HHS Public Access

Author manuscript

Q Rev Biophys. Author manuscript; available in PMC 2017 June 24.

Published in final edited form as:

Q Rev Biophys. 2016 January ; 49: e10. doi:10.1017/S003358351600007X.

Bridging the Gap Between *In Vitro* and *In Vivo* RNA Folding

Kathleen A. Leamy[1,2], **Sarah M. Assmann**[2,3,4], **David H. Mathews**[5], and **Philip C. Bevilacqua**[1,2,4,6,*]

[1]Department of Chemistry, Pennsylvania State University, University Park, PA, 16802, U.S.A

[2]Center for RNA Molecular Biology, Pennsylvania State University, University Park, PA, 16802, U.S.A

[3]Department of Biology, Pennsylvania State University, University Park, PA, 16802, U.S.A

[4]Plant Biology Graduate Program, Pennsylvania State University, University Park, PA, 16802, U.S.A

[5]Department of Biochemistry and Biophysics, Department of Biostatistics and Computational Biology, and Center for RNA Biology, University of Rochester Medical Center, Rochester, NY 14642, U.S.A

[6]Department of Biochemistry and Molecular Biology, Pennsylvania State University, University Park, PA, 16802, U.S.A

Abstract

Deciphering the folding pathways and predicting the structures of complex three-dimensional biomolecules is central to elucidating biological function. RNA is single-stranded, which gives it the freedom to fold into complex secondary and tertiary structures. These structures endow RNA with the ability to perform complex chemistries and functions ranging from enzymatic activity to gene regulation. Given that RNA is involved in many essential cellular processes, it is critical to understand how it folds and functions *in vivo*. Within the last few years, methods have been developed to probe RNA structures *in vivo* and genome-wide. These studies reveal that RNA often adopts very different structures *in vivo* and *in vitro*, and provide profound insights into RNA biology. Nonetheless, both *in vitro* and *in vivo* approaches have limitations: studies in the complex and uncontrolled cellular environment make it difficult to obtain insight into RNA folding pathways and thermodynamics, and studies *in vitro* often lack direct cellular relevance, leaving a gap in our knowledge of RNA folding *in vivo*. This gap is being bridged by biophysical and mechanistic studies of RNA structure and function under conditions that mimic the cellular environment. To date, most artificial cytoplasms have used various polymers as molecular crowding agents and a series of small molecules as cosolutes. Studies under such *in vivo-like* conditions are yielding fresh insights, such as cooperative folding of functional RNAs and increased activity of ribozymes. These observations are accounted for in part by molecular crowding effects and interactions with other molecules. In this review, we report milestones in

*Author for correspondence: Philip C. Bevilacqua, Department of Chemistry, Pennsylvania State University, University Park, PA 16802, USA and Center for RNA Molecular Biology, Pennsylvania State University, University Park, PA 16802, USA. Tel.: 1-814-863-3812; Fax: 1-814-865-2927; pcb5@psu.edu.

RNA folding *in vitro* and *in vivo* and discuss ongoing experimental and computational efforts to bridge the gap between these two conditions in order to understand how RNA folds in the cell.

1. Introduction

According to the classical view of biology, RNA has three roles, as a messenger (mRNA) that shuttles information between DNA and proteins, as an adaptor (tRNA) that translates the information stored in mRNA into protein sequence, and as a structural molecule (rRNA) that is part of the ribosome (Figure 1). Research over the last twenty-five years has revealed that RNA carries out many other essential functions in the cell. RNA regulates gene expression at the transcriptional and translational levels, and this regulation often arises from the structures adopted by various RNA classes, including ribozymes, riboswitches, and RNA-protein complexes (Doudna & Cech, 2002; Serganov & Nudler, 2013). Since RNA is single stranded it can fold back on itself forming a plethora of secondary and tertiary interactions, as well as complex folding motifs, binding pockets, and active site clefts (Figure 1). Misfolding and mutations of RNA are characteristics of many cancers and diseases; for example, triplet repeat expansion diseases are associated with Huntington's disease, myotonic dystrophy, and Fragile X syndrome (Osborne & Thornton, 2006). Single Nucleotide Polymorphisms (SNPs) that alter the structural ensemble of RNA sequences also have been associated with genetic diseases (Halvorsen et al., 2010). Accordingly, understanding RNA structures and their dynamic regulation is an integral aspect of understanding RNA function.

The negatively charged phosphate backbone and diverse folds of RNA lead it to interact with cellular components, including metal ions, ligands, and proteins. Binding interactions with these species can change the fold of the RNA (Figure 2). Monovalent and divalent metal ions are essential for the catalysis of small self-cleaving and large ribozymes both for folding and for active site catalysis (Serganov & Patel, 2007; Swisher et al., 2002). Small molecule binding refolds riboswitches to regulate gene expression in a positive or negative mode (see Figure 2) (Garst et al., 2011; Serganov & Patel, 2007).

Functional RNAs, such as tRNA, ribozymes, and riboswitches, are often found in ribonucleoprotein (RNP) complexes, which can help fold metastable RNA structures or induce a conformational change. The Protein Data Bank (PDB) has over 2,000 annotated RNA-binding proteins, which include RNA chaperones, helicases, dsRNA binding proteins, tRNA synthetases, ribonucleases (RNases) and RNA recognition motifs (RRMs) (Gerstberger et al., 2014). Highly studied RNPs include the ribosome, non-plant RNase P, and the splicesome, which are responsible for the synthesis of proteins, maturation of the 5′-end of tRNAs, and splicing of pre-mRNAs, respectively. Remarkably it is the RNA component that is responsible for catalysis in these three RNPs, while the protein component provides scaffolding (Guerrier-Takada et al., 1983; Nissen et al., 2000).

Crowding plays critical but poorly understood roles in RNA folding. The cellular environment is very complex with up to 40% of the cytosol taken up by macromolecules (Minton, 2001; Zimmerman & Trach, 1991). In addition, small molecule metabolites, polyamines and other species occupy volume and interact with RNAs. Macromolecular

crowding can drive the compaction of RNA and proteins, while small molecules can either stabilize or destabilize RNAs through interactions with the RNA molecule (Minton, 2001).

Nearly all of the biological components that influence RNA structure and function *in vivo*—biological ion compositions, ligands, proteins, and crowding—are missing during typical *in vitro* experiments (expanded upon in Table 1). A major goal of current research is to add back these components in order to more closely mimic *in vivo* conditions. We group studies of RNA folding into three approaches: (1) *in vitro* studies in dilute solutions; (2) *in vivo* studies in living cells; and (3) *in vivo-like* studies that mimic *in vivo* conditions. We also discuss how *in silico* methods facilitate each of these approaches. There are advantages and limitations to working in each of these conditions, and experiments in each can yield unique insights into the biological functions of RNAs. Structures and folding pathways of RNA have been studied mostly in dilute *in vitro* conditions, resulting in fundamental insights into RNA structure and function. However there is a deep desire to understand how Nature works, and the *in vivo* environment is very different from typical *in vitro* solution conditions (Table 1 and Figure 3). In particular, the majority of thermodynamic experiments studying the energetics of folding and associated pathways (Freier et al., 1986b; Schroeder & Turner, 2009) have been conducted in non-biological salt concentrations (London, 1991; Lusk et al., 1968; Minton, 2001; Romani, 2007; Truong et al., 2013). There are also myriad RNA-protein interactions *in vivo*, many of which profoundly affect RNA folding and function.

Over the last few years, *in vivo* experiments probing RNA structure in living cells have revealed significant differences in many RNA structures as compared to *in vitro* (Kwok et al., 2013; Rouskin et al., 2014; Tyrrell et al., 2013). *In vivo* studies, while desirable because of their biological relevance, are at the same time limited in that they typically elucidate only the ensemble structure of each RNA transcript, do not deconvolute RNA-protein interactions versus RNA self-structure, and cannot easily perturb or control solution conditions. In particular, biophysical studies that can be readily conducted under highly controlled *in vitro* conditions are often simply not feasible *in vivo*. In an effort to gain more insight into the structure and function of RNA in the cellular environment, recent studies have focused on the folding pathway, structure, and function of RNAs under *in vivo-like* conditions, which mimic conditions in the cell (Desai et al., 2014; Dupuis et al., 2014; Kilburn et al., 2013; Nakano et al., 2014; Strulson et al., 2012; Strulson et al., 2013).

In silico prediction and modeling of RNA structure is an important tool used in all three of the above approaches to provide additional insight into RNA structure and function. (Dawson & Bujnicki, 2016; Seetin & Mathews, 2012a). Prediction of canonical base pairs, for example, provides testable hypotheses for RNA structure and also provides frameworks for interpreting experimental results. Likewise, experimental data aid in improving *in silico* structure prediction.

In this review we discuss major achievements in describing and understanding RNA folding and structure through *in vitro*, *in vivo* and *in silico* efforts. The next section introduces the reader to *in vitro* studies of RNA folding, which set the stage for *in vivo* and *in vivo-like* studies of RNA folding. We focus on recent efforts to understand how RNA folds in the cell by bridging the gap between knowledge of RNA structure and folding *in vitro* and *in vivo*,

which has led to an emerging field that studies RNA under *in vivo-like* conditions. We also discuss ways in which the accuracy of *in silico* modeling could be improved with experimentally derived *in vivo* structure probing data. We conclude by discussing advances needed under cellular-like conditions to better understand how RNA folds in the cell.

2. Setting the Stage

2.1. *In vitro* studies of RNA folding

Most of what we currently know about RNA structure and folding comes from studies completed *in vitro*, under experimental conditions that favor a folded state. Such studies are typically conducted in dilute solutions with high concentrations (~1 M) of monovalent ions (Freier et al., 1986b) and/or (~10 mM) divalent ions (Herschlag & Cech, 1990), especially Mg^{2+}, or under conditions that facilitate population of a desired folding intermediate, for example by renaturing the RNA at an unusual temperature or salt concentration (Baird et al., 2005). These solution conditions are advantageous for studying folding because they can be chosen such that the RNA folds in an apparent two-state manner or the RNA populates just a single intermediate, but have the drawback that they differ profoundly from *in vivo* conditions, which have predominantly ~140 mM K^+ and 0.5–3 mM Mg^{2+} (expanded upon in Table 1).

An advantage of using high concentrations of monovalent salts is that they compete with trace polyvalent metal ions and hydroxide ions for the phosphate backbone thereby reducing RNA degradation. In addition, high monovalent salt conditions minimize end fraying of RNA hairpins, favoring two-state folding (Freier et al., 1986b). As we describe below, the thermodynamics and kinetics of systems, ranging from simple RNAs, such as hairpins and bulges, to complex RNAs and RNPs, such as ribozymes and the ribosome, have been well characterized. Many aspects of the RNA folding process can be understood by the application of techniques and the systematic manipulation of conditions only possible under *in vitro* or under *in vivo-like* environments.

