

Review

Not peer-reviewed version

The Unpaved Road of Non-Coding RNA Structure-Function Relationships: Current Knowledge, Available Methodologies and Future Trends

[Ana Lúcia Leitão](#) and [Francisco J. Enguita](#) *

Posted Date: 14 January 2025

doi: 10.20944/preprints202501.1068.v1

Keywords: non-coding RNA; X-ray crystallography; cryo-EM; chemical probing; structure-function relationships



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Review

The Unpaved Road of Non-Coding RNA Structure-Function Relationships: Current Knowledge, Available Methodologies and Future Trends

Ana Lúcia Leitão ¹ and Francisco J. Enguita ^{2,*}

¹ Departamento de Química, Faculdade de Ciências e Tecnologia, Universidade NOVA de Lisboa, Campus da Caparica, 2829-516 Caparica, Portugal

² Faculdade de Medicina, Universidade de Lisboa, Av. Prof. Egas Moniz, 1649-028 Lisboa, Portugal

* Correspondence: fenguita@medicina.ulisboa.pt

Abstract: The genomes from complex eukaryotes are enriched in non-coding genes whose transcription products (non-coding RNAs) are involved in the regulation of genomic output at different levels. Non-coding RNA action is predominantly driven by sequence and structural motifs that interact with specific functional partners. Despite the exponential growth in primary RNA sequence data facilitated by next-generation sequencing studies, the availability of tridimensional RNA data is comparatively more limited. The subjacent reasons for this relative lack of information regarding RNA structure are related to the specific chemical nature of RNA molecules and the limitations of the current available methods for structural characterization of biomolecules. In this review, we describe and analyze the different structural motifs involved in non-coding RNA function, and the wet-lab and computational methods used to characterize their structure-function relationships, highlighting the current need for detailed structural studies to explore the molecular determinants of non-coding RNA function.

Keywords: non-coding RNA; X-ray crystallography; cryo-EM; chemical probing; structure-function relationships

1. Prevalence of Non-Coding Genes in the Genomes of Complex Eukaryotes

The genomes of higher eukaryotic organisms contain information units that will be translated into polypeptides (coding DNA) and territories that will not encode for proteins (non-coding DNA). Non-coding DNA harbors a wide array of elements, such as regulatory regions (promoters, enhancers, silencers and insulators), repetitive elements (transposons and satellite DNA), structural regions (telomeres and centromeres), along with functional non-coding RNAs (nucleolar RNAs, ribosomal RNAs, transfer RNAs, microRNAs, PIWI-interacting RNAs and long non-coding RNAs) [1]. The extensive prevalence of non-coding sequences in the human and complex eukaryotic genomes has profound evolutionary implications, reshaping our understanding of genetic complexity, regulatory innovation, and adaptive potential. One of the most striking evolutionary implications of abundant non-coding DNA is its contribution to regulatory complexity [2]. The human genome's vast non-coding regulatory regions include enhancers, silencers, insulators, and ncRNA genes that orchestrate gene expression patterns, which are also supported by the pervasive transcription phenomenon [3]. These elements allow for the fine-tuned spatiotemporal regulation of protein-coding genes, which is also essential for the development and functionality of multicellular organisms [2]. The evolutionary plasticity of non-coding regions has likely facilitated the rapid emergence of adaptive traits. Unlike protein-coding genes, where mutations are often deleterious, mutations in non-coding regions are less constrained and more likely to be neutral or adaptive. This

flexibility enables the exploration of novel regulatory networks and functions without disrupting essential protein-coding sequences [4]. The ncRNA-driven regulatory networks are hypothesized to drive much of the phenotypic complexity observed in higher eukaryotes. For instance, humans and other mammals have a similar number of protein-coding genes to simpler organisms, such as nematodes, yet they exhibit vastly greater biological complexity. This disparity could be partially explained by the regulatory information present within non-coding regions [5]. The human genome contains tens of thousands of non-coding RNA (ncRNA) genes, being long non-coding RNAs (lncRNAs) the most represented family of genome information units. The last version of the human GENCODE database includes approximately 7,500 small ncRNA and over 35,000 lncRNA genes [6].

NcRNA genes are transcribed to produce RNA molecules that will have different regulatory effects over the genomic output, controlling the flow of information from coding genes, and acting either at the nucleus or the cytoplasm [7]. As with any RNA molecule, ncRNAs simultaneously harbor sequence and structural information, however their regulatory functions are largely dictated by their secondary and tertiary structural features rather than primary nucleotide sequences. The structural features of ncRNAs are responsible for their participation in molecular recognition phenomena, catalysis of enzymatic reactions, and the regulation of gene expression [8]. Characterizing the structure of ncRNAs by different techniques, either by wet-lab protocols or computer-based methods, is essential to dissect their functions and mechanisms of action.

2. General Principles and Rules Governing RNA Structure

The functional versatility of RNA relies on its ability to adopt diverse three-dimensional (3D) structures, mainly driven by the single-chain nature of RNA, and the presence of ribose sugars in the building nucleotide units. Understanding the underlying principles and features of RNA 3D architecture is specifically relevant for elucidating the biological roles of ncRNAs. As proteins, RNA presents different organization levels, encompassing primary, secondary, and tertiary structures (Figure 1). The primary structure of RNA refers to its linear sequence of ribonucleotides and encodes the genetic and functional information of the RNA molecule. The RNA sequence determines the potential for base pairing and stacking interactions, which could drive the generation of higher-order structures [9]. Secondary structure arises from hydrogen bonding between complementary bases within the same RNA molecule, giving rise to several RNA motifs such as stems and loops [10,11]. In RNA molecules, due to the presence of a 5'-hydroxyl group in the ribose sugar, the base pairing phenomena can be achieved also by non-canonical rules (non Watson-Crick pairing), expanding the repertoire of possible base interactions when compared to DNA [12]. Additionally, tertiary structure refers to the three-dimensional conformation of an RNA molecule, arising from long-range interactions between secondary structural elements. Key tertiary motifs include **pseudoknots**, formed by base pairing between a loop and a complementary sequence outside the loop; **triple helices**, or **structures formed by** interactions between three strands of RNA; and kissing hairpins, structures that stabilize two hairpins by interactions between the unpaired looped bases [13,14].

The 3D structure of RNA is a result of intricate interactions among its nucleotide components. RNA molecules will fold into conformations that minimize their Gibbs free energy, achieving a stable structure under given environmental conditions [9]. The folding process reduces the conformational entropy of the molecule, being driven by favorable enthalpic interactions, including hydrogen bonding and base stacking. RNA folding typically occurs hierarchically, where secondary structures (e.g., helices and loops) form first, followed by tertiary interactions. The stability of RNA structures is maintained by a combination of noncovalent interactions that include base pairing, base stacking, hydrogen bonding and electrostatic interactions [15]. Canonical base pairing provides significant thermodynamic stability to RNA structures. Hydrogen bonds between the bases in A-U (two hydrogen bonds) and G-C (three hydrogen bonds) pairs contribute to the overall stability of RNA duplexes and other base-paired regions. Canonical base pairing also plays a role in guiding the folding pathways of RNA. The complementarity of A-U and G-C pairs facilitates the correct alignment of RNA strands, minimizing misfolding and the formation of kinetically trapped

structures [16]. This specificity is particularly critical in complex RNA molecules, such as ribosomal RNA, ribozymes and lncRNAs, where precise folding is required for specific functions [17]. Base stacking arises from π - π interactions between the aromatic rings of adjacent nucleobases, complemented by hydrophobic forces that exclude water molecules from the stacked regions. These interactions are stronger in RNA than in DNA due to RNA's 2'-hydroxyl group, which introduces additional electrostatic effects and influences sugar-phosphate backbone geometry [18]. Base stacking provides substantial stabilization to RNA structures by lowering their free energy. Unlike hydrogen bonds in base pairing, stacking interactions are largely sequence-independent but strongly related to the sequence context. Pyrimidine-purine sequences (e.g., C-G or U-G) tend to exhibit stronger stacking due to optimal overlap of aromatic π -electron clouds [19,20]. These interactions contribute to the enhanced stability of RNA helices, particularly in regions with high G-C content, as guanine and cytosine bases have favorable stacking geometries

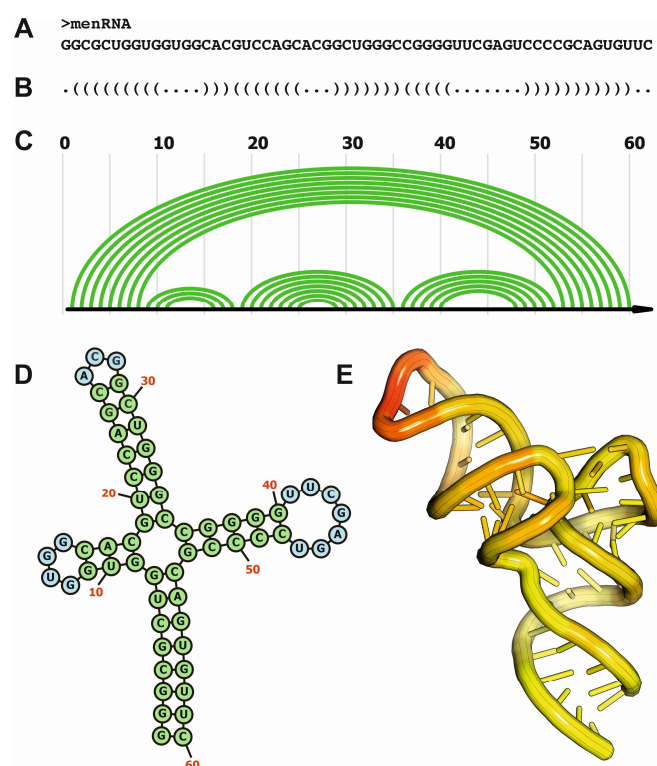


Figure 1. Different organization levels of RNA structure and their representations, exemplified by the human menRNA. A, primary structure of menRNA; B, secondary structure representation of menRNA using the "dot-bracket" notation; C, secondary structure of menRNA using the arc representation; D, secondary structure of menRNA using the 2D diagram format; E, tridimensional structure of menRNA using coordinate data extracted from the experimentally determined structure (PDB code: 8VT5).

In addition to stabilizing duplexes, base stacking also supports the stability of non-canonical structures, such as pseudoknots, internal loops, and junctions. These interactions provide a scaffold for tertiary interactions, such as metal ion coordination and long-range base pairings, that contribute to the overall folding of complex RNAs [21]. Base stacking also influences the conformational dynamics of RNA. Partially stacked regions enable the structural flexibility required for RNA to undergo conformational changes, which is critical for its interactions with proteins, small ligands, and other nucleic acids [22]. Moreover, hydrogen bonding extends beyond base pairing to mediate tertiary interactions that stabilize the three-dimensional architecture of RNA molecules. Hydrogen bonds between unpaired bases in RNA loops facilitate long-range interactions [13], and contribute to the stabilization of intricate tertiary folds when they are formed between bases and the sugar-phosphate backbone [23]. Additionally, the RNA backbone carries a negative charge at each

phosphate group, which contributes to significant electrostatic repulsion within the molecule. This repulsion presents a thermodynamic challenge for folding and stabilization of compact RNA structures [24,25]. The presence of cations in the cellular environment mitigates the repulsion between negatively charged groups. Counterions, such as sodium (Na^+), potassium (K^+), magnesium (Mg^{2+}), and other cationic compounds such as polyamines, could neutralize the phosphate backbone charges, also contributing to stable RNA conformations [26,27].

3. Functional Relevance of Structural Elements in Non-Coding RNAs

RNA structural motifs are arrangements of nucleotide sequences and secondary structures that contribute to the overall architecture and function of RNA molecules. The structural arrangement of RNA motifs arises from specific nucleotide interactions. These motifs are characterized by recurrent patterns of base pairing, stacking, and backbone conformations. They can form within single-stranded regions or emerge from interactions between different secondary structure elements [28]. Structural motifs often stabilize RNA tertiary structures and facilitate interactions with proteins, small molecules, or other RNA molecules, being very relevant in the biogenesis and function of ncRNAs (Figure 2).

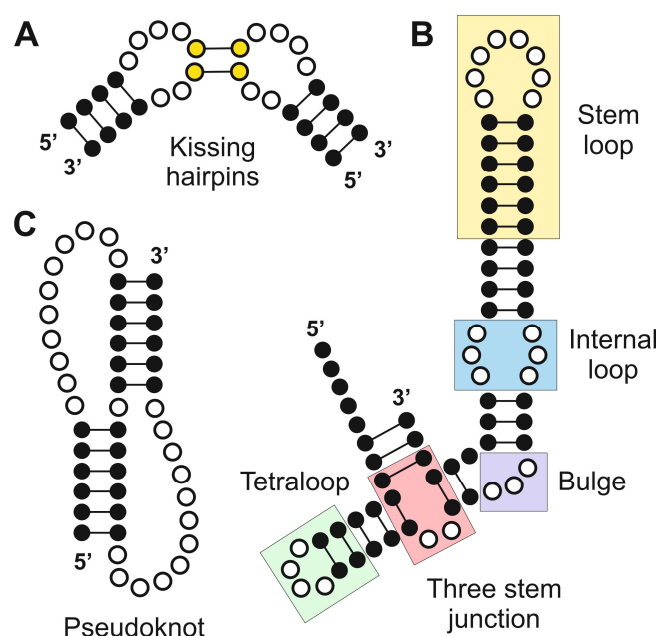


Figure 2. Selected RNA elements generated by secondary and tertiary structure arrangements with relevance in the biogenesis and function of ncRNAs. A, kissing hairpins; B, combined RNA segment that contains a stem loop, and internal loop, a bulge, a three-stem junction and a tetraloop; and C, pseudoknot.

