Received: 16 April 2020 | Revised: 3 September 2020

DOI: 10.1002/bies.202000027

METHODS, MODELS & TECHNIQUES

BioEssays WILEY

From structure to function: Route to understanding lncRNA mechanism

Johannes Graf | Markus Kretz

Institute of Biochemistry, Genetics and Microbiology, University of Regensburg, Regensburg, Germany

Correspondence
Markus Kretz, Institute of Biochemistry, Genetics and Microbiology, University of Regensburg, 93053 Regensburg, Germany.
Email: markus.kretz@vkl.uni-regensburg.de

Abstract

RNAs have emerged as a major target for diagnostics and therapeutics approaches. Regulatory nonprotein-coding RNAs (ncRNAs) in particular display remarkable versatility. They can fold into complex structures and interact with proteins, DNA, and other RNAs, thus modulating activity, localization, or interactome of multi-protein complexes. Thus, ncRNAs confer regulatory plasticity and represent a new layer of regulatory control. Interestingly, long noncoding RNAs (lncRNAs) tend to acquire complex secondary and tertiary structures and their function—in many cases—is dependent on structural conservation rather than primary sequence conservation.

Whereas for many proteins, structure and its associated function are closely connected, for lncRNAs, the structural domains that determine functionality and its interactome are still not well understood. Numerous approaches for analyzing the structural configuration of lncRNAs have been developed recently. Here, will provide an overview of major experimental approaches used in the field, and discuss the potential benefit of using combinatorial strategies to analyze lncRNA modes of action based on structural information.

KEYWORDS
lncRNA, long noncoding RNA

INTRODUCTION

With fundamental cellular functions ranging from energy metabolism to structural components, over signal transduction to being key regulators of gene expression, proteins were attributed great scientific attention, while—with a few exceptions—the RNA was contemplated as the inevitable intermediary required for protein production. However, this picture changed dramatically when high-throughput sequencing data revealed that more than two-thirds of the human genome are actively transcribed into RNA but only < 2% actually encodes for proteins.[1] Several classes of shortnon-coding RNAs (ncRNAs) controlling basic cellular functions such as translation (transfer RNAs, ribosomal RNAs), RNA editing (small nucleolar RNAs), or splicing (small nuclear RNAs) and have been known for quite a long time. More recently, short regulatory ncRNAs (20-30 nt in length), including microRNAs, endogenous short-interfering RNAs, or piwi-associated RNAs, acting as crucial regulators of gene expression were also identified.[2] Long noncoding RNAs (lncRNAs) have lately gained widespread attention and we are only at the beginning to understand their significant roles for a multitude of cellular processes.

Per definition, lncRNAs are over 200 nucleotides in length, lack a protein-coding potential, can be spliced, capped, and/or polyadenylated and are localized either in the nucleus or the cytoplasm of the cell.[1,3] LncRNAs are involved in a multitude of biological processes.

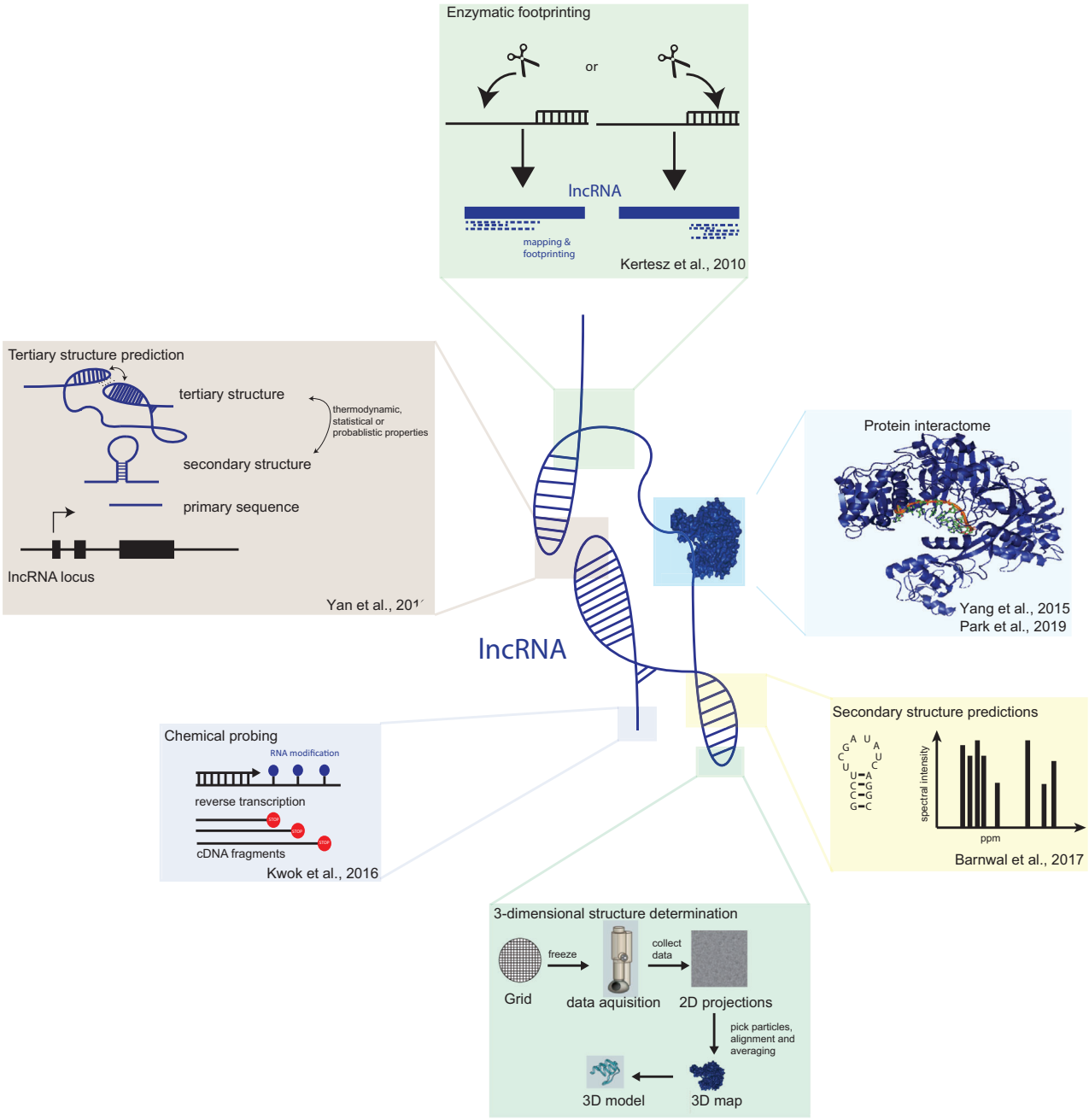

FIGURE 1 Overview of integrative experimental approaches to analyze long non-coding RNA structure

These include (a) interacting with chromatin complexes, and thus contributing to epigenetic gene regulation, (b) serving as modulators of proteins or multiprotein complexes, (c) binding DNA/RNA-associated proteins to regulate transcriptional expression, (d) regulating DNA stability through R-loop and triple helix formation, and (e) contributing to generation of a higher-order chromatin structure.[4]

Since lncRNAs are not only located in dedicated noncoding gene loci, but can be part of protein-coding genes, it became clear that understanding lncRNA function and modes of action is not only a matter for RNA biologists, but of wide interest for all fields of life science.

While a rapidly growing number of lncRNAs have been functionally characterized, studies unveiling modes of actions of lncRNAs are still sparse. This lack of knowledge might be due to several peculiarities inherent to many lncRNAs: on average, they display lower sequence conservation in comparison to proteins across species, showing only portions of conserved bases surrounded by large unconstrained sequences and thus complicating functional predictions based on primary sequence similarity.[5–8] A more likely conservation of ncRNAs includes conservation on (a) secondary or tertiary structure, (b) functional features, or (c) syntenic transcription.[9,10]

Additionally, low abundance of many lncRNAs adds further complexity to the prediction of modes of action. While several RNAs were found to have inherent catalytic activity, most lncRNAs function in concert with proteins, thus acting as part of ribonucleoprotein complexes

(RNPs) to maintain molecular function. Complexation in such particles strongly influences stability of these RNAs as well as broadens their functional spectrum.[11] Therefore, the majority of approaches to analyze modes of actions of lncRNAs are based on interaction studies with protein complexes, DNA, or RNA. For these reasons, most mechanistic studies are done using RNA-protein interaction techniques, which are largely based on structural insights from protein interaction partners, which were the focus of recent and comprehensive review articles.[12–15] Analysis of RNP-complexes has mostly been protein-centered due to the usually highly flexible structure of the RNA constituent. Here, we want to bring awareness to potential benefits of lncRNA-centric structure analyses as a future means to predict lncRNA modes of action based on structural features rather than primary sequence conservation.

