



# THE ELABORATE ARCHITECTURE OF RNA

Cells contain an ocean of twisting and turning RNA molecules. Now researchers are working out the structures — and how important they could be.

BY ELIE DOLGIN

When Philip Bevilacqua decided to work out the shapes of all the RNA molecules in a living plant cell, he faced two problems. First, he had not studied plant biology since high school. And second, biochemists had tended to examine single RNA molecules; tackling the multitudes that waft around in a cell was a much thornier challenge.

Bevilacqua, an RNA chemist at Pennsylvania State University in University Park, was undeterred. He knew that RNA molecules were vital regulators of cell biology and that their structures might offer broad lessons about how they work. He brushed up on plant anatomy in an undergraduate course and worked with molecular plant biologist Sarah Assmann to develop a technique that could cope with RNAs at scale.

In November 2013, they and their teams became the first to describe the shapes of thousands of RNAs in a living cell — revealing a veritable sculpture garden of different forms in the weedy thale cress, *Arabidopsis thaliana*<sup>1</sup>. One month later, a group at the University of California, San Francisco, reported a comparable study of yeast and human cells<sup>2</sup>. The

number of RNA structures they managed to resolve was “unprecedented”, says Alain Laederach, an RNA biologist at the University of North Carolina at Chapel Hill (UNC).

Scientists’ view of RNA has transformed over the past few decades. Once, most RNAs were thought to be relatively uninteresting pieces of limp spaghetti that ferried information between the molecules that mattered, DNA and protein. Now, biologists know that RNAs serve many other essential functions: they help with protein synthesis, control gene activity and modify other RNAs. At least 85% of the human genome is transcribed into RNA, and there is vigorous debate about what, if anything, it does.

But a key mystery has remained: its convoluted structures. Unlike DNA, which forms a predictable double helix, RNA comprises a single strand that folds up into elaborate loops, bulges, pseudo-knots, hammerheads, hairpins and other 3D motifs. These structures flip and twist between different forms, and are thought to be central to the operation of RNA, albeit in ways that are not yet known. “It’s a big missing piece of the puzzle of understanding

how RNAs work,” says Jonathan Weissman, a biophysicist and leader of the yeast and human RNA study.

In the past few years, researchers have begun to get a toehold on the problem. Bevilacqua, Weissman and others have devised techniques that allow them to take snapshots of RNA configurations en masse inside cells — and found that the molecules often look nothing like what is seen when RNA folds under artificial conditions. The work is helping them to decipher some of the rules that govern RNA structure, which might be useful in understanding human variation and disease — and even in improving agricultural crops.

“It gets at the very basic problem of how do living things evolve and how do these molecular rules affect what we look like and how we function,” says Laederach. “And that, fundamentally as a biologist, is really exciting.”

The best-described RNA structures are what Kevin Weeks, a chemical biologist at the UNC, calls “RNA rocks”: molecules that have changed little in their sequence or structure over evolutionary time. These include transfer RNAs and ribosomal RNAs (both involved in

protein synthesis) as well as enzymatic RNAs known as ribozymes. “But in the world of RNAs,” Weeks says, “these are probably huge outliers.”

The bulk of the RNA world is like unexplored, shifting sand. “We know next to nothing about the structure of most RNAs,” says Robert Spitale, a chemist at the University of California, Irvine. RNA molecules typically exist as a linear string of nucleotides — or bases — for only an instant after they are produced from their template DNA. They quickly fold back on themselves, and complementary nucleotides pair up. They then contort further into complex 3D configurations, interact with proteins and other RNAs and change shape to carry out different jobs.

Most techniques for probing RNA structure make use of the reactivity of the nucleotides, or their sensitivity to certain enzymes: regions that are paired up tend to respond differently from those that remain single-stranded. Computer algorithms then help to model the overall structure of the molecule. But these experiments are painstaking and laborious because researchers could interrogate only one part of one molecule at a time.

That changed five years ago, with the arrival of a technique called PARS (parallel analysis of RNA structure), developed by genome scientist Howard Chang at Stanford University in California and computational biologist Eran Segal at the Weizmann Institute of Science in Rehovot, Israel. PARS uses one enzyme to cut RNA where it is single-stranded and another to cleave it at double-stranded sites. Researchers treat a sample of RNA with each enzyme independently to produce two libraries of chopped-up RNA; they then sequence and analyse both collections to work out which nucleotides are paired, and can do this for thousands of RNA types at once.

## RNA RULES

Chang and Segal first used PARS in the budding yeast *Saccharomyces cerevisiae* to reveal the structures of more than 3,000 messenger RNAs (mRNAs)<sup>3</sup>, which bear instructions for building proteins. As well as some weird and wonderful shapes, the scientists also found one of the first clues to the laws that dictate RNA structure: the regions that code for proteins generally contain more base-pairing and have more-elaborate structures than do flanking sequences known as untranslated regions. This pattern makes sense, Chang says, because untranslated regions often interact with regulatory proteins and so need to be in a more-open, accessible orientation.