3.1. Hairpin Loops

Hairpin loops are the most common secondary structural motifs in RNA. They consist of a loop of unpaired nucleotides flanked by a double-stranded stem (Figure 2). The loop typically contains 4–10 nucleotides and stabilizes the RNA structure by closing the end of a stem. Often found at the terminus of a stem, hairpin loops are single-stranded RNA segments that loop back on themselves to form a stable structure [29]. The GNRA tetraloop, where G, N, R, and A represent guanine, any nucleotide, purine, and adenine respectively, is one of the most common. This motif is related to the stabilization of RNA tertiary structures and can participate in long-range tertiary interactions, facilitating RNA folding and functionality [30].

Hairpin loops are central to the maturation process of microRNAs (miRNAs), a conserved class of small ncRNAs involved in post-transcriptional gene silencing. MiRNA precursors, known as primary miRNAs (pri-miRNAs), are transcribed by RNA polymerase II and fold into characteristic

hairpin loop structures [31]. These structures are recognized and processed by the nuclear Microprocessor complex, comprising the RNase III enzyme Drosha and its cofactor DGCR8. The hairpin loop ensures precise cleavage by the Microprocessor, enabling the release of precursor miRNAs (pre-miRNAs) with defined stem-loop structures [32]. The Microprocessor complex has structural flexibility to accommodate different pri-miRNAs [33]. Hairpin loops also regulate the efficiency and specificity of miRNA-target interactions. For instance, mutations in the hairpin structure of let-7 miRNA result in aberrant processing and diminished silencing activity [34]. These findings emphasize how the physical characteristics of hairpin loops contribute to the fidelity of miRNA maturation. Furthermore, auxiliary proteins such as the double-stranded RNA-binding protein DGCR8 and TAR RNA-binding protein (TRBP) recognize and stabilize hairpin loops, enhancing the precision of Drosha and Dicer cleavage events [35]. These protein-RNA interactions highlight the broader role of hairpin loops in facilitating multi-protein complexes essential for miRNA function.

Small interfering RNAs (siRNAs) also rely on hairpin structures for their generation and function. Endogenously derived siRNAs often originate from long double-stranded RNA (dsRNA) precursors, which can include hairpin-forming transcripts. Dicer processes these dsRNA precursors into siRNA duplexes [36]. The precise structure of the hairpin loop can influence Dicer activity, determining the efficiency and accuracy of siRNA generation [37].

Long non-coding RNAs (lncRNAs), which regulate diverse cellular processes such as chromatin remodeling and transcriptional control, also feature hairpin loops as functional motifs. Hairpin structures in lncRNAs can mediate protein-RNA interactions, RNA-RNA duplex formation, and scaffolding of ribonucleoprotein complexes [38]. For instance, the lncRNA HOTAIR utilizes hairpin loops to interact with Polycomb Repressive Complex 2 (PRC2), guiding it to specific genomic loci for histone modification and transcriptional silencing [39]. Similarly, MALAT1, a nuclear retained lncRNA, contains hairpin structures that regulate its stability and interactions with splicing factors, highlighting the versatility of these motifs [40].

3.2. Hairpin Loops

Bulges are regions of unpaired nucleotides on one strand of an RNA duplex, while internal loops consist of unpaired nucleotides on both strands (Figure 2) [41]. Bulges and internal loops are ubiquitous in ncRNAs, appearing in ribosomal RNAs (rRNAs), small nuclear RNAs (snRNAs), small nucleolar RNAs (snoRNAs), miRNA precursors, and lncRNAs. Both structural features can disrupt regular Watson-Crick base pairing, introducing flexibility and asymmetry into RNA structures [42]. This flexibility is critical for enabling ncRNAs to adopt specific conformations required for their functions. **Bulges** often act as hinge points, allowing the RNA to bend or twist into compact tertiary structures. This flexibility can be essential for the formation of higher-order assemblies, such as pseudoknots or long-range base-pairing interactions [8]. **Internal loops**, especially those with specific sequence motifs, can form non-canonical base pairs, such as G–U wobble pairs, or participate in hydrogen bonding networks that stabilize the RNA's tertiary structure [14].

Bulges and internal loops are critical for RNA-protein interactions, which are fundamental to the functions of many ncRNAs. These structural elements often serve as recognition motifs for RNA-binding proteins (RBPs), which preferentially bind to unpaired or irregular regions of RNA [43]. For instance, bulges and internal loops within rRNAs are essential for their interactions with ribosomal proteins and other factors involved in translation [44]. The internal loops in the peptidyl transferase center of the ribosome mediate contacts with transfer RNAs (tRNAs) and catalytic proteins [45]. Additionally, internal loops in miRNA precursors (pre-miRNAs) are recognized by processing enzymes such as Dicer. A study by MacRae and Doudna described how Dicer endonuclease can establish specific interactions with these internal loops, highlighting their role in enzyme orientation for accurate substrate cleavage [46].

Bulges and internal loops in lncRNAs contribute to their ability to act as scaffolds for protein complexes. For example, the lncRNA Xist, which mediates X-chromosome inactivation, contains

internal loops that facilitate interactions with the polycomb repressive complex 2 (PRC2). A detailed study by Zhao and coworkers highlights these interactions, demonstrating how specific structural motifs in Xist recruit PRC2 to silence the X chromosome [47]. Another interesting example involves the lncRNA MALAT1, which contains an internal loop that enhances its ability to interact with the splicing factor SRSF1, facilitating alternative splicing regulation [48].

3.3. Pseudoknots

Pseudoknots are a common and functionally significant structural motif found in RNA molecules. These structures occur when bases in a loop of single-stranded RNA form complementary base pairs with a sequence outside the loop, creating a complex tertiary architecture [49]. This interaction results in a compact, often stable three-dimensional structure. Despite their relative simplicity compared to larger RNA folds, pseudoknots play critical roles in the function of non-coding RNAs (ncRNAs), influencing processes such as translation, catalysis, and genome stability, especially in complex organisms. Pseudoknots can be classified based on their topology and the number of stems and loops involved. The simplest form, known as the H-type pseudoknot, consists of two helical stems connected by loops [50]. More complex pseudoknots may involve additional stems, loops, or interactions with proteins and metal ions. The stability of pseudoknots is influenced by factors such as the length of the loops, the sequence composition, and the presence of stabilizing cations such as magnesium ions. The dynamic nature of RNA allows pseudoknots to fold and unfold, enabling them to serve as molecular switches in various biological processes [51].

In many viral and cellular coding RNAs, pseudoknots promote programmed ribosomal frameshifting, a mechanism that allows the ribosome to change reading frames during translation [52]. Pseudoknots positioned downstream of a slippery sequence in the mRNA create mechanical tension as the ribosome unwinds them, stalling the ribosome and inducing frameshifting [49]. This mechanism is critical for the regulation of gene expression in some RNA viruses and for the synthesis of viral polyproteins [53].

Pseudoknots can also act as scaffolds for assembling multi-molecular complexes. In snoRNAs, pseudoknots help organize and stabilize interactions with proteins and other RNAs involved in ribosomal RNA processing and modification. This scaffolding role is essential in complex organisms where ncRNAs participate in highly coordinated molecular machineries [49]. In lncRNAs, pseudoknots contribute to their ability to act as molecular decoys, scaffolds, or guides for proteins and other RNAs. A relevant example is the Telomerase complex, that contains a 451 nt ncRNA (TERC RNA) together with the enzymatic core of the reverse transcriptase (hTERT). TERC ncRNA contains a series of conserved domains, including a pseudoknot. The TERC pseudoknot has been described as an essential feature for the stabilization of the Telomerase complex and its processive activity [54]. The stability and dynamics of pseudoknots in specific lncRNAs are also involved in cellular responses to some signals, highlighting their importance in the regulation of gene expression networks in complex organisms. For example, in the MEG3 lncRNA, a human maternally expressed tumor suppressor that regulates the p53 pathway, an evolutionary conserved region contains two distal motifs that interact by base complementarity to form alternative, mutually exclusive pseudoknot structures. Mutations that disrupt these interactions impair MEG3-dependent p53 stimulation in vivo and disrupt MEG3 folding in vitro [55].

3.4. Kissing Hairpins

Kissing hairpins refer to a specific tertiary interaction between two RNA hairpins, wherein the single-stranded loop regions of the hairpins interact by base pairing to form a stable duplex (Figure 2). This motif is stabilized by complementary sequences within the loops, leading to the formation of a “kissing” interaction. These structures are stabilized by canonical Watson-Crick base pairing and sometimes further reinforced by stacking interactions and cation coordination [28].

The formation of kissing hairpins contributes to the structural stability of lncRNAs in several ways. First, these interactions reduce the overall entropy of the molecule by locking dynamic regions

into defined conformations [56]. This stabilization is particularly important for lncRNAs exposed to the intracellular milieu, where RNases and other destabilizing factors are prevalent. For example, studies on the lncRNA MALAT1 (metastasis-associated lung adenocarcinoma transcript 1) suggested that such tertiary interactions protect the RNA against exonucleolytic degradation [57]. Additionally, kissing hairpins may influence lncRNA folding pathways, guiding the molecule into functionally active conformations. This folding precision ensures that lncRNAs can interact with their protein, DNA, or RNA targets effectively. For instance, a kissing hairpin motif in the HOTAIR (HOX transcript antisense intergenic RNA) lncRNA facilitates its scaffold function, enabling interactions with chromatin-modifying complexes [58].

Beyond structural stability, kissing hairpins play active roles in the functional repertoire of lncRNAs. These motifs serve as interaction platforms for RNA-binding proteins (RBPs) and other nucleic acids. Their defined geometry and sequence specificity make them ideal docking sites for RBPs, as observed in the interaction of the lncRNA Xist with its associated proteins [59]. Kissing hairpins also facilitate lncRNA-mediated RNA-RNA interactions. In the case of lncRNAs involved in RNA splicing, such as NEAT1, kissing hairpin motifs contribute to the assembly of paraspeckles by mediating interactions between NEAT1 and other RNAs [60]. The conservation of kissing hairpin motifs among different lncRNAs suggested an evolutionary trend, that could have functional implications. Comparative analyses have also revealed conserved loop-loop interactions in lncRNAs across species, underscoring their functional importance [61].

4. Methods and Protocols to Study ncRNA Structures

The availability of ncRNA structures in public domain databases such as the PDB [62], is very limited when compared to proteins [63]. The reasons for this lack of experimental information are complex to analyze, but they should rely on the nature of RNA molecules that limit their analysis by wet-lab protocols. RNA molecules are highly flexible and adopt diverse secondary and tertiary structures. Moreover, RNA often exists in multiple stable conformations, complicating efforts to determine a single, well-defined structure. An additional factor is related to the historical emphasis placed on studying protein structures, as they are directly linked to specific biological events such as enzymatic catalysis [64]. This has also led to a disproportionate number of resources allocated to protein structural studies compared to RNA. Interestingly, an exponential growth of available structures from protein-nucleic acid complexes has been experienced in the last decade due to technical improvements in structure determination by X-ray crystallography and cryo-electron microscopy (Figure 3). Among these structures, there are several entries that include fragments or complete ncRNAs in complex to proteins.

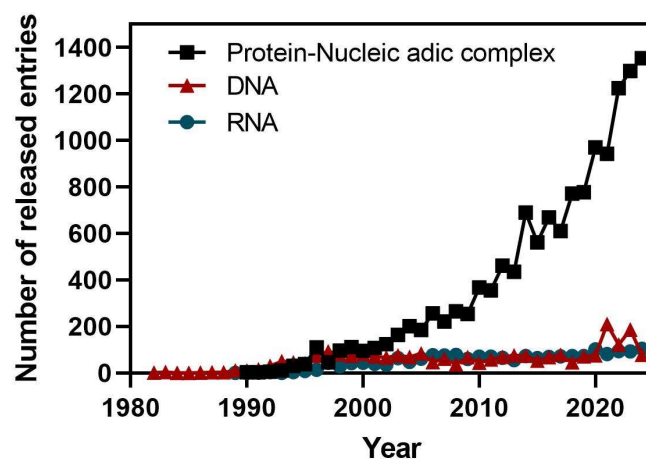


Figure 3. Time course of the deposition statistics in the PDB database showing the number of released entries by year for protein-nucleic acid complexes, and for DNA and RNA molecules (Source: PDB statistics).