For many proteins, structure and its functional relevance are closely connected, whereas for lncRNAs, the structural domains that determine functionality as well as define its interactome are still not well understood. Thus, gaining comprehensive understanding of lncRNA modes of actions might benefit from RNA structural analysis. Below, we will introduce some of the major approaches for analyzing lncRNA structure and, with the help of some exemplary lncRNAs, will discuss the potential benefit of using combinatorial approaches to analyze lncRNA modes of action based on structural information (Figure 1).

METHODS FOR STRUCTURE DETERMINATION OF lncRNAs

In recent years, a wide range of methods have been developed to assess the secondary structure of lncRNAs.[16] This selection of methods can be divided into experimental and computational methods and include, but are not entirely limited to enzymatic footprinting, chemical probing, nuclear magnetic resonance (NMR) spectroscopy, and comparative sequence analysis (Figure 1). The selection below is not meant to be exhaustive, but rather aims at outlining experimental and computational approaches frequently used and developed in the field. Each of these approaches have advantages as well as limitations, and can be combined with molecular biology approaches to gain insights into lncRNA modes of actions (Table 1).

Purification of lncRNAs for structure determination

The purification and analysis of lncRNAs poses a significant challenge for analyses of lncRNA structure, particularly if one wants to preserve the native, structural elements mediating function: Folding of long RNAs in vivo in many cases is not merely based on thermodynamic properties of the RNA, but requires the function of endogenous factors or RNA chaperones, which are not usually present in in vitro approaches.[46–48] Additionally, many RNA purification methods involve heat denaturation and refolding, which can result in misfolding and aggregation of lncRNAs.[49,50] Therefore, several different approaches that avoid RNA denaturation have been developed to overcome these issues in recent years. Most of those approaches utilize affinity tags, which are involved in the immobilization of the target RNAs and ribozymes, allowing for higher specificity during elution.[51,52] Another method developed by Chillón et al. allows for lncRNA purification without involvement of RNA denaturation and affinity tag design, thus aiming to better preserve lncRNA functional elements.[49] After T7 polymerase-based synthesis and removal of DNA and proteins, the desired RNA is obtained by ultrafiltration and purified using size-exclusion chromatography.[53] This semi-native purification protocol allows analysis of long RNAs with preserved cotranscriptional folding patterns as well as maintains potentially functional structural elements.[54] These studies are yielding fresh insights, such as the cooperative folding of functional RNAs or the investigation of new functional elements. Methods that can probe the structure of RNA under complex in vivo like conditions will enhance our understanding of in vivo RNA structural motifs.

Overcoming the challenges of purification of in vivo or in vivo like RNAs will allow the field to accomplish the ultimate goal, to understand how RNA folds in the cell, and thereby might help for the discovery of yet undiscovered mechanisms of lncRNAs.

Experimental methods

Within the last decade, numerous different probing techniques evolved to explore the secondary structures of lncRNAs. These can in general be divided into enzymatic and chemical probing approaches, which both can be combined with high-throughput sequencing and consecutive bioinformatics analysis. Three-dimensional (3-D) structural solution is the ultimate goal in macromolecular structure determination, as it provides atomic resolution data at the molecular level.

Enzymatic footprinting

Enzymatic footprinting is an in vitro approach designed to specifically cleave either single- or double-stranded nucleotides of radioisotope labeled RNAs using ribonucleases (RNases) with different specificities. It can thus be used to map single- versus double-stranded regions of RNA. The fragmented RNA products are typically analyzed alongside a sequencing or an alkaline hydrolysis RNA ladder by denaturing polyacrylamide gel electrophoresis.[17,18]

PARS (parallel analysis of RNA structure) further advances this approach by utilizing next-generation sequencing for the analysis of fragments generated with RNases of different specificities: RNase V1 is typically used to cut base-paired nucleotides and thus digests double-stranded regions within RNAs, while digestion with RNase S1 results in specific digestion of single-stranded RNA domains (Table 1). After a further fragmentation step, the RNA can be analyzed.[55] Therefore, the PARS method gives information about single- or double-stranded conformation on a transcriptome-wide scale.[55]

Frag-Seq (fragmentation sequencing) represents another alternative footprinting methodology, designed to specifically cleave single-stranded RNA, followed by analysis of the resulting fragments through high throughput sequencing.[56] Of note, the Fraq-Seq method appears to create a bias favoring small RNAs since RNase P1—the nuclease used in this method—preferentially cleaves RNAs < 200 nt without a further fragmentation step, thus leaving large RNAs underrepresented.[56] In comparison to PARS, the Fraq-Seq method is less time consuming and there is no need for an additional fragmentation step. Combining complementary data of PARS and Fraq-Seq could improve the accuracy of genome-wide RNA structure measurements.[57]

Chemical probing

Chemical probing makes use of chemicals, which covalently modify atoms of single stranded or flexible nucleotides. The high variety of chemicals (DMS, DEPC, Kethoxal, CMCT) allows probing in vivo and in vitro.[58,18] Reverse transcription introduces mutations into the cDNA, which subsequently can be mapped. Mostly chemical probing is used to measure nucleotide reactivity and thus structural analysis relies on subsequent modeling of the structure.

Recently, in vivo chemical probing strategies were developed, enabling analysis of RNA structure in its native cellular context including endogenous protein binding partners. SHAPE (selective 2′-hydroxyl acylation by primer extension) is based on selective acylation of all flexible and thus unpaired groups of nucleotides.[25] Acylation of nucleotides terminates reverse transcription reactions at modified sites and thus results in truncated cDNA fragments, which can be analyzed by sequencing. Several SHAPE reagents have been tested in order to improve the signal to background ratio,[27] as well as methodical improvements have been developed[21,59,26,60]:

SHAPE MaP[22–24] allows to identify RNA structures at single-nucleotide resolution using a combination of SHAPE-Seq and mutational profiling (MaP) techniques. The method can be customized to interrogate small RNAs, amplicons, or rare RNA species accurately in a mixture of RNAs. After SHAPE mutations are introduced, RT primers are selected depending on the RNA type of interest and the respective RNAs are reverse transcribed. Subsequently, induced mutations can be analyzed by high throughput sequencing.

The cell SHAPE method[28] was developed to provide accurate predictions of secondary structures by modifying non-base paired nucleotides in vivo. The modified RNA is labeled selectively with biotin, enabling purification by streptavidin pull-down. After pulldown, the fragmented RNA is reverse transcribed and can be measured with high throughput sequencing.

A recent development in structure probing methodology is PARIS (Psoralen Analysis of RNA Interactions and Structures), which relies on crosslinking of specific RNA base pairs with the highly specific and reversible nucleic acid crosslinker psoralen-derivative 4′-aminomethyltrioxsalen.[29] Subsequent partial RNA cleavage and complete proteinase digestion results in a set of small crosslinked and directly base-paired RNA fragments. High throughput sequencing reveals the direct base pairing between fragments. Based on these reads, models of RNA structures and interactions can be generated with high specificity and sensitivity.[29] This method allows for structure determination of intra- and inter-molecular RNA-RNA interaction in vivo as well as sequencing of the binding sites and thus helps to resolve the interaction sites.