2.1.1 Major advances: Elucidating RNA folding pathways *in vitro*—With the invention of various enzymological methods, such as PCR, cloning, T7 transcription and chemical synthesis, RNA preparation has advanced to the point where RNA of almost any sequence and length can be studied (Hoseini & Sauer, 2015; Li et al., 2011; Milligan et al., 1987; Mullis, 1990). A wide variety of techniques have been applied to the study of RNA *in vitro* (Table 2). The earliest studies on RNA were conducted on homoribopolymers, such as polyU and polyA, which revealed that stacking—the nonbonded interactions between the surfaces of the bases—contributes to RNA stability (Richards et al., 1963; Suurkuusk et al., 1977). These studies also provided the first indications that individual RNAs adopt structure. An early breakthrough was from studies of tRNA, which could be isolated from living systems owing to its high cellular abundance, which led to insights into RNA tertiary structure. The cloverleaf base pairing of tRNA had been first predicted from sequence alignments of sequence variants (Levitt, 1969). Solving the crystal structure of tRNA confirmed its cloverleaf secondary structure and revealed novel tertiary interactions (Kim et al., 1973; Robertus et al., 1974). The crystal structure of tRNA provided the first direct

evidence that RNAs can form complex structures, akin to those of proteins, and that stacking, base pairing, and tertiary contacts all contribute to the adoption of complex three dimensional structures (Sussman et al., 1978). With the advent of chemical synthesis techniques, ~100–200mer of DNA and eventually ~50mer RNA of any sequence could be made (Matteucci & Caruthers, 1981; Scaringe et al., 1998; Sierzchala et al., 2003), with a plethora of atomic modifications. Semi-synthetic approaches were then developed that combine enzymological and chemical synthesis to facilitate the introduction of mutations both at the nucleotide and functional group levels in RNAs of any size (Moore & Sharp, 1992).

Thermodynamic and kinetic studies under *in vitro* conditions provide insight into the complex folding pathways of many functional RNAs. Ribozymes and riboswitches are ideal for the study of RNA folding because their function serves as a readout for the occupancy of the native state (Banerjee et al., 1993; Crothers et al., 1974; Mitchell et al., 2013; Mitchell & Russell, 2014; Rook et al., 1998). Major themes are that large RNAs fold on a rugged pathway through populated intermediates, largely in a hierarchical manner, where secondary structures form before tertiary contacts, as demanded by the topologies of these complex RNAs (Figure 4) (Brion & Westhof, 1997; Mitchell & Russell, 2014; Solomatin et al., 2010; Tinoco & Bustamante, 1999; Wan et al., 2010). It is informative to consider these principles on several specific RNAs. Using temperature-dependent NMR and relaxation kinetics, the mechanism of tRNA unfolding was elucidated (Crothers et al., 1974; Hilbers et al., 1976; Stein & Crothers, 1976). Five distinct transitions were mapped to the four arms and the tertiary contacts (Crothers et al., 1974). Secondary structures form on a fast time scale (μs to ms) followed by folding of the tertiary structure on a slower time scale (ms to sec). In the presence of monovalent metal ions, multiple thermal unfolding transitions are observed for these processes (Stein & Crothers, 1976). These transitions merge into one as Mg^{2+} concentrations are increased, revealing that Mg^{2+} induces an apparent two-state folding. Larger functional RNAs, ribozymes, and riboswitches also fold in a hierarchical manner *in vitro* (Figure 4A).

The *Azoarcus* group I ribozyme was used to determine the influence of tertiary interactions on RNA folding (Fig. 5). This ribozyme has been shown to fold quickly, with ~80% of the ribozyme folded into the native state in under 50 ms in 15 mM Mg^{2+} (Rangan et al., 2003). To determine the roles of tertiary interactions in ribozyme folding, the tertiary contact between the P9 GAAA tetraloop and its J5/5a receptor were perturbed (Chauhan & Woodson, 2008). While the WT ribozyme folded in a cooperative manner to the native state, the tetraloop mutant occupied many previously hidden intermediates on the folding pathway, even at 50 mM Mg^{2+}. This study indicated that tertiary contacts promote cooperative RNA folding.

More recently, methods have been developed to study RNA folding on the nucleotide level and at the millisecond time scale (Merino et al., 2005; Scalvi et al., 1997; Zhuang et al., 2000). Experiments using hydroxyl radical mapping yielded insight into the pathway of tertiary structure formation and folding kinetics in the *Tetrahymena* Group I Intron (Sclavi et al., 1998). Combined with time resolved SAXS (Roh et al., 2010), hydroxyl radical footprinting on the *Tetrahymena* ribozyme folding pathway uncovered an initial collapse of

structure on the millisecond timescale during the dead time of the instrument. During the subsequent time course, tertiary contacts and several intermediates were elucidated (Sclavi et al., 1998).

The folding pathways of large functional RNAs have proven to be quite complex with intermediates that can be trapped for minutes to hours (Banerjee & Turner, 1995; Chadalavada et al., 2002; Zarrinkar et al., 1996). For example, 90% of the *Tetrahymena* ribozyme is found in a misfolded state that transitions to the native state with hour timescale kinetics (Banerjee & Turner, 1995), and the HDV ribozyme folds through numerous intermediates, some long-lived (Chadalavada et al., 2002). Long-lived misfolded intermediates are often very similar in structure to the native RNA and typically arise from a secondary structure mispairing or an incorrect three-dimensional topology (Mitchell et al., 2013; Treiber et al., 1998; Wan et al., 2010). For instance, a long-lived intermediate occurs in the *Tetrahymena* ribozyme where P3 is docked correctly but the topology of the ribozyme is incorrect (Mitchell et al., 2013; Mitchell & Russell, 2014). To fold into the native state, this misfold needs to undergo a global unwinding of structure. Importantly, the extent to which these pathways and intermediates are populated *in vivo* is unknown. Indeed, some of these folding intermediates are affected by the method by which the RNA is purified. For example, the wild-type HDV ribozyme has the optimal rate of catalysis when the ribozyme is folded co-transcriptionally, as opposed to being renatured prior to assay (Chadalavada et al., 2007). In addition, choice of flanking sequences can profoundly affect the activity of small and large ribozymes (Cao & Woodson, 1998; Chadalavada et al., 2000).

2.1.2 Major advances: Applying biophysical techniques to study RNA folding *in vitro*—

Using optical melting, a set of thermodynamic parameters have been established to estimate folding free energies from sequence and structure alone (Andronescu et al., 2014; Lu et al., 2006; Turner & Mathews, 2010; Xia et al., 1998). The nearest neighbor model predicts the free energy and stability of an RNA from each base pair's nearest neighbor, along with initiation, symmetry, and terminal-AU base pair terms. Nearest neighbor terms for certain loops, those regions without canonical base pairs, have also been determined (Mathews et al., 2004). As noted below, these experimental parameters have been incorporated in RNA structure prediction programs that find the lowest free energy structures for an input RNA sequence (Mathews, 2006; Reeder et al., 2006; Seetin & Mathews, 2012a). Parameters to account for complicated tertiary interactions and loops are still being revised today (Liu et al., 2011; Liu et al., 2010b; Lu et al., 2006). The nearest neighbor parameters currently available were measured under highly folding *in vitro* conditions of 1 M NaCl.

Low-resolution methods provide information about the structure of RNA on both the global and nucleotide length scales. Although these techniques do not give atomic resolution, they have significantly faster throughput than crystallography or NMR structures while still providing insight into the fold and function of RNA. Small Angle X-ray Scattering (SAXS) and Forster Resonance Energy Transfer (FRET) provide low resolution information on the overall fold of an RNA. RNA is particularly amenable to SAXS because the phosphate backbone is electron-rich and scatters X-rays well. Different solution conditions can be prepared and examined quickly by SAXS to elucidate RNA structural changes. The

structures of several functional RNAs and RNA-protein complexes have been explored using SAXS, including ribozymes, riboswitches bound and unbound to ligand, and the spliceosome (Pollack, 2011). FRET studies, in which acceptor and donor fluorophores are attached to the RNA at key locations, have helped elucidate folding intermediates (Walter, 2001). Using single molecule FRET, or smFRET, the *Tetrahymena* ribozyme was found to fold into multiple conformations, nearly all of which were active, indicating that the ribozyme populates multiple native states (Solomatin et al., 2010). Upon exposure to denaturant, the ribozyme re-populated the native conformations, indicating the results are independent of original conformation. Both SAXS and smFRET have been applied to RNA folding under *in vivo-like* conditions, as discussed below (Paudel & Rueda, 2014; Strulson et al., 2013).

Structure probing methods serve essential roles in elucidating the structures of functional RNAs at the nucleotide level. Several chemical probes have been employed to attack and modify the RNA bases, sugar, and backbone, in order to reveal the base pairing status of the nucleotides. Commonly used chemical probes include dimethyl sulfate (DMS), carbodiimide tosylate (CMCT), and SHAPE reagents, which allow selective 2′-hydroxyl acylation—each of which is analyzed by primer extension via reverse transcription. Commonly used enzymatic probes are RNases T1, V1 and S1. Targets of these probes and methods of readout are provided in Figure 6. Structure probing of RNAs *in vitro* has revealed very complex structures, as well as binding sites of ligands, metal ions, and proteins. As discussed below, structure probing with chemical probes can be used *in vivo* as well.

Very recently, RNA structure *in vitro* has been probed genome-wide at the nucleotide level, utilizing the power of next-generation sequencing. Several methods have been developed to map entire transcriptomes. Parallel Analysis of RNA Structure (PARS) cleaves double-stranded regions with RNase V1 and single-stranded regions with RNase S1, and FragSeq cleaves single-stranded regions with nuclease P1 (Kertesz et al., 2010; Underwood et al., 2010). In PARS, RNA is extracted from cells and aliquots are separately exposed to each nuclease, the digested RNA is converted to cDNA through reverse transcription, and then deep sequenced to map the reverse transcriptase stops to the genome. A PARS score is determined from the log ratio of V1/S1 sequencing reads, where a high PARS score indicates more RNA structure (Kertesz et al., 2010). In FragSeq, RNA is extracted from cells, and one aliquot is treated with P1 nuclease and a second aliquot is untreated (Underwood et al., 2010). RNA-seq is then performed on each aliquot, and a cutting score is determined for each mapped nucleotide that indicates the propensity to be cut by P1 nuclease. The cutting score is then used to annotate RNA secondary structures and/or to restrain RNA secondary structure prediction. Genome-wide studies in several organisms, both *in vitro* and *in vivo*, have found that there is significantly more structure in the coding regions than the untranslated regions of RNAs (Ding et al., 2014; Kertesz et al., 2010; Li et al., 2012; Wan et al., 2014; Zheng et al., 2010). There is also less structure in the start and stop codons than in the rest of a transcript, which presumably facilitates read-through by the ribosome.

Using a method similar to PARS but differing in that the RNA structure is probed at several temperatures, PARTE (Parallel Analysis of RNA Structures with Temperature Elevation)

was used to obtain the folding free energies for yeast transcripts genome-wide *in vitro* (Wan et al., 2012). RNA from yeast was folded between 30°C and 75°C and exposed to RNase V1 followed by deep sequencing. By examining the melting temperatures (Tm) of RNAs, noncoding and coding RNAs could be distinguished and RNAs with distinct cellular functions could be identified. Functional ncRNAs were found to have a higher Tm on average than mRNAs.

Three methods that utilize DMS chemistry to determine transcriptome-wide RNA structure were recently published: Structure-seq, DMS-seq, and Mod-seq (Ding et al., 2014; Rouskin et al., 2014; Talkish et al., 2014). To date, only DMS-seq has been applied *in vitro* and all of the methods have been applied *in vivo*. These methods are described in more detail below.