The structure of ncRNAs and other functionally relevant RNAs can be studied by different approaches if a characterization of the secondary or tertiary structures is required. Compared to proteins, RNA molecules can be structurally characterized by a more diverse group of methods due to their chemical nature (Figure 4). In fact, classical methods such as nuclear magnetic resonance, X-ray crystallography and cryo-electron microscopy can be used for in vitro characterization of RNA tertiary structure. Additionally, the secondary structure of RNA molecules can be characterized in vivo and, in a genome-wide manner, by combining chemical probing and next-generation sequencing. Moreover, computer-based methods can be also applied for the prediction and modelling of RNA secondary and tertiary structures.

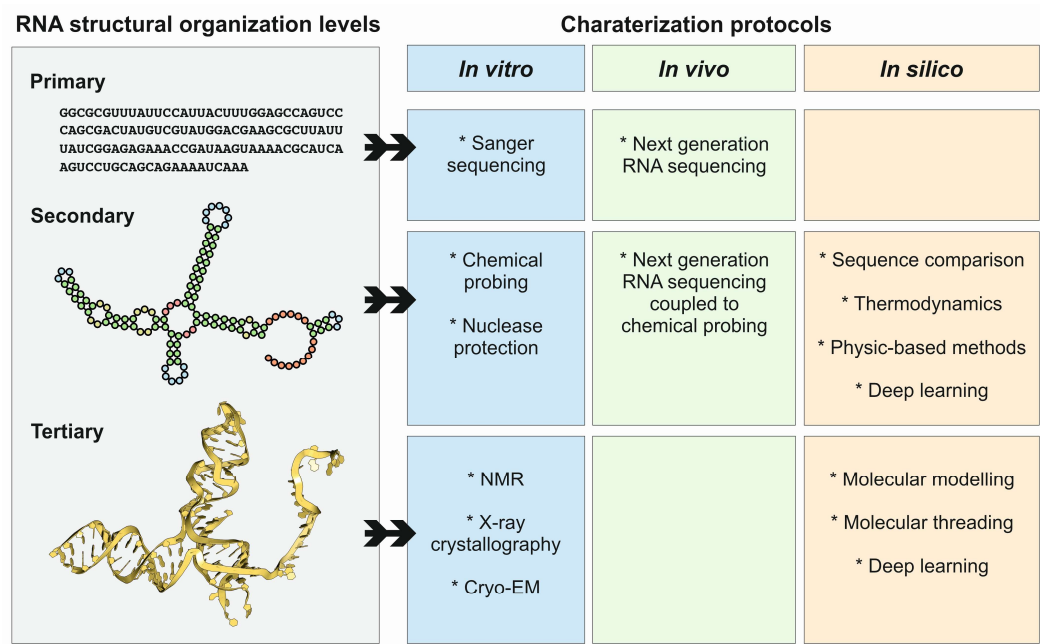


Figure 4. Structural organization levels of RNA molecules, and protocols used for their characterization, in vitro, in vivo and in silico. Note that the different protocols and methods are highly connected, and in many cases are used together as synergistic approaches.

4.1. X-ray Crystallography

X-ray crystallography is a classical technique in structural biology, enabling the elucidation of three-dimensional structures at atomic resolution through X-ray diffraction [65]. The obtention of meaningful X-ray diffraction patterns requires the use of molecular crystals, ordered 3D lattices composed by the spatial repetition of molecular units following geometrical rules [66]. Despite its importance, RNA crystallization faced important challenges due to the intrinsic properties of RNA molecules, such as their structural flexibility, heterogeneity, and biochemical instability [67]. RNA molecules are highly dynamic and flexible, often adopting multiple conformations due to their inherent base-pairing and stacking interactions. This conformational heterogeneity often prevents the formation of ordered crystalline lattices. The presence of non-canonical interactions and tertiary motifs can further exacerbate this challenge by creating structurally diverse populations of RNA. RNA is also prone to degradation by ribonucleases (RNases) and spontaneous hydrolysis under physiological conditions [68]. The susceptibility of RNA to degradation complicates its preparation and handling during crystallization experiments [65]. Another important factor that could prevent RNA crystallization is the relative molecular size. Whereas short RNA molecules (<20 nucleotides) may not contain sufficient structural complexity to support crystal lattice formation, long RNA sequences which are very frequent in ncRNAs (>100 nucleotides) often exhibit excessive flexibility and heterogeneity. Additionally, the sequence composition, such as high uridine or cytidine content, can influence RNA’s propensity to crystallize [69]. RNA tends to form poorly ordered crystals due to

suboptimal packing interactions between molecules. The dominance of electrostatic repulsion between negatively charged phosphate backbones may prevent the close packing required for high-resolution X-ray diffraction [66]. Even when high-quality RNA crystals are obtained, determining phases for X-ray diffraction can be challenging due to the lack of established protocols for introducing heavy-atom derivatives into RNA crystals, which are necessary for phase retrieval using isomorphous replacement (MIR) or multi-wavelength anomalous dispersion (MAD) methods [70].

Some strategies have been developed to increase the probability of successful crystallization of RNA molecules, which have been applied to ncRNAs. The main strategy used was to minimize the RNA conformational heterogeneity by rational design of RNA constructs that can enhance crystallization [68]. These include the truncation of non-essential regions, introducing stabilizing mutations (e.g., replacing uridines with pseudouridines), or incorporating locked nucleic acids (LNAs) that can reduce flexibility and conformational heterogeneity. Moreover, some tertiary structure scaffolds can be engineered, such as kissing loops or tetraloops, to promote the formation of uniform RNA structures that could facilitate crystallization. Incorporation of 2'-O-methyl or phosphorothioate groups can also enhance RNA stability by reducing susceptibility to hydrolysis and RNase degradation [71]. Adding divalent cations (e.g., Mg^{2+} , Mn^{2+}) or polyamines can mitigate electrostatic repulsion between phosphate groups, facilitating tighter packing in crystals. In any case, it is evident that RNA often crystallizes more readily as part of an RNA-protein complex than as a standalone molecule [69]. Proteins can serve as scaffolds that stabilize RNA tertiary structures and promote crystal lattice formation. Co-crystallization with RNA-binding proteins or using engineered proteins with high affinity for RNA motifs, is a commonly employed strategy for structural studies of ncRNAs.

X-ray crystallography was used for solving the structure of small and compact ncRNAs such as catalytic prokaryotic ncRNAs. Ribozymes are a classic example of ncRNAs whose structures have been resolved using X-ray crystallography. The hammerhead ribozyme, which catalyzes site-specific RNA cleavage, was among the first ribozymes to have its structure determined [72]. Early studies used X-ray crystallography to reveal the catalytic core's architecture, including the specific interactions between nucleotide residues critical for its function [73]. These findings provided a molecular basis for understanding RNA catalysis and paved the way for engineering RNA-based catalysts. Another example are riboswitches, regulatory elements in messenger RNAs that control gene expression by binding to small metabolites. X-ray crystallography has been instrumental in uncovering the binding pockets and conformational changes of various riboswitches [74]. For instance, the structure of the thiamine pyrophosphate (TPP) riboswitch was determined at high resolution, revealing how the RNA specifically recognizes the TPP molecule through hydrogen bonding and van der Waals interactions [75]. These studies elucidated the structural basis for ligand-induced RNA folding and regulation.

Small and medium-sized ncRNAs from eukaryotic organisms have been also crystallized, and their structures solved by X-ray diffraction. Pioneer examples include the resolution of the structure of tRNAs from yeast and humans [76]. More recently, X-ray diffraction was used to characterize the structure of specific processing products generated from lncRNAs. The primary transcripts of some lncRNAs such as MALAT1 and MEN β contain a tRNA-like structure at their 3' end, which are processed and cleaved by tRNA processing enzymes RNase P and RNase Z. This maturation pathway generates a cytoplasmic tRNA-like small RNA, known as mascRNA for MALAT1 and menRNA for MEN β , along with mature nuclear retained MALAT1 and MEN β lncRNAs [77]. The molecular structures of menRNA and mascRNA have been solved by X-ray crystallography [78].

4.2. Cryo-Electron Microscopy

Cryo-electron microscopy (cryoEM) has emerged as a transformative technique in structural biology, offering unique advantages for studying RNA molecules, including ncRNAs [79]. Cryo-EM's ability to resolve structures in a near-native environment, without requiring crystallization, has significantly advanced our understanding of ncRNAs, even those with challenging biochemical properties. Recent technological developments in cryoEM, such as direct electron detectors, advanced computational algorithms, and phase plate technology, have greatly improved resolution and contrast, enabling near-atomic structural determination of large and complex ncRNAs as well as ncRNA-protein assemblies [80]. Cryo-EM circumvents the size and flexibility constraints often encountered with other techniques such as X-ray crystallography. Cryo-EM has been successfully applied to elucidate the structures of several ncRNAs and their complexes, offering detailed insights into their mechanisms of action [81] (Figure 5).

One of the landmark applications of cryo-EM to elucidate the function of ncRNAs, was the resolution of the spliceosome structure. The spliceosome is a dynamic ribonucleoprotein complex responsible for the precise removal of introns from pre-mRNA transcripts. High-resolution cryo-EM studies on the yeast spliceosome by Yan and colleagues in 2015 and 2016 have provided detailed insights into its structural organization during various stages of the splicing cycle [82,83]. Cryo-EM analysis solved the structure of the catalytic center of the spliceosome, formed by intricate RNA-RNA and RNA-protein interactions. Key elements include the U2/U6 snRNA duplex, which forms a three-helix junction critical for catalysis, and the U5 snRNA, which interacts with the 5' and 3' splice sites to align the exons for ligation [82]. Additional details of the dynamics of the yeast spliceosome were also determined by cryo-EM experiments, showing that the spliceosome undergoes significant conformational rearrangements to facilitate the two-step splicing reaction where ncRNAs have an important role. In the transition from the Bact (activated) to the C* (catalytically activated) complex, the first and second transesterification reactions are possible due to structural rearrangements that orient the pre-mRNA substrate and the catalytic components of the spliceosome. These rearrangements are mediated by ATP-dependent RNA helicases, which remodel RNA-RNA and RNA-protein interactions to drive the splicing cycle forward [84]. Bertram and coworkers reported a 3D cryo-EM structure of the human pre-catalytic spliceosome B complex. This structure revealed that the U2 snRNP-containing head domain is connected to the main body of the B complex via three primary bridges [85,86]. In a separate study, Zhang and coworkers presented the cryo-EM structure of the human spliceosome just before exon ligation, known as the C* complex. This structure provided insights into the arrangement of splicing factors surrounding the catalytic core of the complex [87]. These studies highlight both conserved and unique features of the human spliceosome compared to its yeast counterpart. For instance, while the overall architecture of the catalytic U2-U6 ribonucleoprotein core is similar between species, the human spliceosome exhibits specific structural adaptations, such as the involvement of metazoan-specific proteins that stabilize key protein and RNA regions.

Telomerase is a ribonucleoprotein complex that extends telomeres, the protective ends of linear chromosomes, by adding repetitive nucleotide sequences. Cryo-EM studies have elucidated the structural characteristics of telomerase in both humans and the ciliate *Tetrahymena thermophila*, providing insights into their functional mechanisms. In 2018, Jiang and colleagues reported the cryo-EM structure of the *Tetrahymena* telomerase holoenzyme at a resolution of approximately 4.8 Å [88]. This study revealed that the catalytic core of the enzyme consists of TERT and telomerase ncRNA (TER). The ncRNA component adopts a conserved pseudoknot structure, essential for its function. In 2021, researchers achieved a significant milestone by constructing the first atomic model of human telomerase complex using cryo-EM, reaching a resolution of approximately 3.0–4.0 Å [89]. This structure revealed a bilobular architecture comprising a **catalytic core lobe, containing the telomerase reverse transcriptase (TERT) and the telomerase ncRNA (hTR)**, and a H/ACA ribonucleoprotein (RNP) lobe, formed by two sets of heterotetrameric H/ACA proteins and the Cajal body protein TCAB1 [89]. While both human and *Tetrahymena* telomerases share a conserved catalytic

core architecture surrounding the ncRNA, differences exist in their accessory components and regulatory mechanisms. The specific molecular interactions involved in telomerase recruitment and regulation differ between the two species, reflecting evolutionary divergence [89,90].