Nuclear magnetic Resonance

Precise base-pairing information, however, still requires high-resolution three-dimensional RNA structures determined by X-ray crystallography or NMR spectroscopy. NMR spectroscopy is a versatile biophysical and structural biology technique that can readily probe base pairing and secondary structure in structured regions of an RNA and thus represents a very powerful tool to solve high-resolution structures and dynamics of RNAs at single base-pair resolution.[32] This information can be used to predict secondary structures. Spectral overlaps due to the similarity in the sugar backbone and line broadening from the movement of flexible RNA regions are the biggest drawbacks of NMR due to their high molecular weight.[30] To overcome this problem, the RNA can also be reduced to smaller substructures and used as restraints in larger models.[31]

Small angle scattering

Small angle scattering (SAS) is the collective name given to various techniques, including X-ray (SAXS) and neutron (SANS) scattering, employed to characterize biological macromolecules, including RNAs. Radiation is scattered by the randomly oriented sample in solution.[33] In SAXS, the scattering pattern describes the distribution of electron density that interacts with X-rays and can be analyzed to provide information about the averaged particle sizes and shapes.[34,35] The effect of averaging many randomly oriented molecules is equivalent to averaging all directions of one molecule, hence for example chirality cannot be determined.[34] Therefore, SAXS methods require combination with other experimental approaches and/or with computational modeling to support the results.

Atomic force microscopy

Atomic force microscopy (AFM) is a type of scanning probe microscopy (SPM), which can be used for various applications, including imaging and force-probing biological samples. AFM instruments collect data to generate images by physically touching the surface of the samples. In AFM, a molecule is immobilized on a solid substrate, and subsequently, the surface is scanned with high precision by an ultra-sharp tip, mounted at the end of a force-sensing cantilever.[38] The tapping mode AFM, in which the cantilever is oscillated close to the sample

TABLE 1 Tools for determining secondary structures of RNA

	Tools for determining secondary structure of RNAs	Description of target	Reference
Experimental methods			
Enzymatic cleavage, footprinting	S1 nuclease (in vitro)	Cleaves all single stranded nucleotides	[17,18]
	RNAse V1 (in vitro)	Cleaves double stranded RNA	
	RNAse A (in vitro)	Cleaves 3′ of single stranded C, U	
	RNAse T1 (in vitro)	Cleaves 3′ of single stranded G	
	RNAse T2 (in vitro)	Cleaves 3′ of single stranded N (with a preference for A)	
	Hydroxyl radicals (in vitro)	Degradation of ribose backbone based on solvent accessibility	
	In-line probing (in vitro)	RNA allowed to degrade over time, single-stranded regions typically degrade faster than structured regions	
Chemical probing	DMS (in vitro, in vivo)	Methylation of single stranded A, C	[17.19,18]
	DEPC (in vitro)	Modifies single stranded A	
	Kethoxal (in vitro)	Modifies single stranded G	
	CMCT (in vitro)	Acylation of single stranded U and G	
	SHAPE (in vitro, in vivo)	Acylation of flexible 2′OH ribose groups, performed with several different molecules	[21–25]
	SHAPE-MaP	Acylation of flexible 2′OH ribose groups, performed with several different molecules, followed by mutational profiling	[26–28,25]
	IcSHAPE	Addition of a custom 2-methylnicotinic acid imidazolide probe in vivo to mark it selectively for biotin tagging	[27,28]
	PARIS (in vitro, in vivo)	Reversible crosslinking of base paired nucleotides	[29]
NMR		probe base pairing and secondary structure	[30–32]
Small angle scattering		provides low-resolution structural information about macromolecules in solution	[33–35]
Atomic force microscopy		enables direct visualization by physically probing native, large molecules under biological conditions	[36–38]
Cryo-EM		utilize a transmission electron microscope to determine the structures of frozen-hydrated samples	[39–41]
Computational methods			
Structural modeling		Prediction of RNA secondary structures based on thermodynamic, statistical, or probabilistic properties	[42,43]
Comparative sequence analysis		Detection of compensatory mutations, which allow an RNA to retain its structure and function, despite evolutionary variation in primary sequence	[44,45]

surface without actually "touching" it, found widespread application in structural biology, and in particular it has been successfully applied to image fragile biological objects, including RNA molecules.[37] AFM has been used to study nucleic acid structures in both fluid and air.[36,61] While AFM is limited in resolution compared with X-ray crystallography or NMR, it enables direct visualization by physically probing native, large molecules under biological conditions. AFM also allows statistical analysis of structurally diverse molecules, as shown for lncRNAs.[62] Structure determination requires computational processing of AFM images. In the case of complex RNA molecules, one of the goals is to identify the RNA sugar-phosphate backbone structures, from which information can be extracted about the presence of double- and single-stranded regions, their connectivity, and mutual orientations.[63] Recently, new protocols were developed to characterize 3D topology of full length lncRNAs in biologically active conformations.[61]

Cryo-electron microscopy

Cryo-electron microscopy (cryo-EM) includes a variety of techniques that utilize a transmission electron microscope to determine the structures of frozen-hydrated samples.[41] In single-particle cryo-EM, images of individual biomolecules are collected and subsequently pro-

cessed to generate 3D reconstructions of biomolecules.[64] In cryo-electron tomography, images of biomolecules within their native cellular environment are collected; however, the resolution in this method is substantially lower than that of all other cryo-EM methods.[39] Single-particle cryo-electron microscopy may provide a new approach to RNA structure determination. Recent advances in the technique allow for high-resolution structure determination of proteins and large RNA–protein complexes. However, functional noncoding RNA molecules that are not part of large RNA–protein complexes are in many cases either too small or too conformationally diverse to characterize with cryo-EM. However, recent improvements of cryo-EM in combination with high-throughput biochemistry and computational 3D structure modeling allowed for RNA-only structure determination.[40]

Computational Methods

Since it was not feasible to experimentally determine the secondary structure of a large amount of long sequences, the in silico prediction was the only realistic source for studying RNA structure for a long time. Common approaches for investigating the secondary structure are comparative methods. They compare evolutionary conserved regions and identify structural elements. This method requires a large input of conserved regions, thus limiting its usability for sparsely conserved lncRNAs. In these cases, the best alternative for computational prediction is thermodynamics.

Comparative sequence analysis

Comparative sequence analysis (also referred to as phylogenetic analysis) is used to detect compensatory mutations, which allow an RNA to retain its structure and function, despite evolutionary variation in primary sequence,[44] thus identifying functional structural elements of RNAs based on evolutionary conservation. Comparative sequence analysis algorithms generate models by aligning RNA sequences based on sequence conservation and single-sequence structure prediction.[65] Since an RNA helix can be formed from two sets of sequences that are not identical to each other, a search for positions (in an alignment of homologous sequences) that covary to maintain Watson-Crick pairing within a potential secondary helix is necessary. The search for coordinated base substitutions in an alignment of homologous sequences is called covariation analysis.[66] If conservation in base pairing could be established, it would provide a powerful indicator that RNA structure plays a role in aspects of lncRNA function, although available tools for structure identification are limited and still have to be improved.[67,45,68] A major challenge is the lack of sequences, which complicates functional predictions based on primary sequence similarity. Therefore, lack of covariation signal in these alignments makes structure prediction for lncRNAs difficult.[45,68]

Structure modeling

Combination of enzymatic footprinting and chemical probing can be used as restraints for RNA folding algorithms designed to predict RNA secondary structure based on thermodynamic, statistical, or probabilistic properties to obtain secondary structure models with lowest free energy. The sole usage of RNA folding approaches is in many cases not sufficient for prediction of lncRNA structures, as they usually cannot take into account transacting factors such as interacting proteins, other RNAs, and small ligands, as well as other physiological conditions.[69] Consequently, they are mostly used to generate in silico secondary structural predictions of a given RNA alone. In addition, computational predictions do not work very well for large RNAs with complex structural elements such as pseudoknots, kissing loops, or long-range interactions.[42] Therefore, a commonly used approach for RNA secondary structure prediction is to divide the RNA in substructures and then employing scoring functions to evaluate the probability of folding into these conformations.[43] Applying structural restraints based on structural features obtained experimentally, based on approaches described above, therefore tremendously improves these scoring functions.[70,71,43]