2.1.3 Benefits and limitations of *in vitro* studies—Many of the foundational experiments on RNA folding and structure have come from *in vitro* experiments, and numerous underlying mechanisms of RNA folding and function have been discovered *in vitro*. Studies *in vitro* have revealed the folding pathways and structures of RNAs. More recently, methods have been developed to probe the structure of RNAs genome-wide. Major advances include elucidating fast formation of secondary structure and slow formation of the tertiary contacts, understanding of RNA folding energetics, establishment of nearest-neighbor parameters, and determination of structures of functional RNA motifs. The complex structures that RNA adopts enable diverse functions. Experimental techniques, ranging from structure probing to kinetic methods, have been applied to RNA across diverse pH, salt, and temperature conditions.

The major limitation of *in vitro* experiments is that the solution conditions are very different from the cellular environment and unavoidably lack many of the components present in cells, which can influence RNA folding and function. These limitations necessitate the development of experiments and techniques under *in vivo* and *in vivo-like* conditions to determine how RNAs fold and respond to cellular environmental conditions.

2.2 *In vivo* studies of RNA folding

In the previous section, we provided an overview of RNA folding *in vitro.* In this section we discuss recent advances made *in vivo* to understand RNA folding. We note that RNA structure has also been explored to a lesser extent in cellular extracts. Experiments in extracts contain more proteins bound to RNA than *in vitro* experiments but less than *in vivo* studies, as supported by recent comparisons of low DMS reactivity assignments amongst *in vitro*, extract, and *in vivo* studies (Ding et al., 2015). Studies in extracts for RNAs with high PPV between reactivities *in vitro* and *in silico*, such as the ribosome, have been shown to be biologically relevant (Ding et al., 2015; Moazed et al., 1986a). Likewise, for RNAs with low PPV between reactivities *in vitro* and *in silico*, studies in extracts might not provide the full complement of interactions. While experiments in cell extracts share many similarities with *in vivo* conditions, thermodynamic assays cannot be easily performed in extracts due to the denaturation and signal of other biomolecules.

An ultimate goal of RNA folding studies is to understand how RNA behaves in the cell. The majority of the methods developed to study RNA *in vivo* are structure probing, where

several chemicals known to penetrate the cell membrane are applied to modify RNA. Structure probing has been used to study the structures of RNAs *in vivo* on both the single gene and genome-wide levels, and has resulted in a breadth of information regarding structures that RNA forms inside living cells. These studies have revealed novel *in vivo* RNA folds, RNA-protein interactions, and novel regulatory roles.

2.2.1 Major advances: Transcript-specific RNA structure mapping *in vivo*— Structure probing of RNA *in vivo* uses small chemicals such as SHAPE reagents, DMS, and CMCT, which penetrate cells and modify solvent-accessible regions of the RNA (Bloomfield et al., 2000; Ehresmann et al., 1987). Structure probing methods using chemicals have revealed that for some transcripts there are significant differences between RNA structures formed *in vivo* and *in vitro*. We first describe *in vivo* structure probing experiments on single transcripts, followed by experiments across a genome.

The first *in vivo* nucleic acid structure probing study was from the Gilbert lab, where binding of multiple proteins to their cognate sites was observed using DMS modification (Nick & Gilbert, 1985). Structure probing is outlined in Figure 6. Briefly, DMS methylates adenine and cytosine on the Watson-Crick face and guanine on the Hoogsteen face. The modification on A and C is read out directly by stops in reverse transcription (RT) one position before the methylated base, while the methylated G is treated with aniline to create an abasic site followed by RT read out, which again stops one position before the modified base (Bloomfield et al., 2000; Ehresmann et al., 1987). The RT can be read out in a gene-specific fashion by polyacrylamide gel electrophoresis (PAGE) or capillary electrophoresis (CE), and in a library fashion with next-generation sequencing (see next section) (Kwok et al., 2013).

The first report of RNA structure comparisons *in vivo* and *in vitro* came from the Cech lab (Zaug & Cech, 1995). Structure probing with DMS was used to map the structures of two known protein-bound RNAs, telomerase RNA and U2 snRNA, as well as the *Tetrahymena* ribozyme. Protections from reactivity *in vivo* compared to *in vitro* indicate either protein protection or gain of base pairing, while enhancements of reactivity indicate refolding to expose RNA bases. Telomerase RNA and U2 snRNA showed different reactivity patterns *in vivo* versus *in vitro*, consistent with the influence of protein binding on DMS reactivity. As expected, the group I ribozyme had very similar nucleotide reactivity *in vivo* and *in vitro*, demonstrating that the ribozyme is not protein-bound and self-splices without protein assistance *in vivo*.

Our group investigated structures of high and low abundance RNAs, also on a gene-specific basis, and compared DMS and SHAPE reactivities *in vivo* and *in vitro*. For low abundance RNAs we developed a gene-specific ligation-mediated PCR (LM-PCR) approach (Kwok et al., 2013). These studies, which were in the model plant species *Arabidopsis thaliana*, revealed *in vivo* footprinting on high abundance 25S rRNA and 5.8S rRNA, as well as on the low abundance U12 snRNA. We showed that different bases in 5.8S rRNA are methylated *in vivo* and *in vitro*, which provided evidence for 5.8S rRNA refolding *in vivo*. These studies also provided critical control reactions that strongly supported DMS modification of RNA

occurring *in vivo* and DMS being completely quenched prior to workup of the *in vivo* reaction. These controls apply equally to the genome-wide studies in the next section.

2.2.2 Major advances: Genome-wide RNA structure mapping *in vivo*—Recently, several groups including ours have developed high-throughput methods to probe RNA structure in living cells transcriptome-wide. These studies revealed significant differences in RNA structure *in vivo* compared to *in vitro* and *in silico* predicted (Ding et al., 2014; Kwok et al., 2015; Rouskin et al., 2014; Talkish et al., 2014). Three separate methods using DMS to probe RNA structure *in vivo* were published in 2014: Structure-seq (Ding et al., 2014), DMS-sequencing (DMS-seq) (Rouskin et al., 2014), and modification sequencing (Mod-seq) (Talkish et al., 2014), each of which utilizes next-generation sequencing to probe RNA structure transcriptome-wide.

Each of these studies revealed novel information on RNA structure and possible regulatory functions of those structures. In Structure-seq, the PPV (positive predictive value) describes the fraction of base pairs in the *in vivo* DMS-restrained predicted structure versus the unrestrained *in silico* predicted structure (Ding et al., 2014). Of the greater than 10,000 mRNAs evaluated in this fashion, most had a PPV value far from unity, with a maximum PPV of the distribution slightly less than 0.4. This observation indicates that the *in vivo* structures of many RNAs cannot be predicted well purely *in silico*, using only sequence information and thermodynamic parameters originally derived *in vitro*. We also observed that the mRNAs with the lowest PPV distribution (bottom 5%) were enriched in annotations of biological function of stress and stimulus response, while the mRNAs with the highest PPV distribution (top 5%) were enriched in housekeeping functions (Ding et al., 2014).

One possibility is that housekeeping RNAs have well-defined folds while stress-related RNAs have ill-defined folds or adopt many folds. DMS-seq in yeast found that certain mRNAs are less structured *in vivo* than naked, protein-free RNA, *in vitro*, and under *in vivo* ATP depletion the mRNAs on a whole become more structured, with the implication that ATP-dependent processes contribute to RNA unfolding. It is likely that a range of factors *in vivo* contribute to RNA structure. (Rouskin et al., 2014). Mod-seq was used to reveal the binding location of the L26 protein by deletion in yeast; upon L26 deletion, 58 nucleotides became more reactive to DMS *in vivo* and most of these nucleotides were located in the 5.8S–25S rRNA interface where L26 is known to bind (Talkish et al., 2014).

Individual copies of a given RNA sequence can adopt different conformations owing to the single-stranded nature of RNA. Indeed, this may be the origin of the low PPV value in the stress-related genes (Ding et al., 2014) in that structure probing methods reveal the average of all populated structures at some instant in time. There is experimental evidence that some transcripts appreciably populate multiple structures *in vitro*. Using the PARS method, ~4% of mRNAs had both high RNase V1 and RNase S1 activity, which cleave paired and unpaired RNA, respectively, under *in vitro* conditions (Wan et al., 2014). The high extents of cleavage by both nucleases suggest that populations of those mRNAs adopt multiple conformations simultaneously *in vitro*, and potentially *in vivo*.

Genome-wide studies revealed a triplet periodicity in mRNA nucleotide reactivity in yeast, mouse, and humans *in vitro* (Incarnato et al., 2014; Wan et al., 2014), as well as in *Arabidopsis in vivo* (Ding et al., 2014; Kertesz et al., 2010). The triplet repeat in reactivity is observed in the coding sequence but not in the untranslated regions. At present the mechanism behind the periodicity is not understood. Observation of the repeat *in vitro* suggests that occupancy of ribosomes is not necessary. Additional studies under *in vitro*, *in vivo*, and *in vivo-like* conditions will be necessary to attain a molecular-level understanding of the triplet periodicity in mRNA.

High-throughput sequencing has been coupled with CLIP (crosslinking and immunoprecipitation) to probe RNA-binding protein sites transcriptome wide in HITS-CLIP (high-throughput sequencing of RNA isolated by crosslinking immunoprecipitation) and PAR-CLIP (photoactivatable-ribonucleoside-enhanced crosslinking and immunoprecipitation) (Hafner et al., 2010; Licatalosi et al., 2008; Weyn-Vanhentenryck et al., 2014). Studies using both of these methods on specific proteins have revealed novel sites of protein binding to RNA as well as possible protein regulatory functions (Hafner et al., 2010; Licatalosi et al., 2008). Briefly, in HITS-CLIP, RNA is crosslinked to proteins, the protein of interest is isolated through IP, the RNA is reverse transcribed and amplified through PCR, then high throughput sequencing is performed and reads are mapped to the genome (Licatalosi et al., 2008). In PAR-CLIP, cells are grown with a photoactivatable nucleoside (4-thiouridine or 5-bromouridine) in the media to facilitate crosslinking with proteins upon exposure to 365 nm radiation (Hafner et al., 2010).

Genome-wide structure data have recently been used to identify certain sites of RNA-protein interactions. The method icSHAPE was used to probe RNA structure in mouse embryonic stem cells *in vivo* and in vitro (Spitale et al., 2015). The difference in nucleotide reactivity *in vitro* and *in vivo* matched binding sites of the protein Rbfox2, previously identified with iCLIP experiments. This methodology was tested again and successfully identified RNA binding sites of another RNA binding protein, HuR. Using this type of analysis, certain RNA-protein interactions and associated RNA structural rearrangements can be distinguished using bioinformatics with experimental genome-wide mapping data.

2.2.3 Major advances: Quantification of cellular factors *in vivo*—*In vivo*

quantification of all the cellular factors known to affect RNA folding would both allow more accurate interpretation of *in vivo* RNA structure datasets and allow design of *in vivo-like* experiments that would more faithfully mimic *in vivo* conditions. Although such a comprehensive view of the inner workings of living cells has yet to be achieved, tremendous strides have been made in technique development for *in vivo* monitoring of cellular parameters relevant to RNA structure, including divalent ion concentrations, pH, ROS (reactive oxygen species), certain cosolutes, and RNA molecules themselves. Almost all of these techniques *in vivo* rely on a fluorescent readout, and thus advances in probe technology have gone hand-in-hand with advances in microscopy, although only the former topic is discussed here.