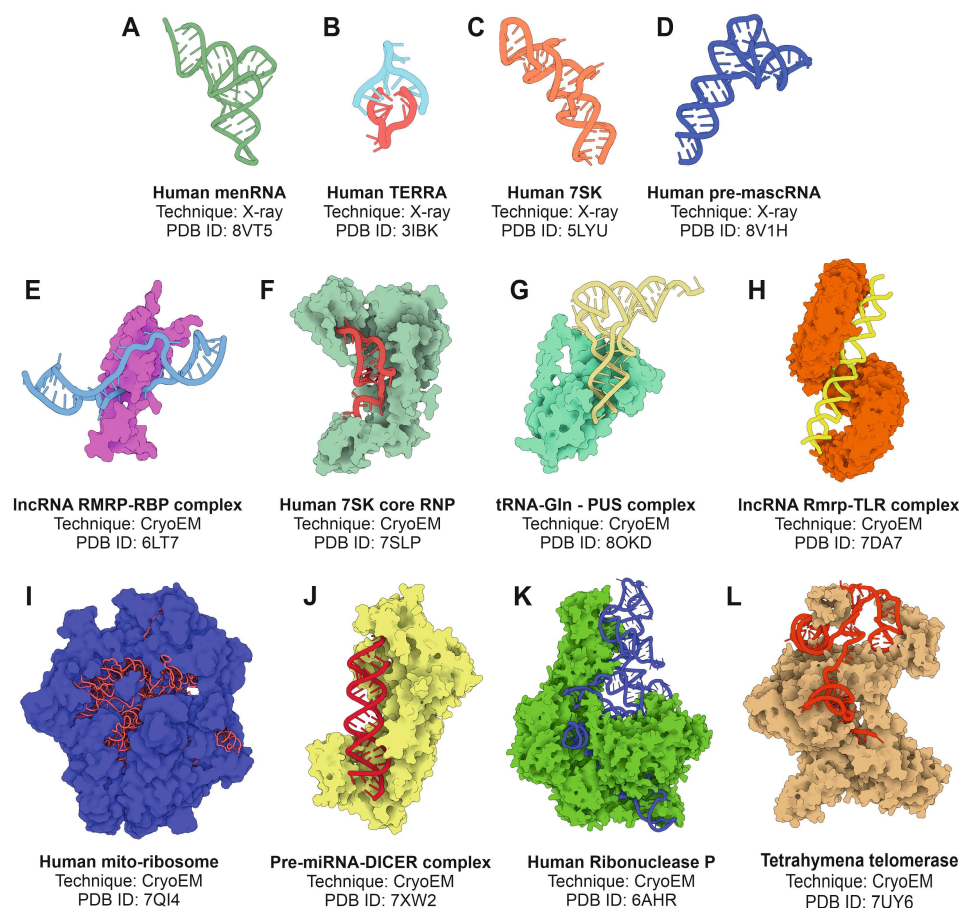


Figure 5. Selected examples of experimentally determined tertiary structures of ncRNAs or complexes of ncRNAs with proteins using X-ray crystallography and cryoEM methods. A, human memRNA; B, human telomere associated RNA; C, human 7SK RNA; D, human pre-mascRNA; E, complex between IncRNA Rmrp and an RNA-binding protein; F, human 7SK core ribonucleoprotein; G, glutamine-tRNA in complex with PUS protein; H, IncRNA Rmrp in complex with TLR; I, human spliceosome; J, human pre-miRNA complex with Dicer endonuclease; K, human ribonuclease P complex; and L, Tetrahymena telomerase. All the figures were prepared from the atomic coordinate data extracted from the PDB database and the 3D Protein Imaging application [91].

Recent advancements in cryo-EM have also provided detailed insights into the biogenesis and mechanism of action of small regulatory RNAs such as miRNAs, based on the structural characterization of RNA-protein complexes [92]. Drosha, an RNase III enzyme, initiates miRNA maturation by cleaving primary miRNAs (pri-miRNAs) into pre-miRNAs within the nucleus. It functions as part of the Microprocessor complex, which includes the DGCR8 protein. High-resolution cryo-EM structures of the full Drosha-DGCR8 complex bound to pri-miRNA have provided insights into its functional domains and interactions. The DGCR8 component contains an RNA-binding domain that recognizes and binds pri-miRNAs, positioning them for precise cleavage by Drosha [93]. This interaction is crucial for the accurate processing of pri-miRNAs into pre-miRNAs, a critical step in miRNA biogenesis. Dicer, another RNase III enzyme, is responsible for processing precursor miRNAs (pre-miRNAs) into mature miRNAs at the cytoplasm. In 2018, Liu and coworkers solved

the structure of human Dicer in complex with a pre-miRNA substrate by cryo-EM [94]. Additional studies revealed that human Dicer, in association with the trans-activation response RNA-binding protein (TRBP), forms a complex that interacts with pre-miRNA [95,96]. The structure highlighted two distinct conformations of the pre-miRNA within the complex: one adopting a perfect base-paired A-form helix and the other exhibiting a partially splayed structure. These conformations suggest a mechanism by which Dicer accommodates various pre-miRNA structures to ensure precise cleavage. Moreover, the DExD/H-box helicase domain of Dicer was observed to have a C-shaped architecture, implicating its role in regulating Dicer's activity during miRNA processing [94,95]. Cryo-EM has been also used to characterize other proteins involved in the miRNA biogenesis pathway, such as the Exportin 5, a protein required for the proper transport of pre-miRNAs from the nucleus to the cytoplasm. The structure of the complex between Exportin 5 and a pre-miRNA has been solved by Cryo-EM, and has been essential to understand the transport mechanism and the unspecific accommodation of miRNA precursors into the core of the shuttle protein [97].

4.3. *In Vivo Methods*

RNA secondary structures play crucial roles in the function and regulation of non-coding RNAs (ncRNAs). The *in vivo* determination of RNA secondary structures is particularly important because cellular environments introduce unique influences, such as molecular crowding, protein binding, and chemical modifications, which can differ significantly from *in vitro* conditions. Chemical probing methods have emerged as indispensable tools for characterizing RNA secondary structures in their native cellular contexts. These approaches exploit the reactivity of RNA nucleotides with small molecules, which differ depending on the nucleotide's structural context (single-stranded, paired, or tertiary interactions). *In vivo* probing captures RNA folding in the presence of cellular cofactors, metabolites, and interacting proteins, providing a more accurate reflection of functional RNA structures. Coupling probing with next-generation sequencing enables transcriptome-wide analysis of RNA structures. The dynamic structural states of ncRNAs can be monitored, shedding light on their roles in processes like translation regulation or splicing.

The original chemical probing method, SHAPE-Seq, involves the use of chemical reagents, such as 1-Methyl-7-nitroisatoic anhydride (1M7) or N-methylisatoic anhydride (NMIA), which preferentially react with flexible nucleotides in single-stranded regions of RNA. These reagents modify the 2'-hydroxyl group of the ribose ring, creating adducts that interfere with reverse transcription. By analyzing reverse transcription termination or mutation events through next-generation sequencing, the structural reactivity of RNA can be quantified. This method is well-suited for probing RNA *in vitro* but is limited to addressing *in vivo* dynamics and structural heterogeneity. Variants such as **icSHAPE** (*in vivo* click SHAPE) were developed to overcome the limitations of traditional SHAPE-Seq in analyzing RNA structure in live cells [98]. This method employs a cell-permeable SHAPE reagent, such as NAI-N3, which can penetrate cellular membranes to modify RNA in its native cellular context. After chemical modification, the azido moiety in NAI-N3 enables "click chemistry" to biotinylate modified RNA, allowing selective enrichment of labeled molecules. icSHAPE is instrumental in studying RNA structures influenced by protein interactions, cellular compartmentalization, and environmental stimuli. Another recent modification of chemical probing methods named SHAPE-MaP (mutational profiling) is a SHAPE-Seq variant that directly integrates chemical modification information into cDNA during reverse transcription [99]. Instead of terminating at modification sites, the reverse transcriptase incorporates mutations or misincorporations at modified nucleotides. These mutations are subsequently identified via sequencing [100]. SHAPE-MaP has several advantages, including reduced noise, improved quantitative accuracy, and compatibility with both *in vitro* and *in vivo* conditions. It also facilitates probing of RNA structure under a wide range of conditions, including complex environments like cell lysates [101].

Chemical probing techniques coupled with RNA sequencing have been applied for the secondary structure characterization of selected lncRNAs (Figure 6). Initial studies were focused on

the structural characterization of relatively small lncRNAs, such as the steroid receptor activator RNA (SRA), with approximately 800 nt in length. SRA is a co-activator of several sex hormone receptors, being strongly associated with some tumors. SHAPE-seq and RNase probing were used to characterize the evolutionary conserved structural domains of SRA [104]. In the following years, the development of new variants of the chemical probing techniques together with the improvement of the analysis software allowed the structural characterization of more complex lncRNAs. For instance, the structure of MALAT1 (approximately 8,400 nt in length), a long non-coding RNA (lncRNA) implicated in gene expression regulation, was probed using SHAPE-seq to identify its structural domains and interactions which are critical for its localization and function. In a recent work, the application of SHAPE-MaP for the characterization of MALAT1 in green monkeys and humans revealed a conserved structural pattern between the lncRNA in both species, despite the sequence differences [105]. Another interesting example is NEAT1, a lncRNA essential for the formation and maintenance of paraspeckles, nuclear bodies involved in RNA metabolism and processing [106]. SHAPE-seq and DMS probing were applied to NEAT1, revealing its highly modular structure and the existence of specific regions that were characterized as essential for paraspeckle assembly [107]. This study also contributed to describe the evolutionary divergences present in the structures of NEAT1 between humans and mouse, and proposed a functional model of NEAT1 lncRNA where the structural elements that contribute to paraspeckle assembly are based on long-range interactions rather than in conserved domains [107]. SHAPE-seq was also used for the structural characterization of HOTAIR (HOX transcript antisense RNA), a lncRNA involved in chromatin remodeling and gene silencing, particularly by guiding the Polycomb Repressive Complex 2 (PRC2) to specific genomic loci [102]. Interestingly, HOTAIR structure reveals a complex degree of structure organization, that is comparable to well-folded ncRNAs such as the group II intron or ribosomal RNAs. HOTAIR comprises four independent folding units, where two of them are predicted as protein-binding domains [102]. A similar modular organization was also observed in MEG3 (maternally expressed gene 3), a lncRNA that acts as a tumor suppressor by modulating the p53-mediated response. In this lncRNA, the folding is stabilized by distal motifs that interact by forming pseudoknot structures. A detailed SHAPE-seq study by Uroda and coworkers, showed that mutations able to disrupt this pseudoknot structures, impaired the MEG3-dependent stimulation of the p53 pathway by decreasing MEG3 stability [55].

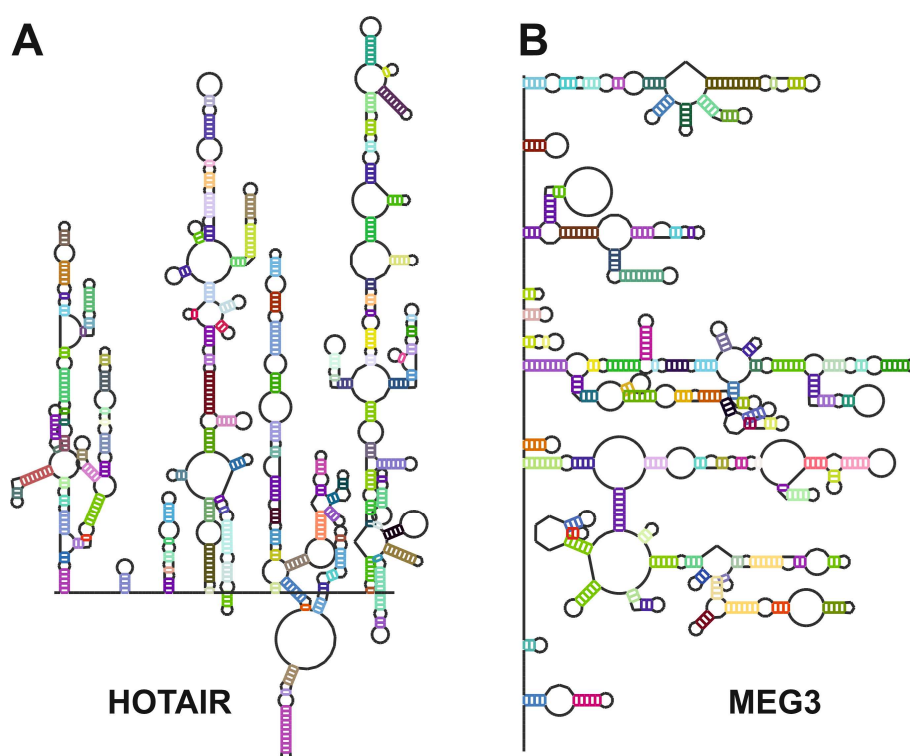


Figure 6. Secondary structures of HOTAIR [102] and MEG3 [55] lncRNAs as determined by chemical probing followed by next-generation sequencing. Sequences and secondary structures were retrieved from the lncRNA-folding repository [103]. Secondary structure maps were generated by the RNArtist software (<https://github.com/fjossinet/RNArtist>).