Hybrid methods

The high variety of lncRNAs in terms of structural flexibility, interaction with other complexes, and stability of these complexes highlights the need for a careful optimization of methods, as well as selection of appropriate structure determination and modeling approaches to match the characteristics of the individual target. Experimental determination of RNA structures at precise atomic resolution is very difficult and rather expensive, whereas sequence-based structure determination is much faster and less expensive. Given the disproportional growth of RNA sequence information in contrast to the limited gain of structural information, it seems unlikely that a majority of lncRNA structures will be solved in the near future solely based on experimental approaches. To address this problem, computational methods were developed to either simulate the process of RNA structure formation or use information derived from other known RNA structures. Many computational methods, however, suffer from various limitations rendering them less reliable for structure prediction of long RNAs. While a few structures could be solved for lncRNAs, the field is still suffering from an incomplete understanding of RNA tertiary structures. In many cases, the limitations of computational and experimental methods can be overcome by combining these two complementary approaches. Individual usage of approaches mentioned above might not entirely capture the dynamic nature of lncRNA structures, whereas combined complementary data of different approaches has improved the accuracy of structural predictions and overcome the weaknesses of separately applied methods or exceed the limitations of a single method in a number of cases.[72,31,73,34,71,74] While in general, experimentally determined structural information can be used as restraint

in computational structure prediction models, not all experimental and computational approaches are compatible with each other.[74] Many of these numerous challenges for connecting experimental and computational methods to determine structures and functions of lncRNAs were covered in detail in recent articles.[75,69,76]

FROM lncRNA STRUCTURE TO FUNCTION

Structure determination of lncRNAs in vivo is very challenging due to their size and high degree of heterogeneity. Typically, lncRNAs display poor conservation across species, harboring only portions of conserved bases surrounded by large unconstrained sequences, which makes functional predictions purely based on structural prediction hardly feasible. Nevertheless, links between structure and function of lncRNAs are emerging and lead to the "RNA modular code" hypothesis, whereby lncRNAs contain distinct structural domains that fold into specific scaffolds, which fulfill distinct functions or interact with specific protein complexes.[77,78]

Structural features of some lncRNAs have been experimentally determined and led to prediction of their modes of action, some of which are summarized below:

Xist

Xist, one of the most well-studied lncRNAs, is a 17-kb transcript responsible for dosage compensation in placental mammals. Xist RNA coats the inactive X chromosome during early development and thus represses its transcription.[78] This occurs with the aid of several protein binding events that take place along the length of the transcript, and which have been well investigated.[80,79] Chromosomal silencing is mediated by domains on the 5′-end of Xist RNA, called A-repeats. In their absence, coating of the inactive X chromosome by Xist still occurs; epigenetic silencing, however, is abolished.[80,81] Since A-repeats serve such a vital role in X chromosome inactivation, their secondary structure has been extensively investigated, but remains not fully understood, as at least eight different structural models have been proposed.[84,79,23] A second structural element, the highly repetitive GC-rich B-repeats, can be found in Xist and was shown to be involved in the recruitment of the chromatin remodeling complex polycomb repressive complex 1 (PRC1).[82,79] Studies also identified many other proteins binding directly to Xist including SHARP and others.[80,86] Recent research suggests a model where Xist initiates compartmentalization and concentration dependent phase separation in vivo.[83,84,80] Induction of Xist in this model *Xist l*eads to a locally higher concentration of many Xist-binding proteins, which can interact with each other. Interaction of many RNA-binding proteins was mediated over repetitive sequences in Xist.[85,90,29,86] This model explains an essential role for Xist itself in initiation of transcriptional repression. Recent studies discovered a biphasic process of the A- and B-repeats where the A-repeats initiate recruitment of the Polycomb complex and the B-repeats subsequently stabilizes them.[87] Further structural elements could be linked to chromatin association of Xist as well as interacting elements with additional RNA-binding proteins.[88] Even though Xist has been extensively studied, many structural modules are still not completely understood.[89,90]

RepA

RepA is a mouse lncRNA 1600 nucleotides in size. It is encoded on the sense strand of the Xist-gene and has been proposed to recruit the histone methyltransferase of polycomb repressive complex 2 (PRC2) to the future inactive X chromosome prior to the expression of Xist.[91] Additionally, RepA appears to upregulate the expression of Xist, which then initiates and spreads silencing across the inactive X-chromosome.[91] SHAPE and DMS chemical probing of RepA in vitro identified a defined tertiary architecture that can form autonomously in the absence of protein partners.[97] This highlights the importance of distinct structural elements necessary for the folding of RNAs without any additional structural restraints of proteins. Phylogenetic analysis and computational 3D modeling supported these findings, thus potentially the solved structure represents the functional state of RepA/Xist and helps to design targeted studies of specific structural elements that will increase the understanding of the complex process of X-inactivation.[97]

HOTAIR

HOTAIR is 2148 nucleotides long lncRNA associated with Sporadic Thoracic Aortic Aneurysm through regulation of extracellular matrix deposition and apoptosis of human aortic smooth muscle cells.[92] The 5t'end of HOTAIR interacts with chromatin remodeling complexes such as polycomb repressive complex 2 (PRC2) or the LSD1/ REST/CoREST complex[78] and leads to silencing of genes on the HoxD cluster. Structural features of HOTAIR were analyzed in vitro using chemical probing, thus revealing four independent structure modules.[93] These experimentally obtained structures correspond to bioinformatically predicted PRC2-binding domains, and thus verify the previously observed interaction between HOTAIR and PRC2. Binding of HOTAIR to PRC2 is mediated by a G-rich region, which closely resembles a putative quadruplex forming sequence and thus differs from repetitive PRC2 binding elements found in XistA-repeats.[100] This lack of consensus may reflect the existence of several lncRNA-PRC2 binding modes, which may each correspond to distinct functions of the PRC2 complex. More recently using the AFM technique, distinct shapes of HOTAIR could be observed.[94]

MEG3

The MEG3 lncRNA consists of three distinct structure modules and represents one of the best-characterized tumor suppressor lncRNAs. Loss-of-function analyses showed that module 2 and 3 are

important for p53 activation.[95] Mechanistic studies have revealed a role for Meg3 in epigenetic regulation by interacting with chromatin-modifying complexes such as PRC2. Based on these findings, MEG3 appears to act as a guide for PRC2 to target chromatin regions in a process involving triplex formation of the lncRNA with the genomic DNA at target sites.[96] SHAPE-MaP and comparative sequence analysis show strong structural conservation of functionally important motifs between species, which overlap partially with the experimentally observed binding sites of PRC2, thus suggesting a conservation of folding motifs.[104] These macroscopic structural motifs were recently further confirmed as highly conserved structured p53-binding domains.[97]

Resolving the entire structure of MEG3 and characterizing its interactions based on structural predictions might help to improve our understanding of its role in tumorigenesis and provide a framework for RNA-based anticancer therapies.

Braveheart

Braveheart is a 590-nucleotide long lncRNA that regulates cardiovascular lineage commitment by interaction with the Zinc-finger motif-containing transcription factor CNBP. The secondary structure of Braveheart was experimentally assessed using SHAPE and DMS probing in vitro. The lncRNA is organized into a modular structure comprising three domains, consisting of 12 helices, eight terminal loops, five sizeable internal loops, and a five-way junction.[98] This 5′ asymmetric G-rich internal loop (RHT/AGIL motif) in vivo is necessary for the interaction with CNBP. This finding is reinforced by recent in vitro SAXS studies,[99] which also suggest a necessary combinatorial binding of the RHT/AGIL motif and other structural elements to CNBP.

As shown above, structural analysis could help to find yet uncharacterized functional relevant modules as well as protein-binding domains in several lncRNAs as shown for HOTAIR and Braveheart.

CONCLUSION: FROM INTERACTOME STUDIES TO STRUCTUROME STUDIES

For proteins, structural methods have been widely used to predict functions and mechanisms. Similar to proteins, some lncRNAs contain only a single domain, while others appear to harbor multiple distinct domain structures. Some of these domains have a clearly defined function associated with them. Such domains, especially those that are recurrent and conserved, might represent crucial functional elements of lncRNAs. As mentioned above, several such examples have already been discovered and helped to elucidate lncRNA modes of action. Therefore, it is of high importance to combine structure determination with genetic manipulations to elucidate the functional domains of lncRNAs. Once increasing numbers of lncRNAs are structurally analyzed, mapping of domain structures might also be a tool to predict function or mechanisms of novel and uncharacterized lncRNAs. Moreover, with recent innovations in CRISPR/Cas9 technologies, it became feasible to introduce targeted modifications in lncRNA gene loci in a streamlined fashion and therefore alter the structure and associated functions of endogenous lncRNAs. Inserting or deleting specific structural domains can lead to gain-of-function or loss-of-function lncRNAs and subsequently results in a more comprehensive understanding of structural organization of lncRNAs and their impact on modes of action.