Fluorescent reporters are of three types: synthetic dyes, genetically encoded reporters, and reporters that incorporate both synthetic dyes and genetically encoded elements. Genetically

encoded reporters typically rely on the cellular factor interacting with and altering the readout from a naturally fluorescent protein from jellyfish, green fluorescent protein (GFP), or its engineered variants (Tsien, 2010), the gene for which can be transformed into the system of interest. The ideal sensor will be minimally invasive and will have high specificity, brightness, and signal to noise ratios, a dynamic range that can accurately report the range of concentrations observed *in vivo*, and response kinetics that are as fast as the natural changes in the probed constituent. The best sensors are also ratiometric, which allows signal normalization to take into account such factors as photobleaching and heterogenous dye distribution. It is important to note that the cellular environment differs among various cellular compartments and organelles. For example, the microenvironments of mitochondria (De Michele et al., 2014) and chloroplasts (Stael et al., 2011) (both of which have their own genomes and thus local RNA transcription) are quite different from the microenvironment of the nucleus, and both differ from the cytosolic environment. Ideally, a sensor would also have the capacity to be specifically targeted to an organelle or subcellular location where RNA folding events of interest occur; for example, sensors that are genetically encoded can be fused to sequences that confer organelle-specific targeting (Choi et al., 2012).

Cations of particular relevance to RNA structure are heavy metals, which tend to destabilize and degrade RNA, Mg^{2+}, which tends to promote RNA folding, and H^{+} (pH), which affects RNA catalysis. In addition, K^{+} and Na^{+} promote formation of the special RNA structure, the G-quadruplex. *In vivo* concentrations of Mg^{2+} (London, 1991; Lusk et al., 1968; Romani, 2007; Truong et al., 2013) and K^{+} as well as pH changes are all within the concentration ranges that can affect RNA structure. Among these cations, sensors based on GFP and its variants are available for Mg^{2+} (Lindenburg et al., 2013), Pb^{2+} (Nadarajan et al., 2014), Hg^{2+} (Hu et al., 2013), and H^{+} (Tantama et al., 2011). A number of synthetic pH sensors are also available (Yang et al., 2014). Both genetically encoded and synthetic sensors of ROS are also available (Pouvreau, 2014; Swanson et al., 2011), which could be applied to study how ROS are associated with genetic diseases (Fimognari, 2015) or environmental conditions (Jaspers & Kangasjärvi, 2010) that affect RNA structure *in vivo*.

As discussed in section 3.2.2, synthetic and biological cosolutes typically destabilize RNA structure. In one early report, sucrose, which is the circulating “energy currency” in plants, was reported to destabilize RNAs *in vitro* (Gao et al., 2016; Lambert & Draper, 2007). While the *in vitro* effects occurred at significantly higher concentrations than prevail in the cytosol proper, in microdomains close to the sites of sugar transporters, sucrose and other sugars could perhaps be present at significantly higher concentrations and consequently affect RNA structure locally; moreover, weakly folded RNAs, such as certain mRNAs, may be more susceptible to such cosolutes. In a possibly analogous situation, while resting Ca^{2+} levels in the cell cytosol are 100–200 nM, Ca^{2+} concentrations as high as 100 mM have been reported at the mouths of Ca^{2+} channels (Tang et al., 2015a). Lipid anchoring of recently developed sucrose and glucose sensors (Fehr et al., 2003; Lager et al., 2006) to probe the near membrane microenvironment of sugar transporters could allow evaluation of this hypothesis.

The physical microenvironment and the localization of RNA, both of which can impact RNA structure, vary across cellular regions and organelles. Accordingly, methods that allow

visualization of the spatial location of any specific RNA of interest are also highly desirable (Buxbaum et al., 2015). One of the first technologies developed for RNA visualization was molecular beacons (Santangelo et al., 2006), which are oligonucleotides tagged with a synthetic fluorophore at one end and a synthetic quencher on the other end. Molecular beacons take on a non-fluorescent stem-loop structure in the absence of a complementary RNA due to the close proximity of the quencher and fluorophore, but exhibit fluorescence upon unfolding and hybridization to the target RNA. Various strategies (Santangelo et al., 2006) can be employed to introduce molecular beacons into mammalian cells but they are not genetically encoded. A more widely used strategy for visualization of specific RNAs employs a genetic approach in which an RNA sequence that binds the bacteriophage MS2 protein is inserted into the UTR of the transcript of interest and the organism is also engineered to express GFP-tagged MS2, which then binds to the transcript of interest, marking its location (Buxbaum et al., 2015).

A different type of RNA marker has been developed recently based on the GFP fluorophore. GFP is fluorescent because the folded protein immobilizes the 4-hydroxy-benzylidene-imidazolinone (HBI) fluorophore encoded by a cyclized and subsequently oxidized Ser–Tyr–Gly tripeptide. RNA aptamers have been identified that analogously immobilize and thus induce fluorescence of a related synthetic fluorophore, DFHBI [(Z)-4-(3,5-difluoro-4-hydroxybenzylidene)-1,2-dimethyl-1H-imidazol-5(4H)-one]. The sequence of the RNA aptamer is genetically incorporated into the gene of interest and upon RNA expression and administration of the membrane-permeant fluorophore and its immobilization by the RNA aptamer, fluorescence is observed that marks the location of the target RNA (Paige et al., 2011). The RNA aptamer, dubbed Spinach, as well as the second generation aptamer Spinach2, both require addition of exogenous Mg^{2+} to fold properly; such addition could obviously also affect native RNA structures. The third generation Spinach reporter, Broccoli, eliminates this requirement (You & Jaffrey, 2015).

Spinach aptamers can be further modified to read out concentrations of cellular metabolites by fusion of the Spinach aptamer with other aptamer sequences (identified by artificial selection) that selectively bind small molecules (Paige et al., 2012), or by incorporation of the Spinach aptamer into prokaryotic riboswitches (You et al., 2015). Riboswitch-based reporters have the advantage of having undergone natural selection that confers high affinity and specificity for the metabolite of interest, but are not currently ratiometric. Ratiometric sensors based on FRET between CFP and YFP, variants of GFP, have been engineered for several metabolites, including those with relevance to RNA structure. For example, FRET-based sensing of ATP concentration (Imamura et al., 2009) could be relevant to RNA structure because of the ATP requirement for the activity of RNA helicases (Rouskin et al., 2014). In summary, the future is bright for *in vivo* quantification of a plethora of the metabolites and physical properties that affect RNA structure. Quantification of cellular factors *in vivo* will play an important role in designing artificial cytoplasms to conduct *in vivo-like* studies of RNA folding.

2.2.4 Benefits and limitations of *in vivo* studies—Studies *in vivo* have shown that RNA can adopt different structures *in vivo* and *in vitro*, and have led to fresh insights on how the cellular environment affects RNA folding across a genome. Novel RNA structure

motifs and RNA-protein interactions have been demonstrated through genome-wide *in vivo* experiments. In addition, novel RNA regulatory pathways have been identified by such studies.

Since some RNAs have been shown to fold and function differently under cellular conditions, "*Why not study RNA solely in living cells instead of in dilute solution conditions?*" The reality is that methods for directly studying RNA folding *in vivo* are limited, and most current *in vivo* approaches rely on structure probing methods that do not probe RNA thermodynamics or folding pathways. Experiments done *in vivo* provide information only on the average RNA structure in a cell or organism and lack information on RNA dynamics, the folding process, and the presence of multiple populated structures of the same transcript. These limitations motivate *in vivo-like* studies to understand the influence of cellular conditions on RNA folding. Before moving to the *in vivo-like* section, we consider the important role that *in silico* studies play in both *in vitro* and *in vivo* studies.

2.3 *In silico* studies of RNA folding

Studies *in vitro* and *in vivo* described above yielded insights into RNA folding and structure that were informed by *in silico* structure prediction tools. Structure probing experiments, for example, typically use *in silico* prediction tools to model structure that is guided by the experimental data. In the subsections below, we describe advances in predicting RNA structure from one sequence, from multiple sequences, and with experimental data. Limitations of each approach are provided as well.

2.3.1 Major advances: RNA structure prediction from one sequence *in silico*—

The most popular approaches to predict RNA structure use dynamic programming algorithms to efficiently search the set of possible structures (Eddy, 2004) and folding free energy nearest neighbor rules to estimate folding stability (Turner & Mathews, 2010). The dynamic programming algorithms guarantee that every structure allowed by the set of folding rules is considered, except for those containing pseudoknots (see below). This means, for example, that the lowest free energy conformation will be found for programs that find lowest free energy structures, i.e. the most probable structure at equilibrium.

The accuracy of RNA structure prediction from sequence alone, in terms of fraction of known pairs correctly predicted, is stubbornly limited to ~70% (Hajiaghayi et al., 2012; Lu et al., 2009), and accuracy is lower for long sequences (>1000 nucleobases) such as small and large ribosomal RNAs and mRNAs (Doshi et al., 2004) or for sequences that fold to more than one conformation at equilibrium. *In silico* predictions of base pairs presently rely on a parameterization of stabilities determined *in vitro* rather than *in vivo*, and these parameters are based on relatively few experiments, as compared to all possible folded sequences.

In response to this moderate success rate, a number of *in silico* methods have been developed to predict alternative structures, as reviewed previously (Mathews, 2006). Programs generate sets of alternative hypotheses for the structure (suboptimal structures) (Wuchty et al., 1999; Zuker, 1989), feasible structures in equilibrium with each other (stochastic samples) (Ding & Lawrence, 2003), or estimates for base pairing probabilities

(partition function calculations) (McCaskill, 1990). Each of these three methods is described in turn. Suboptimal structures are those with similar free energy to the lowest free energy structure. Certain suboptimal structures can sometimes be more representative of the biological structure than the *in silico*-estimated lowest free energy structure, and can be viewed as alternative models or alternative hypotheses for the *in vivo* structure. Stochastic samples are rigorous samples from the equilibrium (Boltzmann) ensemble. They are useful for estimating ensemble statistics for the secondary structure of an RNA. Partition function calculations provide pairing probability estimates; more probable pairs in predicted structures are more likely to occur in the accepted structure (Mathews, 2004).

2.3.2 Major advances: RNA structure prediction from multiple sequences *in silico*—The accuracy of *in silico* folding can be dramatically improved by using additional information to guide the folding. In this section, we discuss using homologous sequences to guide the folding, while in the next section we discuss applying experimental data. Multiple homologous sequences, commonly called an RNA family, can be used to estimate the common secondary structure (Seetin & Mathews, 2012a) because structure is generally conserved to a greater extent than sequence for RNAs. Due to sequence variation, the number of base pairs conserved across a family is smaller than the number of base pairs adopted by each sequence. With enough sequences, conserved pairs stand out as positions of covariation, where compensating base pair changes are observed. Covariation is a change in sequences where one biological species, for example, will have an AU base pair, but another species will have a GC pair at the homologous position. During evolution, two separate changes occurred in sequence (a compensating change) that conserved the base pair.