Other documented examples of the application of chemical probing technologies include the structural characterization of growth arrest specific 5 lncRNA (GAS5) by SHAPE-MaP [108], the application of SHAPE-seq and DMS-seq for the dissection of the secondary structure elements of the Xist lncRNA [109], and the studies of the dynamics of telomeric repeat-containing RNAs as important factors contributing to chromosome stability [90].

4.4. *In Silico Methods*

The hierarchical nature of RNA folding, where secondary structure elements are formed prior their combination into higher order tridimensional structures, makes secondary structure prediction very relevant for understanding RNA function [110]. However, *in silico* prediction methods showed serious limitations, because most RNA secondary structure prediction algorithms rely on thermodynamic models that assume RNA molecules adopt the structure with the lowest free energy [27]. While this assumption often holds, it overlooks kinetic folding pathways and metastable states, which can be functionally relevant. Consequently, predictions may not always capture the biologically active structure [19]. Moreover, our understanding of RNA folding rules is incomplete, particularly for complex motifs such as pseudoknots. Although some algorithms attempt to account for pseudoknots, these structures remain challenging to predict accurately due to their computational complexity and lack of comprehensive thermodynamic parameters. RNA folding is influenced by the cellular environment, including ionic conditions, molecular crowding, and interactions with proteins or other RNAs. Computational tools often operate under simplified conditions that do not fully replicate the cellular milieu, leading to discrepancies between predicted and actual structures [25,26].

4.4.1. Methods for RNA Secondary Structure Prediction

RNA secondary structure elements can be predicted by diverse families of algorithms that are designed with different strategies [111]. Thermodynamic methods predict RNA secondary structure by minimizing the free energy of folding, by applying thermodynamics and base-pairing probabilities. Representative computer tools that use thermodynamics for RNA secondary structure prediction are RNAfold [112], RNAstructure [113], Sfold [114] and Mfold [115]. These methods are especially efficient for the prediction of secondary structures in small RNAs (less than 500 nt). Thermodynamic-based methods have been extensively employed for the characterization of secondary structures in miRNA precursors. For instance, Fernandez and coworkers showed how the genetic variants of the miR-30c precursor modify the secondary structure of the processed hairpin loop, altering the processing efficiency of the Drosha-DGCR8 complex and the levels of mature miRNA. The computer prediction of secondary structures of the different precursors were correlated to the processing efficiency by the Drosha-DGCR8 complex of the different hairpin loops, showing a clear correlation between structural variants, RNA stability and processing outcome [116]. Moreover, the same prediction methods were used to study the influence of the secondary structure elements present in the 3'-UTR of mRNA transcripts and their relationship to miRNA regulatory activity. Analyzing different mutants of the 3'-UTR of lin41 gene from *Caenorhabditis elegans*, Long and coworkers predicted the secondary structure of the let-7 target hybrids using the Sfold tool and validated the regulatory effect of the miRNA. The work determined a potent effect on target recognition driven by the presence or absence of secondary structure elements in the miRNA target [117]. Additionally, comparative sequence analysis methods can be also used to determine secondary structures of ncRNAs. In this approach, covariation analysis can identify the paired residues that evolve in a coordinate manner suggesting base-pairing interactions [118]. Databases such as Rfam

were built based on comparative sequence analysis to annotate ncRNA families and predict conserved secondary structure elements in ncRNAs [119].

Recently, machine learning techniques have transformed RNA secondary structure prediction by leveraging large datasets and powerful algorithms to identify complex patterns. Machine learning models, particularly deep learning approaches, can integrate diverse sources of data, including primary RNA sequences, chemical probing experiments, and evolutionary conservation [120]. These models bypass some of the limitations of thermodynamic methods by learning directly from empirical data. However, the application of these methods to the prediction of secondary structures of ncRNAs is not widespread [121]. Among the different deep learning methods applied to the prediction of RNA secondary structures, supervised models are trained on datasets of RNA sequences with experimentally validated secondary structures. Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) have been employed to capture local and global dependencies in RNA sequences. For example, the SPOT-RNA model uses a combination of CNNs and long short-term memory (LSTM) networks to predict base-pairing probabilities with high accuracy [122]. In **graph neural network models (GNNs)**, RNA molecules can be represented as graphs, with nucleotides as nodes and base-pairing or spatial proximities as edges. GNNs are particularly suited for modeling such relationships. RNAformer, for example, uses graph attention mechanisms to predict secondary structure while accounting for long-range interactions [123].

4.4.2. Methods for RNA Tertiary Structure Prediction

RNA tertiary structure prediction remains a formidable challenge in computational biology due to several intrinsic and extrinsic factors. These difficulties stem from the complex nature of RNA molecules, the limitations of current computational methodologies, and the intricate interplay of physicochemical forces governing RNA folding [124]. Intrinsic challenges related to RNA structure include its **structural flexibility**, extensive presence of **long-range Interactions** and **multitude of stable conformations** [125]. Computer methodologies and protocols for RNA tertiary structure prediction are limited mainly by the suboptimal accuracy of current force fields tailored for RNA, that prevents a proper energy-based modeling of RNA structures [126]. Moreover, the large size of specific families of ncRNA molecules such as lncRNAs and their intricate interactions require extensive computational resources, increasing the time required for accurate calculations. The scientific community, aware of all these experimental hurdles, has established some projects like the RNA-Puzzles initiative, started in 2010, that emerged as a collaborative effort to advance the understanding of RNA structure prediction. The project was inspired by the success of similar initiatives like CASP (Critical Assessment of Protein Structure Prediction) in the field of protein modeling [127]. RNA-Puzzles serves as a blind benchmarking platform where participants are tasked with predicting RNA structures based solely on sequence information, without prior knowledge of experimentally determined structures. These predictions are then evaluated against experimentally derived models, enabling the comparison of different computational approaches [128].

Among the wide range of prediction methods used to model RNA 3D structure, homology modeling relies on the use of structural similarity between RNAs to predict their 3D structure. Tools like ModeRNA are based on the use of already known RNA structures as templates to predict the tertiary configuration of homologous sequences [129]. Other methods such the fragment assembly algorithms predict RNA tertiary structures by assembling smaller structural motifs, combining them into a higher-order structure. Pioneer algorithms such as the MC-Fold/MC-Sym pipeline is a good example, that was used to predict the 3D structures of different pre-miRNAs [130]. More recently, the same experimental strategy was used to develop algorithms such as FARFAR2, a new fragment assembly of RNA with full-atom refinement [131]. This method was successfully applied for the structural prediction of small and well-structured ncRNAs such as ribozymes and riboswitches [132].

Achieving high accuracy in RNA structure prediction often necessitates combining information from different structural scales. Multiscale and hybrid methods that integrate secondary and tertiary structure prediction techniques have emerged as powerful approaches to address this challenge.

RNAComposer is a prominent example that combines thermodynamic principles with fragment assembly techniques to predict RNA 3D structures. RNAComposer uses secondary structure predictions as a starting point, ensuring that the base-pairing interactions are thermodynamically feasible [133]. After these predictions, the software maps predicted secondary structures onto a database of known RNA motifs to assemble a final 3D model. The hybrid methodology allows RNAComposer to incorporate both local base-pairing information and global structural motifs, yielding accurate 3D model [134]. Hybrid and multiscale methods have demonstrated significant improvements in prediction accuracy. They enable the modeling of complex RNA structures, including ribozymes, riboswitches, and large non-coding RNAs [135]. The integration of secondary and tertiary prediction methods is particularly valuable for modeling large RNA molecules where accurate tertiary structure prediction is computationally challenging [136].

Molecular dynamics is a computational technique that models the physical movements of atoms and molecules over time. It relies on solving Newton's equations of motion for a system of particles, where the interactions are described by a potential energy function, or force field. For RNA, MD simulations provide insights into folding pathways, conformational flexibility, and the effects of sequence mutations on structure and function [137]. Accurate modeling of RNA requires specialized force fields that capture the unique chemical and physical properties of nucleotides to accurately reproduce RNA's backbone conformations, hydrogen bonding patterns, and base-stacking interactions [138,139]. RNA folding involves a hierarchical process, where local secondary structures such as helices and loops form first, followed by the establishment of tertiary contacts. MD simulations can capture these folding events by simulating the molecule in explicit solvent conditions, mimicking the cellular environment. Enhanced sampling techniques, such as replica exchange molecular dynamics (REMD) and meta-dynamics, are often employed to overcome the high energy barriers associated with RNA folding and sample a broader conformational space [140]. While all-atom MD simulations offer high-resolution insights, their computational cost limits their implementation in long time scales required to study phenomena such as RNA folding [141]. Coarse-grained models, which simplify RNA representation by grouping atoms into larger units, enable simulations of larger systems and longer timescales [142]. Due to the specific transient nature of the miRNA-based regulation, molecular dynamics has been employed for the characterization of the stability of miRNA-mRNA complexes [143]. Moreover, molecular dynamics simulations were also used to study the stability of the heterologous complex established between the miRNAs, their targets, and the core of the RNA-induced silencing complex (RISC), Ago2 [144]. Interestingly, all the theoretical simulations performed during miRNA-mRNA interactions concluded that the spectrum of a miRNA's potential targets is different from what is currently anticipated by the canonical miRNA seed-model. These findings could explain the relative promiscuity of miRNA regulatory action and its transient nature, which could be based on the observed flexible regulatory effect over mRNA transcripts [143,144].

5. Perspectives and Further Developments

The application of structural biology techniques to non-coding RNAs (ncRNAs) represents a rapidly advancing frontier, driven by the recognition of ncRNAs' diverse biological roles. The development of advanced structural biology, new devices and protocols related to cryo-electron microscopy (cryo-EM) and X-ray crystallography methods, has enabled high-resolution characterization of complex ncRNA structures. These techniques have revealed intricate folding patterns and tertiary interactions critical to ncRNA function. More recently, integrative approaches combining structural data with computational modeling and single-molecule experiments have expanded our capacity to resolve dynamic and transient ncRNA conformations. Cryo-EM has revolutionized the field by allowing the study of large ribonucleoprotein (RNP) complexes, such as the spliceosome and ribosome, at near-atomic resolution. Simultaneously, advancements in computational RNA structure prediction tools, powered by machine learning, have enhanced the accuracy of secondary and tertiary structure predictions. These tools complement experimental

techniques, enabling the characterization of challenging ncRNAs and their complexes. Looking ahead, several key areas promise to drive further advancements in this field.

- **Integration of multiscale data:** combining atomic-level structural information with systemic data on ncRNA localization, interaction networks, and dynamics will enable a holistic understanding of ncRNA function. Hybrid approaches integrating cryo-EM and chemical probing with single-cell RNA sequencing and spatial transcriptomics are particularly promising.
- **High-throughput structural characterization:** automation in cryo-EM and microfluidics-based methods for RNA crystallography could facilitate the high-throughput determination of ncRNA structures, accelerating the discovery of novel functional motifs.
- **Dynamic and contextual studies:** capturing ncRNA structures in their native cellular environment remains a significant challenge. Emerging techniques, such as cryo-electron tomography (cryo-ET) and in situ structural studies, aim to bridge this gap by visualizing ncRNAs within intact cells.
- **Functional modulation and rational drug design:** structural insights into ncRNAs have profound implications for drug discovery. Small molecules targeting ncRNA structures or their interactions with proteins could provide new therapeutic avenues for diseases associated with dysregulated ncRNAs, such as cancer and neurodegenerative disorders.
- **Evolutionary perspectives:** structural comparisons across species can reveal conserved motifs and inform functional hypotheses. Integrating structural biology with evolutionary genomics will help identify universally important ncRNA structures and their roles in diverse organisms.
- **Artificial intelligence:** the use of deep-learning approaches to infer ncRNA structure and function will increase our understanding of their functions, increasing the knowledge about the molecular players related to cell physiology and human disease.

Author Contributions: Conceptualization, A.L.L. and F.J.E.; formal analysis, F.J.E. and A.L.L.; investigation, F.J.E. and A.L.L.; resources, A.L.L. and F.J.E.; data curation, F.J.E. and A.L.L.; writing—original draft preparation, A.L.L. and F.J.E.; writing—review and editing, A.L.L. and F.J.E.; visualization, A.L.L. and F.J.E. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: The authors acknowledge the support and technical assistance from Francisco Enguita jr. in the conceptualization and design of the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Davis, C.A.; Hitz, B.C.; Sloan, C.A.; Chan, E.T.; Davidson, J.M.; Gabdank, I.; Hilton, J.A.; Jain, K.; Baymuradov, U.K.; Narayanan, A.K., et al. The Encyclopedia of DNA elements (ENCODE): data portal update. *Nucleic Acids Res* **2018**, *46*, D794–D801, doi:10.1093/nar/gkx1081.
2. Mattick, J.S.; Gagen, M.J. The evolution of controlled multitasked gene networks: the role of introns and other noncoding RNAs in the development of complex organisms. *Mol Biol Evol* **2001**, *18*, 1611–1630, doi:10.1093/oxfordjournals.molbev.a003951.
3. Villa, T.; Porrua, O. Pervasive transcription: a controlled risk. *FEBS J* **2023**, *290*, 3723–3736, doi:10.1111/febs.16530.
4. Carroll, S.B. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* **2008**, *134*, 25–36, doi:10.1016/j.cell.2008.06.030.
5. Levine, M.; Tjian, R. Transcription regulation and animal diversity. *Nature* **2003**, *424*, 147–151, doi:10.1038/nature01763.
6. Mudge, J.M.; Carbonell-Sala, S.; Diekhans, M.; Martinez, J.G.; Hunt, T.; Jungreis, I.; Loveland, J.E.; Arnan, C.; Barnes, I.; Bennett, R., et al. GENCODE 2025: reference gene annotation for human and mouse. *Nucleic Acids Res* **2024**, 10.1093/nar/gkae1078, doi:10.1093/nar/gkae1078.