Thus, combining structural knowledge with functional data can subsequently enable the discovery of yet undiscovered mechanisms of lncRNAs. Resolving lncRNA structures and characterizing its interactions with cellular binding partners will improve our understanding of lncRNAs in general as well as their role in disease and cancer and might provide frameworks for RNA based therapies.

The presence of multiple dynamic conformations may need to be considered, including structural changes induced by interactions with other molecules. In cells, continuous dynamic rearrangements of lncRNA conformations could appear in processes, including cell division, tissue development, or homeostasis. Therefore, the function of a lncRNA might not merely be defined by its abundance and stability, but could largely be dependent on the current structure conformation—similar to proteins with enzymatic activity. One of the best examples demonstrating the specificity and dynamics of RNA structures is a riboswitch. Riboswitches are RNA sensors that can detect changes in cellular stimuli in the absence of other cofactors, such as proteins.[100,101] As such, some of the first riboswitches were discovered based on changes in RNA structure that had been induced by specific ligands. Furthermore, catalytic RNAs (also known as ribozymes), such as ribosomal RNA (rRNA) and transfer RNA (tRNA), play a role in various biological processes, including RNA processing and protein synthesis.[101,2] Given that many lncRNAs appear to contain complex secondary and tertiary structures, and some of them where already shown to be important for its function and regulation, it stands to reason that many lncRNAs might act based on dynamic regulation of structure conformations.

The development of additional integrative in vivo structural approaches combined with molecular biology/biochemical analyses of lncRNA function is necessary to understand the additional layer of complexity introduced through structure–function dynamics of at least a subset of lncRNAs. Structure, dynamics, and protein interactions can be predicted by combinatorial structural analysis and soon could allow prediction of lncRNA modes of action based on the presence of known functional domains

Given the sheer volume of lncRNAs identified in the human genome, it will be crucial to extend the development of efficient approaches for determining the structure and dynamics of large RNAs. This will speed up the large-scale functional characterization, thus eventually enabling us to predict lncRNA modes of action based on structural features.

ACKNOWLEDGMENTS

We thank Dr. Sonja Hombach and Fabian Schabenberger for critical reading of the manuscript. Our research is supported by the Deutsche Forschungsgemeinschaft (SFB 960 to M.K.).

Open access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ORCID

Markus Kretz https://orcid.org/0000-0003-3420-3925

REFERENCES

1. Djebali, S., Davis, C. A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., … Gingeras, T. R. (2012). Landscape of transcription in human cells. *Nature*, *489*, 101–108. https://doi.org/10.1038/nature11233
2. Matera, A. G, Terns, R. M., & Terns, M. P. (2007). Non-coding RNAs: Lessons from the small nuclear and small nucleolar RNAs. *Nat. Rev. Mol. Cell Biol.*, *8*, 209–220. https://doi.org/10.1038/nrm2124
3. Fatica, A., & Bozzoni, I. (2014). Long non-coding RNAs: New players in cell differentiation and development. *Nat. Rev. Genet.*, *15*, 7–21. https://doi.org/10.1038/nrg3606
4. Kopp, F., & Mendell, J. T. (2018). Functional classification and experimental dissection of long noncoding RNAs. *Cell*, *172*, 393–407. https://doi.org/10.1016/j.cell.2018.01.011
5. Guttman, M., Amit, I., Garber, M., French, C., Lin, M. F., Feldser, D., … Lander, E. S. (2009). Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature*, *458*, 223–227. https://doi.org/10.1038/nature07672
6. Novikova, I. V., Hennelly, S. P., Tung, C.-S., & Sanbonmatsu, K. Y. (2013b). Rise of the RNA machines: Exploring the structure of long non-coding RNAs. *J. Mol. Biol.*, *425*, 3731–3746. https://doi.org/10.1016/j.jmb.2013.02.030
7. Ponjavic, J., Ponting, C. P., & Lunter, G. (2007). Functionality or transcriptional noise? Evidence for selection within long noncoding RNAs. *Genome Res.*, *17*, 556–565. https://doi.org/10.1101/gr.6036807
8. Ponting, C. P., Oliver, P. L., & Reik, W. (2009). Evolution and functions of long noncoding RNAs. *Cell*, *136*, 629–641. https://doi.org/10.1016/j.cell.2009.02.006
9. Diederichs, S. (2014). The four dimensions of noncoding RNA conservation. *Trends Genet.*, *30*, 121–123. https://doi.org/10.1016/j.tig.2014.01.004
10. Tavares, R .C. A., Pyle, A. M., & Somarowthu, S. (2019a). Phylogenetic analysis with improved parameters reveals conservation in lncRNA structures. *J. Mol. Biol.*, *431*, 1592–1603. https://doi.org/10.1016/j.jmb.2019.03.012
11. Yang, Y., Wen, L., & Zhu, H. (2015a). Unveiling the hidden function of long non-coding RNA by identifying its major partner-protein. *Cell Biosci.*, *5*:59 https://doi.org/10.1186/s13578-015-0050-x
12. Blythe, A. J., Fox, A. H., & Bond, C. S. (2016). The ins and outs of lncRNA structure: How, why and what comes next? *Biochim. Biophys. Acta (BBA) – Gene Regulatory Mechanisms*, *1859*, 46–58. https://doi.org/10.1016/j.bbagrm.2015.08.009
13. Dimitrova-Paternoga, L., Jagtap, P. K. A., Chen, Po-C, & Hennig, J. (2020). Integrative structural biology of protein-RNA complexes. *Structure (London, England)*, *28*, 6–28. https://doi.org/10.1016/j.str.2019.11.017
14. Jones, A. N., & Sattler, M. (2019). Challenges and perspectives for structural biology of lncRNAs—the example of the Xist lncRNA A-repeats. *J. Mol. Cell Biol.*, *11*, 845–859. https://doi.org/10.1093/jmcb/mjz086
15. Schlundt, A., Tants, J.-N., & Sattler, M. (2017). Integrated structural biology to unravel molecular mechanisms of protein-RNA recognition. *Methods (San Diego, Calif.), 118-119*:119–136. https://doi.org/10.1016/j.ymeth.2017.03.015
16. Seetin, M. G., & Mathews, D. H. (2012). RNA structure prediction: An overview of methods. In: Keiler, K.C. (ed) *Bacterial Regulatory RNA*. Humana Press, Totowa, NJ, pp 99–122
17. Herschlag, D. (1995). RNA chaperones and the RNA folding problem. *J. Biol. Chem.*, *270*, 20871–20874. https://doi.org/10.1074/jbc.270.36.20871
18. Rajkowitsch, L., Chen, D., Stampfl, S., Semrad, K., Waldsich, C., Mayer, O., … Schroeder, R. (2007). RNA chaperones, RNA annealers and RNA helicases. *RNA Biology*, *4*, 118–130. https://doi.org/10.4161/rna.4.3.5445
19. Weeks, K. M. (2010). Advances in RNA structure analysis by chemical probing. *Curr. Opin. Struct. Biol.*, *20*, 295–304. https://doi.org/10.1016/j.sbi.2010.04.001
20. Storz, G., & Papenfort, K. (eds) (2019). Proteins that chaperone RNA regulation. In: *Regulating with RNA in Bacteria and Archaea*. American Society of Microbiology, pp 385–397
21. Chillón, I., Marcia, M., Legiewicz, M., Liu, F.., Somarowthu, S., & Pyle, A. M. (2015a). Native purification and analysis of long RNAs. In: *Meth. Enzymol.*, Elsevier, pp 3–37
22. Somarowthu, S., Legiewicz, M., Chillón, I., Marcia, M., Liu, F., & Pyle, A. M. (2015a). HOTAIR forms an intricate and modular secondary structure. *Mol. Cell*, *58*, 353–361. https://doi.org/10.1016/j.molcel.2015.03.006
23. Batey, R. T., & Kieft, J. S. (2007). Improved native affinity purification of RNA. *RNA*, *13*, 1384–1389. https://doi.org/10.1261/rna.528007
24. Kanwal, F., & Lu, C. (2019). A review on native and denaturing purification methods for non-coding RNA (ncRNA). *J. Chromatogr. B*, *1120*, 71–79. https://doi.org/10.1016/j.jchromb.2019.04.034
25. Chillón, I., Marcia, M., Legiewicz, M., Liu, F.., Somarowthu, S., & Pyle, A. M. (2015b). Native purification and analysis of long RNAs. In: *Meth. Enzymol.*, Elsevier, pp 3–37
26. Adams, R. L., Huston, N. C., Tavares, R. C. A., & Pyle, A. M. (2019). Sensitive detection of structural features and rearrangements in long, structured RNA molecules. In: *Meth. Enzymol.*, Elsevier, pp 249–289
27. Stern, S., Moazed, D., & Noller, H. F. (1988). [33]Structural analysis of RNA using chemical and enzymatic probing monitored by primer extension. In: *Meth. Enzymol.*, Elsevier, pp 481–489
28. Ziehler, W. A., & Engelke, D. R. (2000). Probing RNA structure with chemical reagents and enzymes. *Curr. Protoc. Nucleic Acid Chem.*, *00*, 6.1.1–6.1.21, 00:6.1.1-6.1.21. https://doi.org/10.1002/0471142700.nc0601s00
29. Kertesz, M., Wan, Y., Mazor, E., Rinn, J. L., Nutter, R. C., Chang, H. Y., & Segal, E. (2010). Genome-wide measurement of RNA secondary structure in yeast. *Nature*, *467*, 103–107. https://doi.org/10.1038/nature09322
30. Underwood, J. G., Uzilov, A. V., Katzman, S., Onodera, C. S., Mainzer, J. E., Mathews, D. H., … Haussler, D. (2010). FragSeq: Transcriptome-wide RNA structure probing using high-throughput sequencing. *Nat. Methods*, *7*, 995–1001. https://doi.org/10.1038/nmeth.1529
31. Wan, Y., Kertesz, M., Spitale, R. C., Segal, E., & Chang, H. Y. (2011). Understanding the transcriptome through RNA structure. *Nat. Rev. Genet.*, *12*, 641–655. https://doi.org/10.1038/nrg3049
32. Ding, Y., Tang, Y., Kwok, C. K., Zhang, Yu, Bevilacqua, P. C., & Assmann, S. M. (2014). In vivo genome-wide profiling of RNA secondary structure reveals novel regulatory features. *Nature*, *505*, 696–700. https://doi.org/10.1038/nature12756
33. Wilkinson, K. A., Merino, E. J., & Weeks, K. M. (2006). Selective 2â€2-hydroxyl acylation analyzed by primer extension (SHAPE): Quantitative RNA structure analysis at single nucleotide resolution. *Nat. Protoc.*, *1*, 1610–1616. https://doi.org/10.1038/nprot.2006.249
34. Lee, B., Flynn, R. A., Kadina, A., Guo, J. K., Kool, E. T., & Chang, H. Y. (2017). Comparison of SHAPE reagents for mapping *RNA* structures inside living cells. *RNA*, *23*, 169–174. https://doi.org/10.1261/rna.058784.116
35. Busan, S., Weidmann, C. A., Sengupta, A., & Weeks, K. M. (2019). Guidelines for SHAPE reagent choice and detection strategy for RNA