Three approaches are used to estimate the biologically conserved structure from a set of homologous sequences (Reeder et al., 2006; Seetin & Mathews, 2012a). In the first approach, the available sequences are aligned, and then used to restrain the *in silico* prediction. This approach is typically the fastest, but generally works best when the pairwise sequence identity of all the homologs is high (75% or higher). These programs are exemplified by RNAalifold (Bernhart et al., 2008) and TurboFold (Harmanci et al., 2011). Programs in the second set predict the structures for each sequence first and then compare the predicted structures to find those common to all sequences. This approach works well when the structure is highly conserved and is exemplified by RNAcast (Reeder & Giegerich, 2005). The third approach is to simultaneously align and fold sequences to find the common structure and sequence alignment. This is the best approach to use when the sequences are diverse (pairwise sequence identity for some sequence pairs below 75%) because low pairwise identity makes sequence alignment challenging. Programs in this class include Dynalign/Multilign (Fu et al., 2014; Xu & Mathews, 2011), Foldalign (Torarinsson et al., 2007), LocARNA (Will et al., 2007), PARTS (Harmanci et al., 2008), and RAF (Do et al., 2008).

The accuracy of *in silico* prediction of conserved structures from a set of homologous sequences can be much higher, than for predictions from single sequences. For example, often an additional 20% or more of the known base pairs can be correctly predicted using multiple homologs as compared to predictions using a single sequence (Xu & Mathews, 2011). For a given set of sequences, however, it is not always obvious which approach or

program to use, and, therefore, it is probably best to try more than one program to develop hypotheses about the *in vivo* structure. To date, no program can completely automate comparative sequence analysis. Manual comparison is still required for the most accurate RNA secondary structure determination.

2.3.3. Major advances: RNA structure prediction *in silico* restrained with experimental data—Another type of information used to guide *in silico* prediction of RNA structure is experimental structure mapping. Such mapping data can come from *in vitro* or *in vivo* experiments and are used to restrain structure prediction (Lorenz et al. 2016; Sloma & Mathews, 2015). The effects of experimental structure restraints have been well studied using *in vitro* probing data on structured non-coding RNAs. Over 85% of known pairs can be correctly predicted using *in vitro* SHAPE, DMS, or enzymatic cleavage data (Cordero et al., 2012; Deigan et al., 2009; Eddy, 2014; Hajdin et al., 2013; Ouyang et al., 2013; Washietl et al., 2012; Wu et al., 2015; Zarringhalam et al., 2012) when the extent of accessibility is quantified using capillary/gel electrophoresis or deep sequencing counts. This is a dramatic improvement over the above-mentioned 70% limit in the absence of mapping data. Using *in vivo* mapping data to improve the accuracy of structure prediction has not yet been well-studied, although mapping data overlaid on known structures suggests that, for structured non-coding RNAs such as rRNAs, the existing methods should improve structure prediction accuracy (Ding et al., 2014). We recently developed a pipeline called StructureFold to fold RNAs across a genome using restraints from experimental data, which works with Structure-seq data (Tang et al., 2015b) and the RNAstructure program and can accommodate other data and folding algorithms.

2.3.4 Challenges with *in silico* modeling of RNA secondary structure—Despite its widespread use, RNA secondary structure prediction has known limitations. First, the nearest neighbor parameters are based on a limited number of experiments measured *in vitro* in 1 M NaCl rather than *in vivo-like* conditions, and there are probably many sequences that are not predicted well with those parameters (Andronescu et al., 2014; Mathews et al., 2004). On the one hand, for a limited number of simple RNAs melted in physiological K^+ and Mg^{2+} concentrations, the stability is often similar to that in 1 M NaCl (Diamond et al., 2001; Jaeger et al., 1990; Jiang et al., 2014; Schroeder & Turner, 2000). However, for a 5S ribosomal RNA loop E motif, for example, an appreciable difference in stabilities was found between buffers with and without Mg^{2+} (Serra et al., 2002). Second, although enthalpy parameters are available for structure prediction between 10 and 60°C (Lu et al., 2006), predictions are generally made at 37°C, which is relevant to humans, but not the majority of organisms. Third, finding lowest free energy structures assumes that RNAs fold to equilibrium, i.e. kinetics do not control folding. In favor of this assumption, an *in vivo* study of ribozymes suggested that RNAs fold to equilibrium to a greater extent in yeast cells than *in vitro* (Mahen et al., 2005). Also, *in vitro* structure mapping studies of annealed ribosomal RNAs were consistent with *in vivo* structures (Moazed et al., 1986b). However, some sequences are kinetically trapped, such as transcriptional riboswitches (Seetin & Mathews, 2012a; Wickiser et al., 2005; Wickiser et al.). Therefore, it is unclear to what extent factors such as non-physiological ionic conditions and cotransciptional folding play roles in shaping the folding of RNA.

A fourth limitation of the most popular programs for *in silico* folding is that they cannot predict pseudoknots (Liu et al., 2010a). A pseudoknot occurs when nucleotides in two loops form base pairs. Formally, a pseudoknot is composed of two or more base pairs, defined by indices i base paired to j and i′ base paired to j′, where the order of the nucleotides is i < i′ < j < j′. Pseudoknotted pairs are a small fraction of total base pairs in known structures but often occur in highly structured and functional RNAs. For programs that predict pseudoknots, the accuracy is shockingly low (<5%) (Bellaousov & Mathews, 2010), although the use of multiple homologous sequences to identify conserved pseudoknots can improve the accuracy (Seetin & Mathews, 2012b). Recently, it was also shown that *in vitro* SHAPE mapping data can guide *in silico* structure prediction, including pseudoknots, and achieve over 90% accuracy at predicting known base pairs (Hajdin et al., 2013). The program that implements this, ShapeKnots, is limited, however, to sequences of 600 nucleotides or fewer.

Although structure mapping data and sequence comparison are each used to guide *in silico* modeling of RNA secondary structure, little has been done until recently to combine the two approaches for additional synergy. The secondary structures of three long non-coding RNAs were modeled with the aid of structure mapping data: HOTAIR (with *in vitro* SHAPE, DMS, and terbium) (Somarowthu et al., 2015), SRA (with *in vitro* SHAPE, DMS, in-line probing, and RNase V1 digestion) (Novikova et al., 2012), and XIST (with *in vivo* DMS mapping) (Fang et al., 2015). For each of these studies, sequence comparison, i.e. the verification that the structures are conserved and the identification of compensating base pair changes, was subsequently used to further support the structure model.

Two software programs were enhanced to combine structure mapping data and sequence comparison to improve structure prediction. Sükösd et al. reported PPfold, a program that uses a probabilistic approach to predict structure and can be guided by SHAPE mapping data and/or sequence covariation as estimated from a sequence alignment (Sükösd et al., 2012). Recently, SHAPE data was used to inform sequence alignment and then RNAalifold to predict the conserved structure for the aligned sequences (Lavender et al., 2015b). The key observation is that homologous nucleotides, i.e. those that align, have similar SHAPE reactivities and thus the differences in SHAPE reactivity can be included as an additional metric in the scoring of alignments. This approach demonstrated an improved accuracy of base pair prediction by RNAalifold as compared to consensus structure prediction or SHAPE guided structure prediction alone. Both of these approaches were used to model HIV RNA structure using mapping data and sequence comparison (Lavender et al., 2015a; Sükösd et al., 2015).

3. Bridging the gap between *in vitro* and *in vivo* RNA folding using *in vivo-like* studies

3.1 The gap

The previous sections outlined major contributions of RNA folding studies *in vitro* and *in vivo* to our understanding of how RNA behaves, while considering the important roles that *in silico* approaches play. *In vitro* studies provide the fundamentals of RNA thermodynamics

and kinetics, RNA structural motifs, and genome-wide RNA structure trends. *In vivo* structure probing methods reveal RNA structural trends related to biological functions and regulatory roles of RNA genome-wide. We discussed how several research teams have used genome-wide *in vivo* structural probing to uncover that, in general, RNAs do not adopt the same structures *in vivo* as *in vitro*. Since structure generally dictates function, understanding differences between RNA folding *in vivo* and *in vitro* can illuminate biological function. Toward accomplishing this goal, RNA folding and function studies have been increasingly conducted under conditions that *mimic* the cellular environment.

The dilute solution conditions traditionally used to study RNA *in vitro* are vastly different from the cellular environment. The cellular environment is a complex solution containing biopolymers, metabolites, dilute free salts, and organelles, with 20 – 40% of the cellular volume occupied by macromolecular crowders (Minton, 2001; Zimmerman & Trach, 1991). As such, there is no single cellular environment to which RNA is exposed. As an mRNA passes from the nucleus to the cytosol, solution conditions change; in eukaryotes, the cell is compartmentalized and as the RNA is transported to different regions its fold can change.

It is of interest to consider the differences between RNA structure in eukaryotic and prokaryotic organisms. Functional RNAs have intricate structures with tertiary contacts that assemble secondary structures close in space. Cations, typically Mg^{2+}, neutralize the negative charge of the phosphate backbone and promote tertiary structures. Free Mg^{2+} concentrations in prokaryotic and eukaryotic cells are different, ~1.5–3.0 mM and 0.5–1.0 mM, respectively (London, 1991; Lusk et al., 1968; Romani, 2007; Truong et al., 2013). Structured RNAs such as ribozymes, riboswitches, and thermosensors, are found frequently in prokaryotes, where free Mg^{2+} levels are higher. Although a few ribozymes and one riboswitch have been identified in eukaryotes, they appear to be rare, and proteins are typically involved in forming requisite tertiary structures (Kubodera et al., 2003; Roth et al., 2014; Salehi-Ashtiani et al., 2006). Lambowitz and co-workers demonstrated that prokaryotic group II introns fold poorly in eukaryotic cells, although they could select variant RNAs that fold into active conformations at eukaryotic low Mg^{2+} concentrations (Truong et al., 2013). Studies in our lab indicate that the eukaryotic innate immune sensor PKR is activated by prokaryotic RNAs under eukaryotic low Mg^{2+} conditions, leading to the speculation that riboswitches and ribozymes may be selected against in eukaryotes to aid in discriminating self and non-self at the RNA level (Hull & Bevilacqua, 2015, 2016; Hull et al., 2016). To date, there are no studies that compare the structures of eukaryotic and prokaryotic RNAs genome-wide, but such information would be valuable.

Historically, *in vitro* experiments lack many of the components of cellular environments, and, moreover, often have high concentrations of salt to fold RNA for thermodynamic and structural studies (Table 1). Thermodynamic studies cannot, however, readily be performed *in vivo*. The cell prohibits wide variations of temperature, pH, salt, and ligand concentration, all of which are necessary to obtain thermodynamic information. As a result, RNA is being increasingly studied in artificial cytoplasms that mimic aspects of the cellular environment while allowing biophysical studies. Several recent studies focused on mimicking aspects of the *in vivo* environment *in vitro*; conditions referred to herein as '*in vivo-like*' conditions (Figure 3). Effects of such conditions as cellular concentrations of monovalent and divalent

ions and molecular crowding agents on the folding of RNAs have been a theme in a number of recent studies (Desai et al., 2014; Dupuis et al., 2014; Nakano et al., 2015; Paudel & Rueda, 2014; Strulson et al., 2014; Tyrrell et al., 2015). Experiments under these *in vivo-like* conditions have the potential to bridge our understanding of observations made *in vitro* and *in vivo*.