7. Chen, L.L.; Kim, V.N. Small and long non-coding RNAs: Past, present, and future. *Cell* **2024**, *187*, 6451-6485, doi:10.1016/j.cell.2024.10.024.
8. Assmann, S.M.; Chou, H.L.; Bevilacqua, P.C. Rock, scissors, paper: How RNA structure informs function. *Plant Cell* **2023**, *35*, 1671-1707, doi:10.1093/plcell/koad026.
9. Tinoco, I., Jr.; Bustamante, C. How RNA folds. *J Mol Biol* **1999**, *293*, 271-281, doi:10.1006/jmbi.1999.3001.
10. Mathews, D.H.; Turner, D.H. Prediction of RNA secondary structure by free energy minimization. *Curr Opin Struct Biol* **2006**, *16*, 270-278, doi:10.1016/j.sbi.2006.05.010.
11. Juan, V.; Wilson, C. RNA secondary structure prediction based on free energy and phylogenetic analysis. *J Mol Biol* **1999**, *289*, 935-947, doi:10.1006/jmbi.1999.2801.
12. Lemieux, S.; Major, F. RNA canonical and non-canonical base pairing types: a recognition method and complete repertoire. *Nucleic Acids Res* **2002**, *30*, 4250-4263, doi:10.1093/nar/gkf540.
13. Brion, P.; Westhof, E. Hierarchy and dynamics of RNA folding. *Annu Rev Biophys Biomol Struct* **1997**, *26*, 113-137, doi:10.1146/annurev.biophys.26.1.113.
14. Westhof, E.; Fritsch, V. RNA folding: beyond Watson-Crick pairs. *Structure* **2000**, *8*, R55-65, doi:10.1016/s0969-2126(00)00112-x.
15. Tinoco, I., Jr.; Li, P.T.; Bustamante, C. Determination of thermodynamics and kinetics of RNA reactions by force. *Q Rev Biophys* **2006**, *39*, 325-360, doi:10.1017/S0033583506004446.
16. Liphardt, J.; Onoa, B.; Smith, S.B.; Tinoco, I., Jr.; Bustamante, C. Reversible unfolding of single RNA molecules by mechanical force. *Science* **2001**, *292*, 733-737, doi:10.1126/science.1058498.
17. Viereg, J.; Cheng, W.; Bustamante, C.; Tinoco, I., Jr. Measurement of the effect of monovalent cations on RNA hairpin stability. *J Am Chem Soc* **2007**, *129*, 14966-14973, doi:10.1021/ja074809o.
18. Guckian, K.M.; Schweitzer, B.A.; Ren, R.X.; Sheils, C.J.; Paris, P.L.; Tahmassebi, D.C.; Kool, E.T. Experimental Measurement of Aromatic Stacking Affinities in the Context of Duplex DNA. *J Am Chem Soc* **1996**, *118*, 8182-8183, doi:10.1021/ja961733f.
19. Serra, M.J.; Turner, D.H. Predicting thermodynamic properties of RNA. *Methods Enzymol* **1995**, *259*, 242-261, doi:10.1016/0076-6879(95)59047-1.
20. Broyde, S. Setting the stage for predicting RNA thermodynamic properties and their structural components. *Biophys J* **1996**, *70*, 1571-1572, doi:10.1016/S0006-3495(96)79722-9.
21. Lilley, D.M. Structures of helical junctions in nucleic acids. *Q Rev Biophys* **2000**, *33*, 109-159, doi:10.1017/s0033583500003590.
22. Herschlag, D. RNA chaperones and the RNA folding problem. *J Biol Chem* **1995**, *270*, 20871-20874, doi:10.1074/jbc.270.36.20871.
23. Lilley, D.M. The origins of RNA catalysis in ribozymes. *Trends Biochem Sci* **2003**, *28*, 495-501, doi:10.1016/S0968-0004(03)00191-9.
24. Draper, D.E.; Grilley, D.; Soto, A.M. Ions and RNA folding. *Annu Rev Biophys Biomol Struct* **2005**, *34*, 221-243, doi:10.1146/annurev.biophys.34.040204.144511.
25. Draper, D.E. RNA folding: thermodynamic and molecular descriptions of the roles of ions. *Biophys J* **2008**, *95*, 5489-5495, doi:10.1529/biophysj.108.131813.
26. Denesyuk, N.A.; Hori, N.; Thirumalai, D. Molecular Simulations of Ion Effects on the Thermodynamics of RNA Folding. *J Phys Chem B* **2018**, *122*, 11860-11867, doi:10.1021/acs.jpcb.8b08142.
27. dell'Erba, M.G.; Zemba, G.R. Thermodynamics of a model for RNA folding. *Phys Rev E Stat Nonlin Soft Matter Phys* **2009**, *79*, 011913, doi:10.1103/PhysRevE.79.011913.
28. Leontis, N.B.; Lescoute, A.; Westhof, E. The building blocks and motifs of RNA architecture. *Curr Opin Struct Biol* **2006**, *16*, 279-287, doi:10.1016/j.sbi.2006.05.009.
29. Gregorian, R.S., Jr.; Crothers, D.M. Determinants of RNA hairpin loop-loop complex stability. *J Mol Biol* **1995**, *248*, 968-984, doi:10.1006/jmbi.1995.0275.
30. Dale, T.; Smith, R.; Serra, M.J. A test of the model to predict unusually stable RNA hairpin loop stability. *RNA* **2000**, *6*, 608-615, doi:10.1017/s1355838200992495.
31. Krol, J.; Sobczak, K.; Wilczynska, U.; Drath, M.; Jasinska, A.; Kaczynska, D.; Krzyzosiak, W.J. Structural features of microRNA (miRNA) precursors and their relevance to miRNA biogenesis and small interfering RNA/short hairpin RNA design. *J Biol Chem* **2004**, *279*, 42230-42239, doi:10.1074/jbc.M404931200.

32. Rice, G.M.; Shivashankar, V.; Ma, E.J.; Baryza, J.L.; Nutiu, R. Functional Atlas of Primary miRNA Maturation by the Microprocessor. *Mol Cell* **2020**, *80*, 892-902 e894, doi:10.1016/j.molcel.2020.10.028.
33. Garg, A.; Shang, R.; Cvetanovic, T.; Lai, E.C.; Joshua-Tor, L. The structural landscape of Microprocessor-mediated processing of pri-let-7 miRNAs. *Mol Cell* **2024**, *84*, 4175-4190 e4176, doi:10.1016/j.molcel.2024.09.008.
34. Xiong, X.; Kang, X.; Zheng, Y.; Yue, S.; Zhu, S. Identification of loop nucleotide polymorphisms affecting microRNA processing and function. *Mol Cells* **2013**, *36*, 518-526, doi:10.1007/s10059-013-0171-1.
35. Dang, T.L.; Le, C.T.; Le, M.N.; Nguyen, T.D.; Nguyen, T.L.; Bao, S.; Li, S.; Nguyen, T.A. Select amino acids in DGCR8 are essential for the UGU-pri-miRNA interaction and processing. *Commun Biol* **2020**, *3*, 344, doi:10.1038/s42003-020-1071-5.
36. McIntyre, G.J.; Yu, Y.H.; Lomas, M.; Fanning, G.C. The effects of stem length and core placement on shRNA activity. *BMC Mol Biol* **2011**, *12*, 34, doi:10.1186/1471-2199-12-34.
37. Zhang, D.; Jue, D.; Smith, N.; Zhong, C.; Finnegan, E.J.; de Feyter, R.; Wang, M.B.; Greaves, I. Asymmetric bulges within hairpin RNA transgenes influence small RNA size, secondary siRNA production and viral defence. *Nucleic Acids Res* **2024**, *52*, 9904-9916, doi:10.1093/nar/gkae573.
38. Wang, C.; Wang, L.; Ding, Y.; Lu, X.; Zhang, G.; Yang, J.; Zheng, H.; Wang, H.; Jiang, Y.; Xu, L. LncRNA Structural Characteristics in Epigenetic Regulation. *Int J Mol Sci* **2017**, *18*, doi:10.3390/ijms18122659.
39. Cerase, A.; Tartaglia, G.G. Long non-coding RNA-polycomb intimate rendezvous. *Open Biol* **2020**, *10*, 200126, doi:10.1098/rsob.200126.
40. Brown, J.A.; Bulkley, D.; Wang, J.; Valenstein, M.L.; Yario, T.A.; Steitz, T.A.; Steitz, J.A. Structural insights into the stabilization of MALAT1 noncoding RNA by a bipartite triple helix. *Nat Struct Mol Biol* **2014**, *21*, 633-640, doi:10.1038/nsmb.2844.
41. Walter, A.E.; Turner, D.H.; Kim, J.; Lyttle, M.H.; Muller, P.; Mathews, D.H.; Zuker, M. Coaxial stacking of helices enhances binding of oligoribonucleotides and improves predictions of RNA folding. *Proc Natl Acad Sci U S A* **1994**, *91*, 9218-9222, doi:10.1073/pnas.91.20.9218.
42. Mondal, M.; Mukherjee, S.; Halder, S.; Bhattacharyya, D. Stacking geometry for non-canonical G:U wobble base pair containing dinucleotide sequences in RNA: dispersion-corrected DFT-D study. *Biopolymers* **2015**, *103*, 328-338, doi:10.1002/bip.22616.
43. Hermann, T.; Patel, D.J. RNA bulges as architectural and recognition motifs. *Structure* **2000**, *8*, R47-54, doi:10.1016/s0969-2126(00)00110-6.
44. Correll, C.C.; Munishkin, A.; Chan, Y.L.; Ren, Z.; Wool, I.G.; Steitz, T.A. Crystal structure of the ribosomal RNA domain essential for binding elongation factors. *Proc Natl Acad Sci U S A* **1998**, *95*, 13436-13441, doi:10.1073/pnas.95.23.13436.
45. Sargsyan, K.; Lim, C. Arrangement of 3D structural motifs in ribosomal RNA. *Nucleic Acids Res* **2010**, *38*, 3512-3522, doi:10.1093/nar/gkq074.
46. Macrae, I.J.; Zhou, K.; Li, F.; Repic, A.; Brooks, A.N.; Cande, W.Z.; Adams, P.D.; Doudna, J.A. Structural basis for double-stranded RNA processing by Dicer. *Science* **2006**, *311*, 195-198, doi:10.1126/science.1121638.
47. Zhao, J.; Sun, B.K.; Erwin, J.A.; Song, J.J.; Lee, J.T. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science* **2008**, *322*, 750-756, doi:10.1126/science.1163045.
48. Tripathi, V.; Ellis, J.D.; Shen, Z.; Song, D.Y.; Pan, Q.; Watt, A.T.; Freier, S.M.; Bennett, C.F.; Sharma, A.; Bubulya, P.A., et al. The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol Cell* **2010**, *39*, 925-938, doi:10.1016/j.molcel.2010.08.011.
49. Staple, D.W.; Butcher, S.E. Pseudoknots: RNA structures with diverse functions. *PLoS Biol* **2005**, *3*, e213, doi:10.1371/journal.pbio.0030213.
50. Huang, F.W.; Li, L.Y.; Reidys, C.M. Sequence-structure relations of pseudoknot RNA. *BMC Bioinformatics* **2009**, *10 Suppl 1*, S39, doi:10.1186/1471-2105-10-S1-S39.
51. Green, L.; Kim, C.H.; Bustamante, C.; Tinoco, I., Jr. Characterization of the mechanical unfolding of RNA pseudoknots. *J Mol Biol* **2008**, *375*, 511-528, doi:10.1016/j.jmb.2007.05.058.
52. Hsu, C.F.; Chang, K.C.; Chen, Y.L.; Hsieh, P.S.; Lee, A.I.; Tu, J.Y.; Chen, Y.T.; Wen, J.D. Formation of frameshift-stimulating RNA pseudoknots is facilitated by remodeling of their folding intermediates. *Nucleic Acids Res* **2021**, *49*, 6941-6957, doi:10.1093/nar/gkab512.