structure probing studies. *Biochemistry*, *58*, 2655–2664. https://doi.org/10.1021/acs.biochem.8b01218
36. Homan, P. J., Favorov, O. V., Lavender, C. A., Kursun, O., Ge, X., Busan, S., … Weeks, K. M. (2014). Single-molecule correlated chemical probing of RNA. *Proc. Natl. Acad. Sci. USA*, *111*, 13858–13863. https://doi.org/10.1073/pnas.1407306111
37. Kwok, C. K., Sahakyan, A. B., & Balasubramanian, S. (2016). Structural analysis using SHALiPE to reveal RNA G-quadruplex formation in human precursor MicroRNA. *Angew. Chem., Int. Ed.*, *55*, 8958–8961. https://doi.org/10.1002/anie.201603562
38. Smola, M. J., & Weeks, K. M. (2018). In-cell RNA structure probing with SHAPE-MaP. *Nat. Protoc.*, *13*, 1181–1195. https://doi.org/10.1038/nprot.2018.010
39. Siegfried, N. A., Busan, S., Rice, G. M., Nelson, J. A. E., & Weeks, K. M. (2014). RNA motif discovery by SHAPE and mutational profiling (SHAPE-MaP). *Nat. Methods*, *11*, 959–965. https://doi.org/10.1038/nmeth.3029
40. Smola, M. J., Christy, T. W., Inoue, K., Nicholson, C. O., Friedersdorf, M., Keene, J. D., … Weeks, K. M. (2016). SHAPE reveals transcript-wide interactions, complex structural domains, and protein interactions across the Xist lncRNA in living cells. *Proc. Natl. Acad. Sci. U.S.A.*, *113*, 10322–10327. https://doi.org/10.1073/pnas.1600008113
41. Smola, M. J., Rice, G. M., Busan, S., Siegfried, N. A., & Weeks, K. M. (2015). Selective 2′-hydroxyl acylation analyzed by primer extension and mutational profiling (SHAPE-MaP) for direct, versatile and accurate RNA structure analysis. *Nat. Protoc.*, *10*, 1643–1669. https://doi.org/10.1038/nprot.2015.103
42. Spitale, R. C., Flynn, R. A., Zhang, Q. C., Crisalli, P., Lee, B., Jung, J.-W., … Chang, H. Y. (2015). Structural imprints in vivo decode RNA regulatory mechanisms. *Nature*, *519*, 486–490. https://doi.org/10.1038/nature14263
43. Lu, Z., Zhang, Q. C., Lee, B., Flynn, R. A., Smith, M. A., Robinson, J. T., … Chang, H. Y. (2016). RNA duplex map in living cells reveals higher-order transcriptome structure. *Cell*, *165*, 1267–1279. https://doi.org/10.1016/j.cell.2016.04.028
44. Xue, Y., Kellogg, D., Kimsey, I. J., Sathyamoorthy, B., Stein, Z. W., McBrairty, M., & Al-Hashimi, H. M. (2015). Characterizing RNA excited states using NMR relaxation dispersion. *Meth Enzymol*, *558*, 39–73. https://doi.org/10.1016/bs.mie.2015.02.002
45. Barnwal, R. P., Yang, F., & Varani, G. (2017). Applications of NMR to structure determination of RNAs large and small. *Arch. Biochem. Biophys.*, *628*, 42–56. https://doi.org/10.1016/j.abb.2017.06.003
46. Kotar, A., Foley, H. N., Baughman, K. M., & Keane, S. C. (2020). Advanced approaches for elucidating structures of large RNAs using NMR spectroscopy and complementary methods. *Methods S104620231930235X*, https://doi.org/10.1016/j.ymeth.2020.01.009
47. Liu, Li, Boldon, L., Urquhart, M., & Wang, X. (2013). Small and wide angle X-ray scattering studies of biological macromolecules in solution. *JoVE*, 4160. https://doi.org/10.3791/4160
48. Plumridge, A., Katz, A. M., Calvey, G. D., Elber, R., Kirmizialtin, S., & Pollack, L. (2018). Revealing the distinct folding phases of an RNA three-helix junction. *Nucleic Acids Res.*, *46*, 7354–7365. https://doi.org/10.1093/nar/gky363
49. Svergun, D. I., & Koch, M. H. J. (2003). Small-angle scattering studies of biological macromolecules in solution. *Rep. Prog. Phys.*, *66*, 1735–1782. https://doi.org/10.1088/0034-4885/66/10/R05
50. Shahin, V., & Barrera, N. P. (2008). Providing unique insight into cell biology via atomic force microscopy. In: *Int. Rev. Cytol. Suppl.*, Elsevier, pp 227–252
51. Schön, P. (2018). Atomic force microscopy of RNA: State of the art and recent advancements. *Semin. Cell Dev. Biol.*, *73*, 209–219. https://doi.org/10.1016/j.semcdb.2017.08.040
52. Lyubchenko, Y. L., Shlyakhtenko, L. S., & Ando, T. (2011). Imaging of nucleic acids with atomic force microscopy. *Methods (San Diego, Calif.)*, *54*, 274–283. https://doi.org/10.1016/j.ymeth.2011.02.001
53. Uroda, T., Chillón, I., Annibale, P., Teulon, J.-M., Pessey, O., Karuppasamy, M., … Marcia, M. (2020). Visualizing the functional 3D shape and topography of long noncoding RNAs by single-particle atomic force microscopy and in-solution hydrodynamic techniques. *Nat. Protoc.*, *15*, 2107–2139. https://doi.org/10.1038/s41596-020-0323-7
54. Novikova, I., Hennelly, S., & Sanbonmatsu, K. (2013a). Tackling structures of long noncoding RNAs. *IJMS*, *14*, 23672–23684. https://doi.org/10.3390/ijms141223672
55. Gilmore, J. L., Yoshida, A., Hejna, J. A., & Takeyasu, K. (2017). Visualization of conformational variability in the domains of long single-stranded RNA molecules. *Nucleic Acids Res.*, *45*, 8493–8507. https://doi.org/10.1093/nar/gkx502
56. Nannenga, B. L., & Gonen, T. (2019). The cryo-EM method microcrystal electron diffraction (MicroED). *Nat. Methods*, *16*, 369–379. https://doi.org/10.1038/s41592-019-0395-x
57. Cheng, Y., Grigorieff, N., Penczek, P. A., & Walz, T. (2015). A primer to single-particle cryo-electron microscopy. *Cell*, *161*, 438–449. https://doi.org/10.1016/j.cell.2015.03.050
58. Beck, M., & Baumeister, W. (2016). Cryo-electron tomography: Can it reveal the molecular sociology of cells in atomic detail? *Trends Cell Biol.*, *26*, 825–837. https://doi.org/10.1016/j.tcb.2016.08.006
59. Kappel, K., Zhang, K., Su, Z., Watkins, A. M., Kladwang, W., Li, S., … Das, R. (2020). Accelerated cryo-EM-guided determination of three-dimensional RNA-only structures. *Nat. Methods*, *17*, 699–707. https://doi.org/10.1038/s41592-020-0878-9
60. Eddy, S. R. (2014). Computational analysis of conserved RNA secondary structure in transcriptomes and genomes. *Annu. Rev. Biophys.*, *43*, 433–456. https://doi.org/10.1146/annurev-biophys-051013-022950
61. Jaeger, J. A., Turner, D. H., & Zuker, M. (1989). Improved predictions of secondary structures for RNA. *Proc. Natl. Acad. Sci. USA*, *86*, 7706–7710. https://doi.org/10.1073/pnas.86.20.7706
62. Parsch, J., Braverman, J. M., & Stephan, W. (2000). Comparative sequence analysis and patterns of covariation in RNA secondary structures. *Genetics*, *154*:909–921
63. Reuter, J. S., & Mathews, D. H. (2010). RNAstructure: Software for RNA secondary structure prediction and analysis. *BMC Bioinformatics*, *11*:129 https://doi.org/10.1186/1471-2105-11-129
64. Rivas, E., Clements, J., & Eddy, S. R. (2020). Estimating the power of sequence covariation for detecting conserved RNA structure. *Bioinformatics*, *36*, 3072–3076. https://doi.org/10.1093/bioinformatics/btaa080
65. Tavares, R. C. A., Pyle, A. M., & Somarowthu, S. (2019b). Phylogenetic analysis with improved parameters reveals conservation in lncRNA structures. *J. Mol. Biol.*, *431*, 1592–1603. https://doi.org/10.1016/j.jmb.2019.03.012
66. Ponce-Salvatierra, A., Astha, Merdas, K., Nithin, C., Ghosh, P., Mukherjee, S., & Bujnicki, J. M. (2019a). Computational modeling of RNA 3D structure based on experimental data. *Biosci. Rep.*, *39*. https://doi.org/10.1042/BSR20180430
67. Mathews, D. H., & Turner, D. H. (2006). Prediction of RNA secondary structure by free energy minimization. *Curr. Opin. Struct. Biol.*, *16*, 270–278. https://doi.org/10.1016/j.sbi.2006.05.010
68. Yan, K., Arfat, Y., Li, D., Zhao, F., Chen, Z., Yin, C., … Qian, A. (2016). Structure prediction: New insights into decrypting long noncoding RNAs. *IJMS*, *17*:132 https://doi.org/10.3390/ijms17010132
69. Low, J. T., & Weeks, K. M. (2010). SHAPE-directed RNA secondary structure prediction. *Methods (San Diego, Calif.)*, *52*, 150–158. https://doi.org/10.1016/j.ymeth.2010.06.007
70. Rice, G. M., Leonard, C. W., & Weeks, K. M. (2014). RNA secondary structure modeling at consistent high accuracy using differential SHAPE. *RNA*, *20*, 846–854. https://doi.org/10.1261/rna.043323.113