3.2 Design of artificial cytoplasms and early experiments

In this section, we discuss various methods of mimicking cytoplasmic conditions, including the use of polymers and cosolutes as crowding agents and the use of protocells and synthetic membranes. We also discuss the outcomes of early experiments under these *in vivo-like* conditions. Finally, directions in which the field needs to move to understand the fold and function of RNA *in vivo* are suggested.

3.2.1 Polymers—Synthetic crowding agents such as polyethylene glycol, dextran, and ficoll, and small cosolute additives such as methanol, proline, and TMAO have been used to mimic the crowded environment of the living cell. Functional RNAs that are well-studied *in vitro* have been used to test the effects crowding agents have on RNA folding. Various methods, including UV melts, SAXS, kinetic techniques, and smFRET, have been used to study RNA under these *in vivo-like* conditions. Several studies have shown that synthetic crowding agents affect the thermodynamics and function of several RNAs (Dupuis et al., 2014; Kilburn et al., 2013; Kilburn et al., 2010; Lambert et al., 2010; Strulson et al., 2014). Findings of these studies are that RNAs fold cooperatively, structure becomes compact, and ribozymes cleave faster under *in vivo-like* conditions (Kilburn et al., 2013; Nakano et al., 2009; Strulson et al., 2014; Strulson et al., 2013).

The kinetics of several small and large ribozymes have been probed under *in vivo-like* conditions and in all reported cases, rates of catalysis have increased in the presence of molecular crowders as compared to dilute solution conditions (Desai et al., 2014; Nakano et al., 2009; Paudel & Rueda, 2014; Strulson et al., 2012; Strulson et al., 2013). For example, the hammerhead ribozyme has higher catalytic activity, between 3.5–6.5 faster than in dilute solutions, in the presence of 10%–30% (wt %) PEG200 or PEG8000, suggesting a more populated active state in crowded conditions (Nakano et al., 2009). In addition, i*n vivo-like* solution conditions can stabilize ribozymes even in the presence of denaturants. For example, the rate of catalysis of the CPEB3 ribozyme in the presence of 2.5 M of the denaturant urea was recovered by the addition of 30% (w/v) PEG200, PEG8000, or Dextran10, at a rate higher than in buffer alone (Strulson et al., 2013). SAXS experiments have provided insight into the structural basis for enhanced catalysis, showing that the natively folded state adopts a more compact structure in the presence of molecular crowders under conditions of biological Mg^{2+} concentrations (Kilburn et al., 2010; Strulson et al., 2013).

The thermal stability of several functional RNAs has been reported to increase under *in vivo-like* conditions as compared to *in vitro* experiments. For instance, in 20% PEG200 or PEG8000 the hammerhead ribozyme retains catalytic activity up to 60°C, a temperature that thermally denatures the ribozyme in dilute solutions (Nakano et al., 2009). Observation of

increased hammerhead catalytic activity, up to 270-fold, at high temperatures in crowded conditions indicates a more thermostable RNA under *in vivo-like* conditions. Interestingly, the individual secondary structure elements of the ribozyme were observed, through optical melting experiments, to be thermally destabilized in molecular crowding agents, indicating tertiary structure is stabilized and resulting in more cooperative folding of the ribozyme (Nakano et al., 2009). A thermodynamic study from our lab using SHAPE structure probing on $tRNA^{phe}$ under *in vivo-like* conditions showed that tRNA folds in a cooperative manner at biological Mg^{2+} concentrations in the presence of molecular crowding (Strulson et al., 2014). The observed increase in folding cooperativity with crowding was accompanied by an increase in the temperature of the melting transition for tertiary structure. When the tertiary interactions were removed by mutation of nucleotides in tertiary contacts to uridine, cooperativity was lost and the RNA folded with multiple transitions under all solution conditions, thus indicating that tertiary interactions are vital to cooperative RNA folding under *in vivo-like* conditions. This effect is similar to that observed under *in vitro* conditions mentioned above (Chauhan & Woodson, 2008).

The contribution of molecular crowding agents to RNA catalysis and folding has been found to be largest in a background of physiologically low ionic conditions rather than high ionic conditions (Kilburn et al., 2013; Strulson et al., 2013). In the absence of crowding, physiological concentrations of Mg^{2+} are not high enough to fold functional RNAs in a two-state manner. This is apparent from the observation of long-lived intermediates and slow folding under these conditions (Banerjee & Turner, 1995; Chadalavada et al., 2002; Mitchell et al., 2013). However, in the presence of biological crowding conditions and physiological Mg^{2+}, functional RNAs tend to fold in a cooperative manner into compact structures (Desai et al., 2014; Dupuis et al., 2014; Strulson et al., 2014; Tyrrell et al., 2015), and ribozymes and riboswitches tend to have higher rates of cleavage and higher ligand binding affinity (Paudel & Rueda, 2014). The addition of more Mg^{2+} to these conditions does not result in a further increase in the rate of activity or more cooperative RNA folding, indicating that together physiological crowding and Mg^{2+} conditions fold RNA optimally (Figure 7).

A recent study explored the structural effects of the molecular crowding agent PEG (ranging in size from the monomer to 35000 kDa) on the adenine riboswitch (Tyrrell et al., 2015). Using SHAPE chemistry, the reactivity of the riboswitch under *in vitro*, *in vivo*, and *in vivo-like* conditions was explored. The authors found that in low molecular weight PEG (<3350 kDa) the riboswitch had low correlation between reactivity *in vivo* and *in vivo-like* conditions, whereas in higher molecular weight PEG (12000 kDa-35000 kDa) the RNA had a similar reactivity under *in vivo* and *in vivo-like* conditions. While this study was limited to a single molecular crowding agent, it is significant because it showed that certain *in vivo-like* conditions are not accurate cellular mimics.

Recently, the folding of a model RNA was examined in vivo. The *Salmonella* fourU RNA thermometer hairpin containing a FRET pair was injected into live mammalian cells and reported to have similar melting temperatures and unfolding free energy *in vivo* and *in vitro* (Gao et al., 2016). The addition of 30% (w/v) PEG of varying sizes and Ficoll70 was shown to modify the thermodynamics of the hairpin, and higher molecular weight polymers were found to have similar effects on the RNA as the *in vivo* environment. The *in vivo* data had a

very broad distribution of melting temperatures and free energy between both different cells and different cellular compartments, leading to some uncertainty about how the cellular environment is affecting RNA folding.

3.2.2 Cosolutes—While molecular crowding agents generally facilitate the folding of functional RNAs, small cosolutes have varying and complicated effects on RNA thermostability and folding cooperativity. This arises in part because the effect on stability depends strongly on the interactions between the particular cosolute and RNA considered. Cosolutes, also known as osmolytes, regulate osmotic pressure in cells (Record et al., 1998; Yancey et al., 1982). Effects of cosolutes on RNA folding were not significantly investigated until the last decade. Studies on RNAs with either secondary and/or tertiary structures report that cosolutes such as betaine, proline, and methanol, almost always destabilize secondary structures, while having mixed effects on tertiary structure (Lambert & Draper, 2007; Lambert & Draper, 2012; Lambert et al., 2010; Soto et al., 2007). Several osmolytes have been shown to interact with the nucleobase, sugar, and phosphate of RNAs, with examples of both favorable and unfavorable interactions (Lambert & Draper, 2007). Stabilizing osmolytes have unfavorable interactions with the unfolded state of RNA, resulting in RNA compaction that buries functional groups and stabilization of the native state, while destabilizing osmolytes have favorable interactions with the unfolded state of RNA, driving unfolding (Holmstrom et al., 2015; Lambert & Draper, 2007; Lambert et al., 2010).

There are a limited number of studies available for the effect of cosolutes on RNA function. The hammerhead ribozyme was shown to have increased rates of cleavage in 20% cosolutes, such as glycerol and 1,2-dimethoxyethane, in the presence of physiological Mg^{2+}, which was attributed to enhanced electrostatic interactions with Mg^{2+} (Nakano et al., 2015). The secondary and tertiary structures of the hammerhead ribozyme were destabilized in the presence of several cosolutes (Nakano et al., 2009). In crowded conditions, ribozyme activity also increased, while secondary structure was destabilized and tertiary structure was stabilized (see above).

The influence of the cosolute trimethylamine oxide (TMAO) on RNA secondary and tertiary structure, as well as on the phosphate backbone, has been studied (Denning et al., 2013; Lambert & Draper, 2007; Lambert et al., 2010). TMAO is unusual in having almost no effect on secondary structure stability, while generally stabilizing tertiary structure. A small 58mer rRNA was found to exhibit cooperative two-state folding in the presence of TMAO, observed by a single transition in an optical melting experiment (Lambert et al., 2010).

3.2.3 Protocells and synthetic membranes—There are several groups focusing on how to model RNA function and structure in early Earth conditions which also relate to compartmentalization in modern cells. Coacervates and synthetic membranes are often used to mimic early Earth protocells, and RNA function in these protocells is often studied through ribozyme cleavage. Our lab studied the activity of a two-piece hammerhead ribozyme in aqueous two-phase systems (ATPS) made of polyethylene glycol and dextran (Strulson et al., 2012). The system forms a dextran-rich phase droplet in which the ribozyme preferentially localizes at a concentration up to 3,000 times that of the aqueous phase, resulting in a 70-fold increase in the rate of catalysis. This study suggested that RNA

catalysis in the early Earth environment could have arisen from compartmentalization increasing the local concentration of RNA, possibly accelerating very slow reactions so that they could occur on a biologically relevant time scale.

Similar to these droplets, mononucleotides will form microdroplets when mixed with cationic peptides in water (Koga et al., 2011). Inside these droplets, nucleotides and peptides can reach concentrations as high as 1.6 M and 400 mM respectively, which is much more concentrated than in the aqueous phase. Cationic and anionic dyes and certain nanoparticles were shown to partition into the droplets, indicating that the droplets are permeable to charged molecules (Koga et al., 2011). These droplet phases are another indicator that early life could have arisen in non-membranous compartments. More recently we made coacervates from nucleotides and poly(allylamine) that contain molar concentrations of Mg^{2+} and nucleotides, which could facilitate RNA catalysis in an early life scenario (Frankel et al., 2016).

4. Future Directions

The majority of what is known about RNA folding and structure comes from studies that were performed *in vitro* on small model systems and highly structured RNAs. In contrast, little is known about how RNA folds and functions *in vivo*. Current *in vivo* methods probe the RNA structure ensemble. While providing a benchmark for new prediction parameters, ensemble methods cannot themselves generate thermodynamic parameters.

The current thermodynamic parameters for RNA structure prediction were established in 1 M NaCl. However, several-transcript specific and genome-wide studies have shown that certain RNAs do not fold into the same structures *in vivo* and *in vitro* (Kwok et al., 2013; Rouskin et al., 2014; Tyrrell et al., 2013; Tyrrell et al., 2015), so improved software and prediction parameters are needed to model *in vivo* structure. In particular, genome-wide *in vivo* structure probing data sets (Ding et al., 2014; Rouskin et al., 2014) contain a wealth of information that has not yet been completely realized or understood. One barrier to taking full advantage of the data is that most of the available *in silico* methods assume that RNAs fold to a single structure (Cordero et al., 2012; Deigan et al., 2009; Hajdin et al., 2013; Ouyang et al., 2013; Wu et al., 2015), while probing data averages across all structures populated by sequences for the duration of the experiment.