53. Su, L.; Chen, L.; Egli, M.; Berger, J.M.; Rich, A. Minor groove RNA triplex in the crystal structure of a ribosomal frameshifting viral pseudoknot. *Nat Struct Biol* **1999**, *6*, 285-292, doi:10.1038/6722.
54. Theimer, C.A.; Blois, C.A.; Feigon, J. Structure of the human telomerase RNA pseudoknot reveals conserved tertiary interactions essential for function. *Mol Cell* **2005**, *17*, 671-682, doi:10.1016/j.molcel.2005.01.017.
55. Uroda, T.; Anastasakou, E.; Rossi, A.; Teulon, J.M.; Pellequer, J.L.; Annibale, P.; Pessey, O.; Inga, A.; Chillon, I.; Marcia, M. Conserved Pseudoknots in lncRNA MEG3 Are Essential for Stimulation of the p53 Pathway. *Mol Cell* **2019**, *75*, 982-995 e989, doi:10.1016/j.molcel.2019.07.025.
56. Turner, D.H. Thermodynamics of base pairing. *Curr Opin Struct Biol* **1996**, *6*, 299-304, doi:10.1016/s0959-440x(96)80047-9.
57. Wilusz, J.E.; Freier, S.M.; Spector, D.L. 3' end processing of a long nuclear-retained noncoding RNA yields a tRNA-like cytoplasmic RNA. *Cell* **2008**, *135*, 919-932, doi:10.1016/j.cell.2008.10.012.
58. Tsai, M.C.; Manor, O.; Wan, Y.; Mosammaparast, N.; Wang, J.K.; Lan, F.; Shi, Y.; Segal, E.; Chang, H.Y. Long noncoding RNA as modular scaffold of histone modification complexes. *Science* **2010**, *329*, 689-693, doi:10.1126/science.1192002.
59. Yang, L.; Froberg, J.E.; Lee, J.T. Long noncoding RNAs: fresh perspectives into the RNA world. *Trends Biochem Sci* **2014**, *39*, 35-43, doi:10.1016/j.tibs.2013.10.002.
60. Naganuma, T.; Nakagawa, S.; Tanigawa, A.; Sasaki, Y.F.; Goshima, N.; Hirose, T. Alternative 3'-end processing of long noncoding RNA initiates construction of nuclear paraspeckles. *EMBO J* **2012**, *31*, 4020-4034, doi:10.1038/emboj.2012.251.
61. Rivas, E.; Clements, J.; Eddy, S.R. A statistical test for conserved RNA structure shows lack of evidence for structure in lncRNAs. *Nat Methods* **2017**, *14*, 45-48, doi:10.1038/nmeth.4066.
62. Sussman, J.L.; Lin, D.; Jiang, J.; Manning, N.O.; Prilusky, J.; Ritter, O.; Abola, E.E. Protein Data Bank (PDB): database of three-dimensional structural information of biological macromolecules. *Acta Crystallogr D Biol Crystallogr* **1998**, *54*, 1078-1084, doi:10.1107/s0907444998009378.
63. Lawson, C.L.; Berman, H.M.; Chen, L.; Vallat, B.; Zirbel, C.L. The Nucleic Acid Knowledgebase: a new portal for 3D structural information about nucleic acids. *Nucleic Acids Res* **2024**, *52*, D245-D254, doi:10.1093/nar/gkad957.
64. Neidle, S. Beyond the double helix: DNA structural diversity and the PDB. *J Biol Chem* **2021**, *296*, 100553, doi:10.1016/j.jbc.2021.100553.
65. Lietzke, S.E.; Barnes, C.L.; Kundrot, C.E. Crystallization and structure determination of RNA. *Curr Opin Struct Biol* **1995**, *5*, 645-649, doi:10.1016/0959-440x(95)80057-3.
66. Ferre-D'Amare, A.R.; Zhou, K.; Doudna, J.A. A general module for RNA crystallization. *J Mol Biol* **1998**, *279*, 621-631, doi:10.1006/jmbi.1998.1789.
67. Golden, B.L.; Kundrot, C.E. RNA crystallization. *J Struct Biol* **2003**, *142*, 98-107, doi:10.1016/s1047-8477(03)00042-x.
68. Vicens, Q.; Gooding, A.R.; Laederach, A.; Cech, T.R. Local RNA structural changes induced by crystallization are revealed by SHAPE. *RNA* **2007**, *13*, 536-548, doi:10.1261/rna.400207.
69. Zhang, J.; Ferre-D'Amare, A.R. Post-crystallization Improvement of RNA Crystal Diffraction Quality. *Methods Mol Biol* **2015**, *1316*, 13-24, doi:10.1007/978-1-4939-2730-2_2.
70. Carrasco, N.; Buzin, Y.; Tyson, E.; Halpert, E.; Huang, Z. Selenium derivatization and crystallization of DNA and RNA oligonucleotides for X-ray crystallography using multiple anomalous dispersion. *Nucleic Acids Res* **2004**, *32*, 1638-1646, doi:10.1093/nar/gkh325.
71. Zhang, J.; Ferre-D'Amare, A.R. Post-crystallization Improvement of RNA Crystals by Synergistic Ion Exchange and Dehydration. *Bio Protoc* **2015**, *5*.
72. Pley, H.W.; Flaherty, K.M.; McKay, D.B. Three-dimensional structure of a hammerhead ribozyme. *Nature* **1994**, *372*, 68-74, doi:10.1038/372068a0.
73. Scott, W.G.; Finch, J.T.; Klug, A. The crystal structure of an all-RNA hammerhead ribozyme: a proposed mechanism for RNA catalytic cleavage. *Cell* **1995**, *81*, 991-1002, doi:10.1016/s0092-8674(05)80004-2.
74. Edwards, A.L.; Garst, A.D.; Batey, R.T. Determining structures of RNA aptamers and riboswitches by X-ray crystallography. *Methods Mol Biol* **2009**, *535*, 135-163, doi:10.1007/978-1-59745-557-2_9.

75. Serganov, A.; Polonskaia, A.; Phan, A.T.; Breaker, R.R.; Patel, D.J. Structural basis for gene regulation by a thiamine pyrophosphate-sensing riboswitch. *Nature* **2006**, *441*, 1167-1171, doi:10.1038/nature04740.
76. Itoh, Y.; Chiba, S.; Sekine, S.; Yokoyama, S. Crystal structure of human selenocysteine tRNA. *Nucleic Acids Res* **2009**, *37*, 6259-6268, doi:10.1093/nar/gkp648.
77. Zong, X.; Nakagawa, S.; Freier, S.M.; Fei, J.; Ha, T.; Prasanth, S.G.; Prasanth, K.V. Natural antisense RNA promotes 3' end processing and maturation of MALAT1 lncRNA. *Nucleic Acids Res* **2016**, *44*, 2898-2908, doi:10.1093/nar/gkw047.
78. Skeparnias, I.; Zhang, J. Structural basis of NEAT1 lncRNA maturation and menRNA instability. *Nat Struct Mol Biol* **2024**, *31*, 1650-1654, doi:10.1038/s41594-024-01361-z.
79. Ma, H.; Jia, X.; Zhang, K.; Su, Z. Cryo-EM advances in RNA structure determination. *Signal Transduct Target Ther* **2022**, *7*, 58, doi:10.1038/s41392-022-00916-0.
80. Bonilla, S.L.; Kieft, J.S. The promise of cryo-EM to explore RNA structural dynamics. *J Mol Biol* **2022**, *434*, 167802, doi:10.1016/j.jmb.2022.167802.
81. Chen, X.; Wang, L.; Xie, J.; Nowak, J.S.; Luo, B.; Zhang, C.; Jia, G.; Zou, J.; Huang, D.; Glatt, S., et al. RNA sample optimization for cryo-EM analysis. *Nat Protoc* **2024**, 10.1038/s41596-024-01072-1, doi:10.1038/s41596-024-01072-1.
82. Yan, C.; Hang, J.; Wan, R.; Huang, M.; Wong, C.C.; Shi, Y. Structure of a yeast spliceosome at 3.6-angstrom resolution. *Science* **2015**, *349*, 1182-1191, doi:10.1126/science.aac7629.
83. Yan, C.; Wan, R.; Bai, R.; Huang, G.; Shi, Y. Structure of a yeast activated spliceosome at 3.5 Å resolution. *Science* **2016**, *353*, 904-911, doi:10.1126/science.aag0291.
84. Yan, C.; Wan, R.; Bai, R.; Huang, G.; Shi, Y. Structure of a yeast step II catalytically activated spliceosome. *Science* **2017**, *355*, 149-155, doi:10.1126/science.aak9979.
85. Bertram, K.; Agafonov, D.E.; Dybkov, O.; Haselbach, D.; Leelaram, M.N.; Will, C.L.; Urlaub, H.; Kastner, B.; Luhrmann, R.; Stark, H. Cryo-EM Structure of a Pre-catalytic Human Spliceosome Primed for Activation. *Cell* **2017**, *170*, 701-713 e711, doi:10.1016/j.cell.2017.07.011.
86. Bertram, K.; Agafonov, D.E.; Liu, W.T.; Dybkov, O.; Will, C.L.; Hartmuth, K.; Urlaub, H.; Kastner, B.; Stark, H.; Luhrmann, R. Cryo-EM structure of a human spliceosome activated for step 2 of splicing. *Nature* **2017**, *542*, 318-323, doi:10.1038/nature21079.
87. Zhang, X.; Yan, C.; Hang, J.; Finci, L.I.; Lei, J.; Shi, Y. An Atomic Structure of the Human Spliceosome. *Cell* **2017**, *169*, 918-929 e914, doi:10.1016/j.cell.2017.04.033.
88. Jiang, J.; Wang, Y.; Susac, L.; Chan, H.; Basu, R.; Zhou, Z.H.; Feigon, J. Structure of Telomerase with Telomeric DNA. *Cell* **2018**, *173*, 1179-1190 e1113, doi:10.1016/j.cell.2018.04.038.
89. Liu, B.; He, Y.; Wang, Y.; Song, H.; Zhou, Z.H.; Feigon, J. Structure of active human telomerase with telomere shelterin protein TPP1. *Nature* **2022**, *604*, 578-583, doi:10.1038/s41586-022-04582-8.
90. Palka, C.; Forino, N.M.; Hentschel, J.; Das, R.; Stone, M.D. Folding heterogeneity in the essential human telomerase RNA three-way junction. *RNA* **2020**, *26*, 1787-1800, doi:10.1261/rna.077255.120.
91. Tomasello, G.; Armenia, I.; Molla, G. The Protein Imager: a full-featured online molecular viewer interface with server-side HQ-rendering capabilities. *Bioinformatics* **2020**, *36*, 2909-2911, doi:10.1093/bioinformatics/btaa009.
92. Leitao, A.L.; Enguita, F.J. A Structural View of miRNA Biogenesis and Function. *Noncoding RNA* **2022**, *8*, doi:10.3390/ncrna8010010.
93. Partin, A.C.; Zhang, K.; Jeong, B.C.; Herrell, E.; Li, S.; Chiu, W.; Nam, Y. Cryo-EM Structures of Human Drosha and DGCR8 in Complex with Primary MicroRNA. *Mol Cell* **2020**, *78*, 411-422 e414, doi:10.1016/j.molcel.2020.02.016.
94. Liu, Z.; Wang, J.; Cheng, H.; Ke, X.; Sun, L.; Zhang, Q.C.; Wang, H.W. Cryo-EM Structure of Human Dicer and Its Complexes with a Pre-miRNA Substrate. *Cell* **2018**, *173*, 1191-1203, doi:10.1016/j.cell.2018.05.031.
95. Lee, H.; Roh, S.H. Cryo-EM structures of human DICER dicing a pre-miRNA substrate. *FEBS J* **2024**, *291*, 3072-3079, doi:10.1111/febs.17048.
96. Nguyen, T.D.; Trinh, T.A.; Bao, S.; Nguyen, T.A. Secondary structure RNA elements control the cleavage activity of DICER. *Nat Commun* **2022**, *13*, 2138, doi:10.1038/s41467-022-29822-3.