71. Dans, P. D., Gallego, D., Balaceanu, A., Darré, L., Gómez, H., & Orozco, M. (2019). Modeling, simulations, and bioinformatics at the service of RNA structure. *Chem*, *5*, 51–73. https://doi.org/10.1016/j.chempr.2018.09.015
72. Patel, T. R., Chojnowski, G., Astha, Koul, A., Mckenna, S. A., & Bujnicki, J. M. (2017). Structural studies of RNA-protein complexes: A hybrid approach involving hydrodynamics, scattering, and computational methods. *Methods (San Diego, Calif.)*, *118-119*:146–162. https://doi.org/10.1016/j.ymeth.2016.12.002
73. Williams, B., Zhao, Bo, Tandon, A., Ding, F., Weeks, K. M., Zhang, Qi, & Dokholyan, N. V. (2017). Structure modeling of RNA using sparse NMR constraints. *Nucleic Acids Res.*, *45*, 12638–12647. https://doi.org/10.1093/nar/gkx1058
74. Ponce-Salvatierra, A., Astha, Merdas, K., Nithin, C., Ghosh, P., Mukherjee, S., & Bujnicki, J. M. (2019b). Computational modeling of RNA 3D structure based on experimental data. *Biosci. Rep.*, *39*. https://doi.org/10.1042/BSR20180430
75. Magnus, M., Matelska, D., Łach, G., Chojnowski, G., Boniecki, M. J., Purta, E., … Bujnicki, J. M. (2014). Computational modeling of RNA 3D structures, with the aid of experimental restraints. *RNA Biology*, *11*, 522–536. https://doi.org/10.4161/rna.28826
76. Schlick, T., & Pyle, A. M. (2017). Opportunities and challenges in RNA structural modeling and design. *Biophys. J.*, *113*, 225–234. https://doi.org/10.1016/j.bpj.2016.12.037
77. Guttman, M., & Rinn, J. L. (2012). Modular regulatory principles of large non-coding RNAs. *Nature*, *482*, 339–346. https://doi.org/10.1038/nature10887
78. Simon, M. D., Pinter, S. F., Fang, R., Sarma, K., Rutenberg-Schoenberg, M., Bowman, S. K., … Lee, J. T. (2013). High-resolution Xist binding maps reveal two-step spreading during X-chromosome inactivation. *Nature*, *504*, 465–469. https://doi.org/10.1038/nature12719
79. Pintacuda, G., Young, A. N., & Cerase, A. (2017). Function by structure: Spotlights on Xist long non-coding RNA. *Front. Mol. Biosci.*, *4*:90. https://doi.org/10.3389/fmolb.2017.00090
80. Strehle, M., & Guttman, M. (2020). Xist drives spatial compartmentalization of DNA and protein to orchestrate initiation and maintenance of X inactivation. *Curr. Opin. Cell Biol.*, *64*, 139–147. https://doi.org/10.1016/j.ceb.2020.04.009
81. Wutz, A., Rasmussen, T. P., & Jaenisch, R. (2002). Chromosomal silencing and localization are mediated by different domains of Xist RNA. *Nat. Genet.*, *30*, 167–174. https://doi.org/10.1038/ng820
82. Colognori, D., Sunwoo, H., Kriz, A. J., Wang, C. -Yu, & Lee, J. T. (2019). Xist deletional analysis reveals an interdependency between Xist RNA and polycomb complexes for spreading along the inactive X. *Mol. Cell*, *74*:101–117.e10.e10. https://doi.org/10.1016/j.molcel.2019.01.015
83. Cerase, A., Armaos, A., Neumayer, C., Avner, P., Guttman, M., & Tartaglia, G. G. (2019). Phase separation drives X-chromosome inactivation: A hypothesis. *Nat. Struct. Mol. Biol.*, *26*, 331–334. https://doi.org/10.1038/s41594-019-0223-0
84. Shin, Y., & Brangwynne, C. P. (2017). Liquid phase condensation in cell physiology and disease. *Science*, *357*, eaaf4382 https://doi.org/10.1126/science.aaf4382
85. Cirillo, D., Blanco, M., Armaos, A., Buness, A., Avner, P., Guttman, M., … Tartaglia, G. G. (2017). Quantitative predictions of protein interactions with long noncoding RNAs. *Nat. Methods*, *14*, 5–6. https://doi.org/10.1038/nmeth.4100
86. Maenner, S., Blaud, M., Fouillen, L., Savoye, A., Marchand, V., Dubois, A., … Branlant, C. (2010). 2-D structure of the A region of Xist RNA and Its implication for PRC2 association. *PLoS Biol.*, *8*:e1000276 https://doi.org/10.1371/journal.pbio.1000276
87. Colognori, D., Sunwoo, H., Wang, D., Wang, C.-Y., & Lee, J. T. (2020). Xist repeats A and B account for two distinct phases of X inactivation establishment. *Dev. Cell*, *54*:21–32.e5.e5. https://doi.org/10.1016/j.devcel.2020.05.021
88. Galupa, R., & Heard, E. (2015). X-chromosome inactivation: New insights into cis and trans regulation. *Curr. Opin. Genet. Dev.*, *31*, 57–66. https://doi.org/10.1016/j.gde.2015.04.002
89. Tsai, M.-C., Manor, O., Wan, Y., Mosammaparast, N., Wang, J. K., Lan, F., … Chang, H. Y. (2010b). Long noncoding RNA as modular scaffold of histone modification complexes. *Science*, *329*, 689–693. https://doi.org/10.1126/science.1192002
90. Yang, Y., Wen, L., & Zhu, H. (2015b). Unveiling the hidden function of long non-coding RNA by identifying its major partner-protein. *Cell Biosci.*, *5*:59 https://doi.org/10.1186/s13578-015-0050-x
91. Zhao, J., Sun, B. K., Erwin, J. A., Song, J.-J., & Lee, J. T. (2008). Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science*, *322*, 750–756. https://doi.org/10.1126/science.1163045
92. Guo, X., Chang, Q., Pei, H., Sun, X., Qian, X., Tian, C., & Lin, H. (2017). Long non-coding RNAâ€"mRNA correlation analysis reveals the potential role of HOTAIR in pathogenesis of sporadic thoracic aortic aneurysm. *Eur. J. Vasc. Endovasc. Surg.*, *54*, 303–314. https://doi.org/10.1016/j.ejvs.2017.06.010
93. Somarowthu, S., Legiewicz, M., Chillón, I., Marcia, M., Liu, F., & Pyle, A. M. (2015b). HOTAIR forms an intricate and modular secondary structure. *Mol. Cell*, *58*, 353–361. https://doi.org/10.1016/j.molcel.2015.03.006
94. Spokoini-Stern, R., Stamov, D., Jessel, H., Aharoni, L., Haschke, H., Giron, J., … Bachelet, I. (2020). Visualizing the structure and motion of the long noncoding *RNA* HOTAIR. *RNA*, *26*, 629–636. https://doi.org/10.1261/rna.074633.120
95. Zhang, X., Rice, K., Wang, Y., Chen, W., Zhong, Y., Nakayama, Y., … Klibanski, A. (2010). Maternally expressed gene 3 (MEG3) noncoding ribonucleic acid: Isoform structure, expression, and functions. *Endocrinology*, *151*, 939–947. https://doi.org/10.1210/en.2009-0657
96. Mondal, T., Subhash, S., Vaid, R., Enroth, S., Uday, S., Reinius, B., … Kanduri, C. (2015). MEG3 long noncoding RNA regulates the TGF-β pathway genes through formation of RNA–DNA triplex structures. *Nat. Commun.*, *6*, 7743 https://doi.org/10.1038/ncomms8743
97. Uroda, T., Anastasakou, E., Rossi, A., Teulon, J.-M., Pellequer, J.-L., Annibale, P., … Marcia, M. (2019). Conserved pseudoknots in lncRNA MEG3 Are essential for stimulation of the p53 pathway. *Mol. Cell*, *75*:982–995.e9.e9. https://doi.org/10.1016/j.molcel.2019.07.025
98. Xue, Z., Hennelly, S., Doyle, B., Gulati, A. A., Novikova, I. V., Sanbonmatsu, K. Y., & Boyer, L. A. (2016). A G-rich motif in the lncRNA braveheart interacts with a zinc-finger transcription factor to specify the cardiovascular lineage. *Mol. Cell*, *64*, 37–50. https://doi.org/10.1016/j.molcel.2016.08.010
99. Kim, D. N., Thiel, B. C., Mrozowich, T., Hennelly, S. P., Hofacker, I. L., Patel, T. R., & Sanbonmatsu, K. Y. (2020). Zinc-finger protein CNBP alters the 3-D structure of lncRNA Braveheart in solution. *Nat. Commun.*, *11*, 148 https://doi.org/10.1038/s41467-019-13942-4
100. Breaker, R. R. (2012). Riboswitches and the RNA world. *Cold Spring Harbor Perspect. Biol.*, *4*, a003566–a003566. https://doi.org/10.1101/cshperspect.a003566
101. Henkin, T. M. (2008). Riboswitch RNAs: Using RNA to sense cellular metabolism. *Genes Dev.*, *22*, 3383–3390. https://doi.org/10.1101/gad.1747308
102. Mchugh, C. A., Chen, C.-K., Chow, A., Surka, C. F., Tran, C., Mcdonel, P., … Guttman, M. (2015). The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature*, *521*, 232–236. https://doi.org/10.1038/nature14443
103. Sherpa, C., Rausch, J. W., & Le Grice, S. F. (2018). Structural characterization of maternally expressed gene 3 RNA reveals conserved motifs and potential sites of interaction with polycomb repressive complex 2. *Nucleic Acids Res.*, *46*, 10432–10447. https://doi.org/10.1093/nar/gky722