Modeling a single structure works well for non-coding RNA sequences that function with a single structure, such as ribosomal RNAs, but there are many RNAs for which this assumption is not correct, such as RNA switches and open reading frames. A key challenge is developing methods to use the probing data to model ensembles of relevant structures. Three recent papers highlight work to address this challenge. Cordero and Das report an *in silico* method (M^2-REEFFIT) that models complex mixtures of multiple structures, aided by *in vitro* SHAPE mapping of the wild type sequence and also a set of mutant sequences, which reveal nucleotide interactions (Cordero & Das, 2015). Multiple structures for the 5′ UTR of an mRNA were modeled using *in vitro* SHAPE mapping of a mixture of structures (Kutchko et al., 2015). The multiple conformations were modeled *in silico* using stochastic sampling, restrained using the standard SHAPE restraints expressed as free energy terms

(Deigan et al., 2009). A third *in vitro* approach separated multiple conformations of HIV RNA using native gel electrophoresis, and mapped the structures with SHAPE in the gel (Sherpa et al., 2015). This simplified the *in silico* analysis because the SHAPE mapping data were acquired for each conformation independently.

Another key challenge for mapping studies is determining the best way to discover or model interactions of RNAs with proteins or other RNAs. *In vivo*, all RNAs can interact with macromolecules and metabolites. These interactions generally result in protection from probing agents. Deconvoluting *in silico* whether a nucleotide is unreactive because of intramolecular structure or intermolecular interactions is a grand challenge that will likely require new types of experimental information to address. Modeling and predicting three-dimensional RNA structures *in vitro* is an ongoing challenge. A recent RNA puzzle tested blind 3D folding predictions by providing research teams with RNA sequences and chemical probing data for those RNAs (Miao et al., 2015). The structures that the teams modeled were compared with crystal structures, and most teams could predict Watson-Crick base pairs, but struggled in predicting non-canonical WC base pairing and stacking interactions. A long-range of goal is to predict relevant RNA 3D structures *in vivo* to understand the biologically relevant confirmation(s).

The study of RNA under *in vivo-like* conditions is relatively young. To better mimic the cellular environment, more complex cytoplasm mimics should be developed. To date artificial cytoplasms have focused on synthetic polymers and cosolutes, but more accurate ionic conditions, biopolymers and even cell extracts need to be applied. In addition, studies under *in vivo-like* conditions have focused on single-transcripts in synthetic crowding and cosolute conditions. Genome-wide comparisons of RNA folding under *in vivo* and *in vivo-like* conditions are needed. Lastly, methods that can probe the thermodynamics and kinetics of RNA folding under complex *in vivo-like* conditions will enhance our understanding of *in vivo* RNA folding. Overcoming the challenges outlined herein will allow the field to accomplish the ultimate goal, to understand how RNA folds in the cell.

Acknowledgments

The authors would like to thank the NIH for funding under R01-GM110237 and the NSF for funding under IOS-1339282.

References

ALBERTS, B., BRAY, D., LEWIS, J., ROBERTS, K., WATSON, JD. Molecular biology of the cell. 3. Garland Publishing; New York and London: 1994.

ANDRONESCU M, CONDON A, TURNER DH, MATHEWS DH. The Determination of RNA Folding Nearest Neighbor Parameters. Methods in Molecular Biology. 2014; 1097:45–70. [PubMed: 24639154]

BAIRD NJ, WESTHOF E, QIN H, PAN T, SOSNICK TR. Structure of a folding intermediate reveals the interplay between core and peripheral elements in RNA folding. Journal of Molecular Biology. 2005; 352:712–722. [PubMed: 16115647]

BANERJEE AR, JAEGER JA, TURNER DH. Thermal unfolding of a group 1 ribozyme: the low-temperature transition is primarily disruption of the tertiary structure. Biochemistry. 1993; 32(1): 153–163. [PubMed: 8418835]

BANERJEE AR, TURNER DH. The time dependence of chemical modification reveals slow steps in the folding of a Group I ribozyme. Biochemistry. 1995; 34:6504–6512. [PubMed: 7756281]

BELLAOUSOV S, MATHEWS DH. ProbKnot: Fast prediction of RNA secondary structure including pseudoknots. RNA. 2010; 16:1870–1880. [PubMed: 20699301]

BERNHART SH, HOFACKER IL, WILL S, GRUBER AR, STADLER PF. RNAalifold: improved consensus structure prediction for RNA alignments. BMC Bioinformatics. 2008; 9:474. [PubMed: 19014431]

BLOOMFIELD, VA., CROTHERS, DM., TINOCO, IJ. Nucleic Acids: Structures, Properties, and Functions. Sausalito, California: University Science Books; 2000.

BRION P, WESTHOF E. Hierarchy and dynamics of RNA folding. Annual Review of Biophysics and Biomolecular Structure. 1997; 26:113–137.

BUXBAUM AR, HAIMOVICH G, SINGER RH. In the right place at the right time: visualizing and understanding mRNA localization. Nature reviews Molecular cell biology. 2015; 16(2):95–109. [PubMed: 25549890]

CAO Y, WOODSON SA. Destabilizing effect of an rRNA stem-loop on an attenuator hairpin in the 5′ exon of the *Tetrahymena* pre-rRNA. RNA. 1998; 4:901–914. [PubMed: 9701282]

CHADALAVADA DM, CERRONE-SZAKAL AL, BEVILACQUA PC. Wild-type is the optimal sequence of the HDV ribozyme under cotranscriptional conditions. RNA. 2007; 13:2189–2201. [PubMed: 17956974]

CHADALAVADA DM, KNUDSEN SM, NAKANO SI, BEVILACQUA PC. A role for upstream RNA structure in facilitating the catalytic fold of the genomic hepatitis delta virus ribozyme. Journal of Molecular Biology. 2000; 301:349–367. [PubMed: 10926514]

CHADALAVADA DM, SENCHAK SE, BEVILACQUA PC. The folding pathway of the genomic hepatitis delta virus ribozyme is dominated by slow folding of the pseudoknots1. Journal of Molecular Biology. 2002; 317(4):559–575. [PubMed: 11955009]

CHAUHAN S, WOODSON SA. Tertiary interactions determine the accuracy of RNA folding. Journal Of The American Chemical Society. 2008; 130:1296–1303. [PubMed: 18179212]

CHOI WG, SWANSON SJ, GILROY S. High-resolution imaging of Ca2+, redox status, ROS, and pH using GFP biosensors. The Plant Journal. 2012; 70(1):118–128. [PubMed: 22449047]

CLATTERBUCK SOPER SF, DATOR RP, LIMBACH PA, WOODSON SA. In vivo X-ray footprinting of pre-30S ribosomes reveals chaperone-dependent remodeling of late assembly intermediates. Molecular Cell. 2013; 52:506–516. [PubMed: 24207057]

CORDERO P, DAS R. Rich RNA Structure Landscapes Revealed by Mutate-and-Map Analysis. PLoS Comput Biol. 2015; 11(11):e1004473. [PubMed: 26566145]

CORDERO P, KLADWANG W, VANLANG CC, DAS R. Quantitative dimethyl sulfate mapping for automated RNA secondary structure inference. Biochemistry. 2012; 51(36):7037–7039. [PubMed: 22913637]

CROTHERS DM, COLE PE, HILBERS CW, SHULMAN RG. The molecular mechanism of thermal unfolding of *Escherichia coli* formylmethionine transfer RNA. Journal of Molecular Biology. 1974; 87:63–88. [PubMed: 4610153]

DAWSON WK, BUJNICKI JM. Computational modeling of RNA 3D structures and interactions. Current Opinion in Structural Biology. 2016; 37:22–28. [PubMed: 26689764]

DE MICHELE R, CARIMI F, FROMMER WB. Mitochondrial biosensors. The International Journal of Biochemistry & Cell Biology. 2014; 48:39–44. [PubMed: 24397954]

DEIGAN KE, LI TW, MATHEWS DH, WEEKS KM. Accurate SHAPE-directed RNA structure determination. Proceedings of the National Academy of Sciences. 2009; 106(1):97–102.

DENNING EJ, THIRUMALAI D, MACKERELL AD. Protonation of Trimethylamine N-Oxide (TMAO) is required for stabilization of RNA tertiary structure. Biophysical Chemistry. 2013; 184:8–16. [PubMed: 24012912]

DESAI R, KILBURN D, LEE HT, WOODSON S. Increased ribozyme acitivty in crowded solutions. Journal of Biological Chemistry. 2014; 289(5):2972–2977. [PubMed: 24337582]

DIAMOND JM, TURNER DH, MATHEWS DH. Thermodynamics of three-way multibranch loops in RNA. Biochemistry. 2001; 40:6971–6981. [PubMed: 11389613]

DING Y, KWOK CK, TANG Y, BEVILACQUA PC, ASSMANN SM. Genome-wide profiling of *in vivo* RNA structure at single-nucleotide resolution using Structure-seq. Nature Protocals. 2015; 10:1050–1066.

DING Y, LAWRENCE CE. A statistical sampling algorithm for RNA secondary structure prediction. Nucleic Acids Res. 2003; 31(24):7280–7301. [PubMed: 14654704]

DING Y, TANG Y, KWOK CK, ZHANG Y, BEVILACQUA PC, ASSMANN SM. *In vivo* genome-wide profiling of RNA secondary structure reveals novel regulatory features. Nature. 2014; 505:696–700. [PubMed: 24270811]

DO CB, FOO CS, BATZOGLOU S. A max-margin model for efficient simultaneous alignment and folding of RNA sequences. Bioinformatics. 2008; 24(13):i68–76. [PubMed: 18586747]

DOSHI KJ, CANNONE JJ, COBAUGH CW, GUTELL RR. Evaluation of the suitability of free-energy minimization using nearest-neighbor energy parameters for RNA secondary structure prediction. BMC Bioinformatics. 2004; 5(1):105. [PubMed: 15296519]

DOUDNA JA, CECH TR. The chemical repertoire of natural ribozymes. Nature. 2002; 418:222–228. [PubMed: 12110898]

DUPUIS NF, HOLMSTROM ED, NESBITT DJ. Molecular-crowding effects on single-molecule RNA folding/unfolding thermodynamics and kinetics. Proceedings of the National Academy of Sciences. 2014; 111(23):8464–8469.

DYER RB, BRAUNS EB. Laser-Induced Temperature Jump Infrared Measurements of RNA Folding. Methods in Enzymology. 2009; 469:353–372. [PubMed: 20946798]

EDDY SR. How do RNA folding algorithms work? Natuer Biotechnology. 2004; 22(11):1457–1458.

EDDY SR. Computational analysis of conserved RNA secondary structure in transcriptomes and genomes. Annual review of biophysics. 2014; 43:433–456.

EHRESMANN C, BAUDIN F, MOUGEL M, ROMBY P, EBEL JP, EHRESMANN B. Probing the structure of RNAs in solution. Nucleic Acids Research. 1987; 15:9109–9128. [PubMed: 2446263]

FANG R, MOSS WN, RUTENBERG-SCHOENBERG M, SIMON MD. Probing Xist RNA Structure in Cells Using Targeted Structure-Seq. PLoS Genet. 2015; 11(12):e1005668. [PubMed: 26646615]

FEHR M, LALONDE S, LAGER I, WOLFF MW, FROMMER WB. *In vivo* imaging of the dynamics fo glucose uptake in the cytosol of COS-7 cells by fluorescent nanosensors. Journal of Biological Chemistry. 2003; 278:19127–19133. [PubMed: 12649277]

FEIG, AL., UHLENBECK, OC. The role of metal ions in RNA biochemistry. In: Gesteland, RF.Cech, TR., Atkins, JF., editors. The RNA World. 2. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press; 1999. p. 287-320.