97. Okada, C.; Yamashita, E.; Lee, S.J.; Shibata, S.; Katahira, J.; Nakagawa, A.; Yoneda, Y.; Tsukihara, T. A high-resolution structure of the pre-microRNA nuclear export machinery. *Science* **2009**, *326*, 1275-1279, doi:10.1126/science.1178705.
98. Spitale, R.C.; Flynn, R.A.; Zhang, Q.C.; Crisalli, P.; Lee, B.; Jung, J.W.; Kuchelmeister, H.Y.; Batista, P.J.; Torre, E.A.; Kool, E.T., et al. Structural imprints in vivo decode RNA regulatory mechanisms. *Nature* **2015**, *519*, 486-490, doi:10.1038/nature14263.
99. Baek, S.C.; Kim, B.; Jang, H.; Kim, K.; Park, I.S.; Min, D.H.; Kim, V.N. Structural atlas of human primary microRNAs generated by SHAPE-MaP. *Mol Cell* **2024**, *84*, 1158-1172 e1156, doi:10.1016/j.molcel.2024.02.005.
100. Mustoe, A.M.; Busan, S.; Rice, G.M.; Hajdin, C.E.; Peterson, B.K.; Ruda, V.M.; Kubica, N.; Nutiu, R.; Baryza, J.L.; Weeks, K.M. Pervasive Regulatory Functions of mRNA Structure Revealed by High-Resolution SHAPE Probing. *Cell* **2018**, *173*, 181-195 e118, doi:10.1016/j.cell.2018.02.034.
101. Zeller, M.J.; Favorov, O.; Li, K.; Nuthanakanti, A.; Hussein, D.; Michaud, A.; Lafontaine, D.A.; Busan, S.; Serganov, A.; Aube, J., et al. SHAPE-enabled fragment-based ligand discovery for RNA. *Proc Natl Acad Sci U S A* **2022**, *119*, e2122660119, doi:10.1073/pnas.2122660119.
102. Somarowthu, S.; Legiewicz, M.; Chillon, I.; Marcia, M.; Liu, F.; Pyle, A.M. HOTAIR forms an intricate and modular secondary structure. *Mol Cell* **2015**, *58*, 353-361, doi:10.1016/j.molcel.2015.03.006.
103. Bugnon, L.A.; Edera, A.A.; Prochetto, S.; Gerard, M.; Raad, J.; Fenoy, E.; Rubiolo, M.; Chorostecki, U.; Gabaldon, T.; Ariel, F., et al. Secondary structure prediction of long noncoding RNA: review and experimental comparison of existing approaches. *Brief Bioinform* **2022**, *23*, doi:10.1093/bib/bbac205.
104. Novikova, I.V.; Hennelly, S.P.; Sanbonmatsu, K.Y. Structural architecture of the human long non-coding RNA, steroid receptor RNA activator. *Nucleic Acids Res* **2012**, *40*, 5034-5051, doi:10.1093/nar/gks071.
105. Monroy-Eklund, A.; Taylor, C.; Weidmann, C.A.; Burch, C.; Laederach, A. Structural analysis of MALAT1 long noncoding RNA in cells and in evolution. *RNA* **2023**, *29*, 691-704, doi:10.1261/rna.079388.122.
106. Clemson, C.M.; Hutchinson, J.N.; Sara, S.A.; Ensminger, A.W.; Fox, A.H.; Chess, A.; Lawrence, J.B. An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. *Mol Cell* **2009**, *33*, 717-726, doi:10.1016/j.molcel.2009.01.026.
107. Lin, Y.; Schmidt, B.F.; Bruchez, M.P.; McManus, C.J. Structural analyses of NEAT1 lncRNAs suggest long-range RNA interactions that may contribute to paraspeckle architecture. *Nucleic Acids Res* **2018**, *46*, 3742-3752, doi:10.1093/nar/gky046.
108. Frank, F.; Kavousi, N.; Bountali, A.; Dammer, E.B.; Mourtada-Maarabouni, M.; Ortlund, E.A. The lncRNA Growth Arrest Specific 5 Regulates Cell Survival via Distinct Structural Modules with Independent Functions. *Cell Rep* **2020**, *32*, 107933, doi:10.1016/j.celrep.2020.107933.
109. Fang, R.; Moss, W.N.; Rutenberg-Schoenberg, M.; Simon, M.D. Probing Xist RNA Structure in Cells Using Targeted Structure-Seq. *PLoS Genet* **2015**, *11*, e1005668, doi:10.1371/journal.pgen.1005668.
110. Shcherbakova, I.; Brenowitz, M. Perturbation of the hierarchical folding of a large RNA by the destabilization of its Scaffold's tertiary structure. *J Mol Biol* **2005**, *354*, 483-496, doi:10.1016/j.jmb.2005.09.032.
111. Reeder, J.; Hochsmann, M.; Rehmsmeier, M.; Voss, B.; Giegerich, R. Beyond Mfold: recent advances in RNA bioinformatics. *J Biotechnol* **2006**, *124*, 41-55, doi:10.1016/j.jbiotec.2006.01.034.
112. Hofacker, I.L. RNA secondary structure analysis using the Vienna RNA package. *Curr Protoc Bioinformatics* **2009**, *Chapter 12*, 12 12 11-12 12 16, doi:10.1002/0471250953.bi1202s26.
113. Ali, S.E.; Mittal, A.; Mathews, D.H. RNA Secondary Structure Analysis Using RNAstructure. *Curr Protoc* **2023**, *3*, e846, doi:10.1002/cpz1.846.
114. Ding, Y.; Chan, C.Y.; Lawrence, C.E. Sfold web server for statistical folding and rational design of nucleic acids. *Nucleic Acids Res* **2004**, *32*, W135-141, doi:10.1093/nar/gkh449.
115. Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* **2003**, *31*, 3406-3415, doi:10.1093/nar/gkg595.
116. Fernandez, N.; Cordiner, R.A.; Young, R.S.; Hug, N.; Macias, S.; Caceres, J.F. Genetic variation and RNA structure regulate microRNA biogenesis. *Nat Commun* **2017**, *8*, 15114, doi:10.1038/ncomms15114.
117. Long, D.; Lee, R.; Williams, P.; Chan, C.Y.; Ambros, V.; Ding, Y. Potent effect of target structure on microRNA function. *Nat Struct Mol Biol* **2007**, *14*, 287-294, doi:10.1038/nsmb1226.

118. Nawrocki, E.P.; Eddy, S.R. Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics* **2013**, *29*, 2933-2935, doi:10.1093/bioinformatics/btt509.
119. Ontiveros-Palacios, N.; Cooke, E.; Nawrocki, E.P.; Triebel, S.; Marz, M.; Rivas, E.; Griffiths-Jones, S.; Petrov, A.I.; Bateman, A.; Sweeney, B. Rfam 15: RNA families database in 2025. *Nucleic Acids Res* **2024**, *10.1093/nar/gkae1023*, doi:10.1093/nar/gkae1023.
120. Sato, K.; Hamada, M. Recent trends in RNA informatics: a review of machine learning and deep learning for RNA secondary structure prediction and RNA drug discovery. *Brief Bioinform* **2023**, *24*, doi:10.1093/bib/bbad186.
121. Zhao, Q.; Zhao, Z.; Fan, X.; Yuan, Z.; Mao, Q.; Yao, Y. Review of machine learning methods for RNA secondary structure prediction. *PLoS Comput Biol* **2021**, *17*, e1009291, doi:10.1371/journal.pcbi.1009291.
122. Singh, J.; Hanson, J.; Paliwal, K.; Zhou, Y. RNA secondary structure prediction using an ensemble of two-dimensional deep neural networks and transfer learning. *Nat Commun* **2019**, *10*, 5407, doi:10.1038/s41467-019-13395-9.
123. Townshend, R.J.L.; Eismann, S.; Watkins, A.M.; Rangan, R.; Karelina, M.; Das, R.; Dror, R.O. Geometric deep learning of RNA structure. *Science* **2021**, *373*, 1047-1051, doi:10.1126/science.abe5650.
124. Miao, Z.; Westhof, E. RNA Structure: Advances and Assessment of 3D Structure Prediction. *Annu Rev Biophys* **2017**, *46*, 483-503, doi:10.1146/annurev-biophys-070816-034125.
125. Thiel, B.C.; Flamm, C.; Hofacker, I.L. RNA structure prediction: from 2D to 3D. *Emerg Top Life Sci* **2017**, *1*, 275-285, doi:10.1042/ETLS20160027.
126. Li, J.; Chen, S.J. RNA 3D Structure Prediction Using Coarse-Grained Models. *Front Mol Biosci* **2021**, *8*, 720937, doi:10.3389/fmolb.2021.720937.
127. Kryshchovych, A.; Schwede, T.; Topf, M.; Fidelis, K.; Moult, J. Critical assessment of methods of protein structure prediction (CASP)-Round XV. *Proteins* **2023**, *91*, 1539-1549, doi:10.1002/prot.26617.
128. Miao, Z.; Adamiak, R.W.; Antczak, M.; Boniecki, M.J.; Bujnicki, J.; Chen, S.J.; Cheng, C.Y.; Cheng, Y.; Chou, F.C.; Das, R., et al. RNA-Puzzles Round IV: 3D structure predictions of four ribozymes and two aptamers. *RNA* **2020**, *26*, 982-995, doi:10.1261/rna.075341.120.
129. Rother, M.; Milanowska, K.; Puton, T.; Jeleniewicz, J.; Rother, K.; Bujnicki, J.M. ModeRNA server: an online tool for modeling RNA 3D structures. *Bioinformatics* **2011**, *27*, 2441-2442, doi:10.1093/bioinformatics/btr400.
130. Parisien, M.; Major, F. The MC-Fold and MC-Sym pipeline infers RNA structure from sequence data. *Nature* **2008**, *452*, 51-55, doi:10.1038/nature06684.
131. Watkins, A.M.; Das, R. RNA 3D Modeling with FARFAR2, Online. *Methods Mol Biol* **2023**, *2586*, 233-249, doi:10.1007/978-1-0716-2768-6_14.
132. Watkins, A.M.; Rangan, R.; Das, R. FARFAR2: Improved De Novo Rosetta Prediction of Complex Global RNA Folds. *Structure* **2020**, *28*, 963-976 e966, doi:10.1016/j.str.2020.05.011.
133. Biesiada, M.; Purzycka, K.J.; Szachniuk, M.; Blazewicz, J.; Adamiak, R.W. Automated RNA 3D Structure Prediction with RNAComposer. *Methods Mol Biol* **2016**, *1490*, 199-215, doi:10.1007/978-1-4939-6433-8_13.
134. Biesiada, M.; Pachulska-Wieczorek, K.; Adamiak, R.W.; Purzycka, K.J. RNAComposer and RNA 3D structure prediction for nanotechnology. *Methods* **2016**, *103*, 120-127, doi:10.1016/j.ymeth.2016.03.010.
135. Purzycka, K.J.; Popena, M.; Szachniuk, M.; Antczak, M.; Lukasiak, P.; Blazewicz, J.; Adamiak, R.W. Automated 3D RNA structure prediction using the RNAComposer method for riboswitches. *Methods Enzymol* **2015**, *553*, 3-34, doi:10.1016/bs.mie.2014.10.050.
136. Sarzynska, J.; Popena, M.; Antczak, M.; Szachniuk, M. RNA tertiary structure prediction using RNAComposer in CASP15. *Proteins* **2023**, *91*, 1790-1799, doi:10.1002/prot.26578.
137. Palermo, G.; Casalino, L.; Magistrato, A.; Andrew McCammon, J. Understanding the mechanistic basis of non-coding RNA through molecular dynamics simulations. *J Struct Biol* **2019**, *206*, 267-279, doi:10.1016/j.jsb.2019.03.004.
138. Kuhrova, P.; Mlynsky, V.; Zgarbova, M.; Krepl, M.; Bussi, G.; Best, R.B.; Otyepka, M.; Sponer, J.; Banas, P. Improving the Performance of the Amber RNA Force Field by Tuning the Hydrogen-Bonding Interactions. *J Chem Theory Comput* **2019**, *15*, 3288-3305, doi:10.1021/acs.jctc.8b00955.
139. Tan, D.; Piana, S.; Dirks, R.M.; Shaw, D.E. RNA force field with accuracy comparable to state-of-the-art protein force fields. *Proc Natl Acad Sci U S A* **2018**, *115*, E1346-E1355, doi:10.1073/pnas.1713027115.

140. Ebrahimi, P.; Kaur, S.; Baronti, L.; Petzold, K.; Chen, A.A. A two-dimensional replica-exchange molecular dynamics method for simulating RNA folding using sparse experimental restraints. *Methods* **2019**, *162-163*, 96-107, doi:10.1016/j.ymeth.2019.05.001.
141. Bell, D.R.; Cheng, S.Y.; Salazar, H.; Ren, P. Capturing RNA Folding Free Energy with Coarse-Grained Molecular Dynamics Simulations. *Sci Rep* **2017**, *7*, 45812, doi:10.1038/srep45812.
142. Dawson, W.K.; Maciejczyk, M.; Jankowska, E.J.; Bujnicki, J.M. Coarse-grained modeling of RNA 3D structure. *Methods* **2016**, *103*, 138-156, doi:10.1016/j.ymeth.2016.04.026.
143. Paciello, G.; Acquaviva, A.; Ficarra, E.; Deriu, M.A.; Macii, E. A molecular dynamics study of a miRNA:mRNA interaction. *J Mol Model* **2011**, *17*, 2895-2906, doi:10.1007/s00894-011-0991-x.
144. Xia, Z.; Clark, P.; Huynh, T.; Loher, P.; Zhao, Y.; Chen, H.W.; Ren, P.; Rigoutsos, I.; Zhou, R. Molecular dynamics simulations of Ago silencing complexes reveal a large repertoire of admissible 'seed-less' targets. *Sci Rep* **2012**, *2*, 569, doi:10.1038/srep00569.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.