the transcriptome has revealed the complexity of post-transcriptional regulatory networks. The structural dimension of this complexity is now accessible with the publication of these three studies that elegantly leverage the power of next generation sequencing.. Although the three studies reveal similar general structural features of transcripts, there are key differences in the specific features found by each approach. Such discrepancies may arise due to differences in experimental design, which can cause changes to the inherently-dynamic structure of RNA. In this case, each study used different protocols for RNA extraction, library preparation and, in particular, determining levels of background noise. These experimental details must be taken into account when comparing structures discovered using the different approaches.

These three studies provide our first insight into the secondary structure of an entire transcriptome in eukaryotes — the class of organisms comprising plants, animals and fungi. However, a full characterization of transcriptome structure will require a concerted community effort, with an emphasis on standardization to allow quantitative comparisons of these data sets. Only then will it be possible to fully integrate these findings to determine the structural elements that are consequential in the transcriptome[12].

References

1. Crick F. Nature. 1970; 227:561–563. [PubMed: 4913914]
2. Rouskin S, Zubradt M, Washietl S, Kellis M, Weissmann JS. Nature. 2014 XXX.
3. Wan Y, et al. Nature. 2014 XXX.
4. Ding Y, et al. Nature. 2014 XXX.
5. Woodson SA. Curr. Opin. Chem. Biol. 2008; 12:667–673. [PubMed: 18935976]
6. Dominski Z, Marzluff WF. Gene. 2007; 396:373–390. [PubMed: 17531405]
7. Halvorsen M, Martin JS, Broadaway S, Laederach A. PLoS Genet. 2010; 6:e1001074. [PubMed: 20808897]
8. Tijerina P, Mohr S, Russell R. Nature Protocols. 2007; 2:2608–2623.
9. Shcherbakova I, Mitra S, Laederach A, Brenowitz M. Curr. Opin. Chem. Biol. 2008; 12:655–666. [PubMed: 18926923]
10. Martin JS, et al. RNA. 2012; 18:77–87. [PubMed: 22109839]
11. Tucker BJ, Breaker RR. Curr. Opin. Struct. Biol. 2005; 15:342–348. [PubMed: 15919195]
12. Ulitsky I, Bartel DP. Cell. 2013; 154:26–46. [PubMed: 23827673]

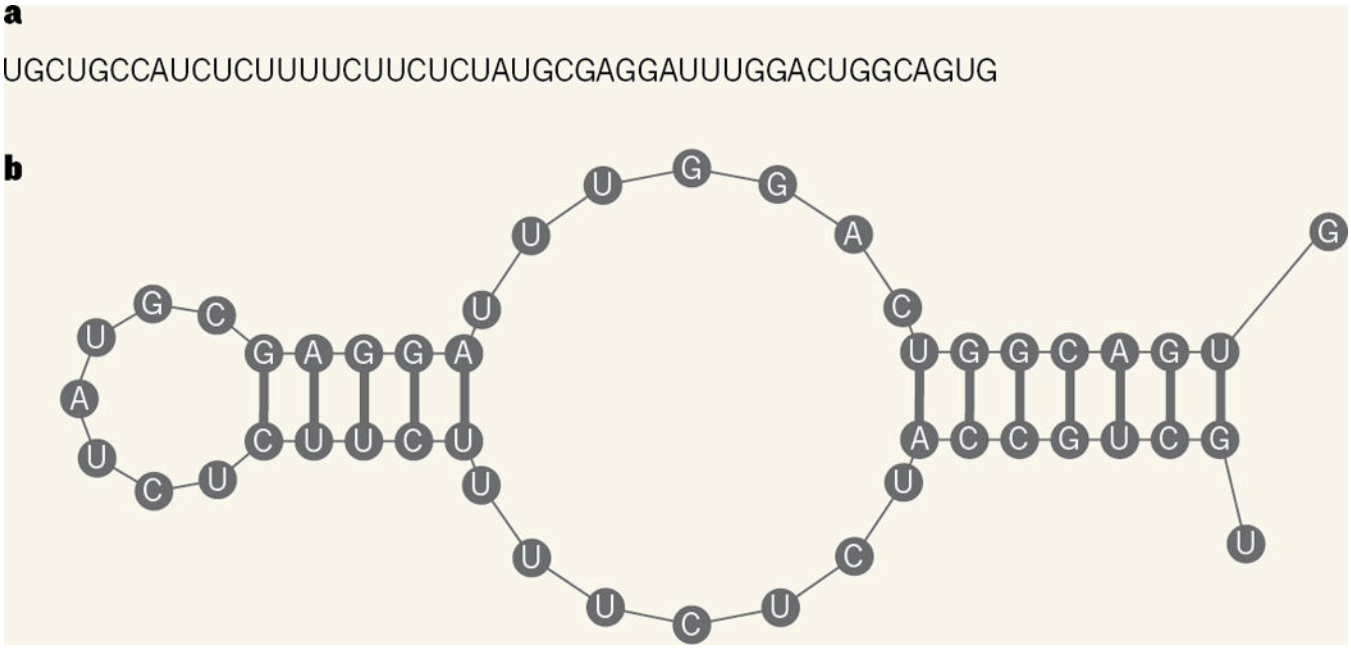

Figure 1. Principles of RNA primary sequence and secondary structure

a, RNA is a single-stranded polymer, with bases adenine (A), cytosine (C), guanine (G) and uracil (U). b, Unlike DNA, RNA molecules do not pair up to form helices. The bases of an individual molecule can therefore pair with one another (G–C and A–U), causing the RNA to fold into secondary structures. Three reports[2,3,4] find that such folding is commonplace in humans, plants and yeast.