104. Fang, R., Moss, W. N., Rutenberg-Schoenberg, M., & Simon, M. D. (2015). Probing Xist RNA structure in cells using targeted structure-Seq. *PLoS Genet.*, *11*, e1005668 https://doi.org/10.1371/journal.pgen.1005668
105. Liu, F., Somarowthu, S., & Pyle, A. M. (2017). Visualizing the secondary and tertiary architectural domains of lncRNA RepA. *Nat. Chem. Biol.*, *13*, 282–289. https://doi.org/10.1038/nchembio.2272
106. Wu, L., Murat, P., Matak-Vinkovic, D., Murrell, A., & Balasubramanian, S. (2013). Binding interactions between long noncoding RNA HOTAIR and PRC2 proteins. *Biochemistry*, *52*, 9519–9527. https://doi.org/10.1021/bi401085h
107. Jain, A., & Vale, R. D. (2017). RNA phase transitions in repeat expansion disorders. *Nature*, *546*, 243–247. https://doi.org/10.1038/nature22386
108. Chu, Ci, Zhang, Q. C., Da Rocha, S. T., Flynn, R. A., Bharadwaj, M., Calabrese, J. M., … Chang, H. Y. (2015). Systematic discovery of Xist RNA binding proteins. *Systematic Discovery of Xist RNA Binding Proteins. Cell*, *161*, 404–416. https://doi.org/10.1016/j.cell.2015.03.025
109. Tsai, M.-C., Manor, O., Wan, Y., Mosammaparast, N., Wang, J. K., Lan, F., … Chang, H. Y. (2010a). Long noncoding RNA as modular scaffold of histone modification complexes. *Science*, *329*, 689–693. https://doi.org/10.1126/science.1192002

How to cite this article: Graf, J., & Kretz, M. (2020). From structure to function: Route to understanding lncRNA mechanism. *BioEssays*, 42, e2000027. https://doi.org/10.1002/bies.202000027