The pair followed this up last year with a study of human mRNA. Led by graduate student Yue Wan, the researchers looked at more than 20,000 mRNA structures from blood cell lines generated from

**NATURE.COM**  
To hear a podcast  
on RNA architecture:  
[go.nature.com/oqfooo](http://go.nature.com/oqfooo)

two parents and their child, and discovered around 1,900 single-nucleotide variations in regions that do not code for protein that had altered RNA structure<sup>4</sup>. The question now is whether these affect what the RNAs do, or whether they are mostly background noise.

At least some evidence suggests that they matter. In May, Laederach and his team reported on variants in the untranslated region of an mRNA that is linked to a rare form of eye cancer called retinoblastoma. In healthy individuals, this mRNA simultaneously

## THE BULK OF THE RNA WORLD IS LIKE UNEXPLORED, SHIFTING SAND.

adopts three structures, but in two people with the disease, nucleotide variants force the molecule to collapse into a single conformation<sup>5</sup>. Laederach thinks that such variations in mRNA folding could be a general mechanism of disease and a source of human variation in common traits such as height.

A major limitation of the PARS method is that the enzymes used cannot easily penetrate the cell membrane, so scientists must extract the RNA from the cell and, in doing so, disrupt the native structure. In principle, base-pairing should ensure that RNAs spring back into roughly the same shape when they are allowed to refold in a test tube. But in fact, the technique strips away RNA-binding proteins, a process that can dramatically alter a molecule's structure.

To get at RNA structures *in vivo*, many scientists have turned to dimethylsulfate (DMS). This chemical penetrates cells, where it reacts with two of the four RNA nucleotides — adenine and cytosine — but only when they are in an unpaired state. Researchers then convert the RNA into DNA and sequence it. Any nucleotides that have been altered by DMS block the conversion into DNA, so scientists can use prematurely shortened bits of DNA to identify nucleotides that were unpaired.

Weissman and his colleagues deployed this method to analyse the full complement of mRNAs in yeast and humans, both in living cells and after the molecules had been extracted and allowed to refold<sup>2</sup>. “It was very exciting at first because we really didn't know what the differences would be *in vivo* and *in vitro*,” says Silvi Rouskin, a graduate student who worked on the project and is now at the Whitehead Institute in Cambridge, Massachusetts.

Many scientists had expected to see more

RNA folding inside a cell because they thought that interacting proteins would stabilize RNA structures there. But Weissman and his team saw the opposite. This, they now think, could be because mRNAs inside cells are actively generating proteins — and looser molecules are more available to the cell's protein-building machinery.

Bevilacqua and Assmann saw something curious when they used the DMS approach in their study of mRNA in *A. thaliana*<sup>1</sup>. mRNAs from genes that are involved in stress responses — ones activated during drought, say — tended to be folded more loosely inside a cell than predicted by computer modelling. By contrast, mRNAs of ‘housekeeping’ genes, which are involved in essential cell maintenance, mostly matched the predictions. The team proposes that stress-response RNAs are folded loosely so that they can shift shape easily inside a cell and thereby change the level of protein production in the face of changing conditions. By contrast, the housekeeping RNAs have to churn out relatively stable levels of protein. “That was just an amazing moment to see that dichotomy,” Assmann says.

The trouble with the DMS method is that it reveals the pairing of only two types of nucleotide, and computer modelling fills in the rest. To obtain pairing information for every letter of RNA inside the cell, Chang and Spitale adapted a structure-probing technique called SHAPE<sup>6</sup>. This allowed them to deduce the structures of more than 19,000 RNAs in mouse embryonic stem cells, an effort they published earlier this year<sup>7</sup>. The researchers showed that a common chemical modification to mRNA unfurls the molecule's structure, and they detected distinctive structural ‘signatures’ that predict where proteins will bind to control RNA shape.

Some researchers are already mulling over ways to put these revelations to use. Assmann and Bevilacqua are probing the structures of RNAs in rice, one of the world's most important staple foods, and plan to do the same for other agriculturally important plants. They would like to find ways to manipulate RNA shapes to improve stress tolerance and ultimately crop yield.

Rouskin, meanwhile, is looking at the RNAs of fruit flies to improve understanding of how these molecules' structures affect embryonic development. “Now we finally have the tools,” she says. “And we can ask all these questions that we never even thought about asking.” ■

**Elie Dolgin** is a science writer in Somerville, Massachusetts.

1. Ding, Y. *et al. Nature* **505**, 696–700 (2014).
2. Rouskin, S. *et al. Nature* **505**, 701–705 (2014).
3. Kertesz, M. *et al. Nature* **467**, 103–107 (2010).
4. Wan, Y. *et al. Nature* **505**, 706–709 (2014).
5. Kutchko, K. M. *et al. RNA* **21**, 1274–1285 (2015).
6. Spitale, R. C. *et al. Nature Chem. Biol.* **9**, 18–20 (2013).
7. Spitale, R. C. *et al. Nature* **519**, 486–490 (2015).