FIMOGNARI C. Role of Oxidative RNA Damage in Chronic-Degenerative Diseases. Oxidative Medicine and Cellular Longevity. 2015; 2015:8.

FRANKEL EA, BEVILACQUA PC, KEATING CD. Polyamine/Nucleotide Coacervates Provide Strong Compartmentalization of Mg2+, Nucleotides, and RNA. Langmuir. 2016; 32(8):2041–2049. [PubMed: 26844692]

FREIER SM, KIERZEK R, CARUTHERS MH, NEILSON T, TURNER DH. Free energy contributions of G.U and other terminal mismatches to helix stability. Biochemistry. 1986a; 25:3209–3223. [PubMed: 3730356]

FREIER SM, KIERZEK R, JAEGER JA, SUGIMOTO N, CARUTHERS MH, NEILSON T, TURNER DH. Improved free-energy parameters for predictions of RNA duplex stability. Proceedings of the National Academy of Sciences. 1986b; 83:9373–9377.

FU Y, SHARMA G, MATHEWS DH. Dynalign II: common secondary structure prediction for RNA homologs with domain insertions. Nucleic Acids Research. 2014; 42(22):13939–13948. [PubMed: 25416799]

GAO M, GNUTT D, ORBAN A, APPEL B, RIGHETTI F, WINTER R, NARBERHAUS F, MÜLLER S, EBBINGHAUS S. RNA hairpin folding in the crowded cell. Angewandte Chemie International Edition. 2016; 55:3224–3228. [PubMed: 26833452]

GARST AD, EDWARDS AL, BATEY RT. Riboswitches: Structures and Mechanisms. Cold Spring Harb Perspectives in Biology. 2011; 3:a0033533.

GERSTBERGER S, HAFNER M, TUSCHL T. A census of human RNA-binding proteins. Nature Reviews Genetics. 2014; 15:829–845.

GUERRIER-TAKADA C, GARDINER K, MARSH T, PACE N, ALTMAN S. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. Cell. 1983; 35:849–857. [PubMed: 6197186]

HAFNER M, LANDTHALER M, BURGER L, KHORSHID M, HAUSSER J, BERNINGER P, ROTHBALLER A, ASCANO M JR, JUNGKAMP AC, MUNSCHAUER M, ULRICH A, WARDLE GS, DEWELL S, ZAVOLAN M, TUSCHL T. Transcriptome-wide Identification of RNA-Binding Protein and MicroRNA Target Sites by PAR-CLIP. Cell. 2010; 141(1):129–141. [PubMed: 20371350]

HAJDIN CE, BELLAOUSOV S, HUGGINS W, LEONARD CW, MATHEWS DH, WEEKS KM. Accurate SHAPE-directed RNA secondary structure modeling, including pseudoknots. Proceedings of the National Academy of Sciences. 2013; 110(14):5498–5503.

HAJIAGHAYI M, CONDON A, HOOS HH. Analysis of energy-based algorithms for RNA secondary structure prediction. BMC Bioinformatics. 2012; 13:22. [PubMed: 22296803]

HALVORSEN M, MARTIN JS, BROADAWAY S, LAEDERACH A. Disease-associated mutations that alter the RNA structural ensemble. Plos Genetics. 2010; 6(8):e1001074. [PubMed: 20808897]

HARMANCI AO, SHARMA G, MATHEWS DH. PARTS: Probabilistic Alignment for RNA joinT Secondary structure prediction. Nucleic Acids Research. 2008; 36:2406–2417. [PubMed: 18304945]

HARMANCI AO, SHARMA G, MATHEWS DH. TurboFold: Iterative probabilistic estimation of secondary structures for multiple RNA sequences. BMC Bioinformatics. 2011; 12(1):108. [PubMed: 21507242]

HERSCHLAG D, CECH TR. Catalysis of RNA cleavage by the *Tetrahymena thermophil* ribozyme. 1. Kinetic description of the reaction of an RNA substrate complementary to the active site. Biochemistry. 1990; 29:10159–10171. [PubMed: 2271645]

HILBERS CW, ROBILLARD GT, SHULMAN RG, BLAKE RD, WEBB PK, FRESCO R, RIESNER D. Thermal unfolding of yeast glycine transfer RNA. Biochemistry. 1976; 15:1874–1882. [PubMed: 773427]

HOLMSTROM ED, DUPUIS NF, NESBITT DJ. Kinetic and thermodynamic origins of osmolyte-influenced nucleic acid folding. Journal of Physical Chemistry B. 2015; 119:3686–3696.

HOSEINI SS, SAUER MG. Molecular cloning using polymerase chain reaction, an educational guide for cellular engineering. Journal of Biological Engineering. 2015; 9(1):1–13. [PubMed: 25745515]

HU B, HU LL, CHEN ML, WANG JH. A FRET ratiometric fluorescence sensing system for mercury detection and intracellular colorimetric imaging in live Hela cells. Biosensors and Bioelectronics. 2013; 49:499–505. [PubMed: 23811485]

HULL CM, ANMANGANDLA A, BEVILACQUA PC. Bacterial riboswitches and ribozymes potently activate the human innate immune sensor PKR. ACS Chemical Biology. 2016; 11:1118–1127. [PubMed: 27011290]

HULL CM, BEVILACQUA PC. Mechanistic analysis of activation of the innate Immune sensor PKR by bacterial RNA. Journal of Molecular Biology. 2015; 427:3501–3515. [PubMed: 26026708]

HULL CM, BEVILACQUA PC. Discriminating self and non-self by RNA: roles for RNA structure, misfolding, and modification in regulating the innate immune sensor PKR. Accounts of Chemical Research. 2016; doi: 10.1021/acs.accounts.6b00151

IMAMURA H, HUYNH NHAT KP, TOGAWA H, SAITO K II, NO R, KATO-YAMADA Y, NAGAI T, NOJI H. Visualization of ATP levels inside single living cells with fluorescence resonance energy transfer-based genetically encoded indicators. Proceedings of the National Academy of Sciences of the United States of America. 2009; 106(37):15651–15656. [PubMed: 19720993]

INCARNATO D, NERI F, ANSELMI F, OLIVIERO S. Genome-wide profiling of mouse RNA secondary structures reveals key features of the mammalian transcriptome. Genome biology. 2014; 15:1–13.

JAEGER JA, ZUKER M, TURNER DH. Melting and chemical modification of a cyclized self-splicing group I intron: similarity of structures in 1 M Na^{+}, in 10 mM Mg^{2+}, and in the presence of substrate. Biochemistry. 1990; 29(44):10147–10158. [PubMed: 2271644]

JASPERS P, KANGASJÄRVI J. Reactive oxygen species in abiotic stress signaling. Physiologia Plantarum. 2010; 138(4):405–413. [PubMed: 20028478]

JIANG T, KENNEDY SD, MOSS WN, KIERZEK E, TURNER DH. Secondary Structure of a Conserved Domain in an Intron of Influenza A M1 mRNA. Biochemistry. 2014; 53(32):5236–5248. [PubMed: 25026548]

KERTESZ M, WAN Y, MAZOR E, RINN JL, NUTTER RC, CHANG HY, SEGAL E. Genome-wide measurement of RNA secondary structure in yeast. Nature. 2010; 467(2):103–107. [PubMed: 20811459]

KILBURN D, ROH JH, BEHROUZI R, BRIBER RM, WOODSON SA. Crowders perturb the entropy of RNA energy landscapes to favor folding. Journal Of The American Chemical Society. 2013; 135:10055–10063. [PubMed: 23773075]

KILBURN D, ROH JH, GUO L, BRIBER R, WOODSON S. Molecular crowding stabilizes folded RNA structure by the excluded volume efect. Journal Of The American Chemical Society. 2010; 132(25):8690–8696. [PubMed: 20521820]

KIM SH, QUIGLEY GJ, SUDDATH FL, MCPHERSON A, SNEDEN D, KIM JJ, WEINZIERL J, RICH A. Three-Dimensional Structure of Yeast Phenylalanine Transfer RNA: Folding of the Polynucleotide Chain. Science. 1973; 179(4070):285–288. [PubMed: 4566654]

KLOSTERMEIER D, MILLAR DP. RNA Conformation and Folding Studied with Fluorescence Resonance Energy Transfer. Methods. 2001; 23(3):240–254. [PubMed: 11243837]

KOGA S, WILLIAMS DS, PERRIMAN AW, MANN S. Peptide-nucleotide microdroplets as a step towards a membrane-free protocell model. Nature Chemistry. 2011; 3:720–724.

KUBODERA T, WATANABE M, YOSHIUCHI K, YAMASHITA N, NISHIMURA A, NAKAI S, GOMI K, HANAMOTO H. Thiamine-regulated gene expression of *Aspergillus oryzae thi*A requires splicing of the intron containing a riboswitch-like domain in the 5′-UTR. FEBS letters. 2003; 555:516–520. [PubMed: 14675766]

KUTCHKO KM, SANDERS W, ZIEHR B, PHILLIPS G, SOLEM A, HALVORSEN M, WEEKS KM, MOORMAN N, LAEDERACH A. Multiple conformations are a conserved and regulatory feature of the RB1 5′ UTR. RNA. 2015; 21(7):1274–1285. [PubMed: 25999316]

KWOK CK, DING Y, TANG Y, ASSMANN SM, BEVILACQUA PC. Determination of in vivo RNA structure in low-abundance transcripts. Nature Communications. 2013; 4doi: 10.1038/ncomms3971

KWOK CK, TANG Y, ASSMANN SM, BEVILACQUA JM. The RNA structurome: transcriptome-wide structure probing with next-generation sequencing. Trends in biochemical sciences. 2015; 40:221–232. [PubMed: 25797096]

LAGER I, LOOGER LL, HILPERT M, LALONDE S, FROMMER WB. Conversion of a putative agrobacterium sugar-ginding protein into a FRET sensor with high selectivity for sucrose. Journal of Biological Chemistry. 2006; 281:30875–30883. [PubMed: 16912038]

LAMBERT D, DRAPER DE. Effects of osmolytes on RNA secondary and tertiary structure stabilities and RNA-Mg2+ ion interactions. Journal of Molecular Biology. 2007; 370(5):993–1005. [PubMed: 17555763]

LAMBERT D, DRAPER DE. Denaturation of RNA secondary and tertiary structure by urea: simple unfolded state models and free energy parameters account for measured *m*-values. Biochemistry. 2012; 51:9014–9026. [PubMed: 23088364]

LAMBERT D, LEIPPLY D, DRAPER DE. The osmolyte TMAO stabilizes native RNA tertiary structures in the absence of Mg2+: evidence for a large barrier to folding form phosphate dehydration. Journal of Molecular Biology. 2010; 404(1):138–157. [PubMed: 20875423]

LAVENDER CA, GORELICK RJ, WEEKS KM. Structure-Based Alignment and Consensus Secondary Structures for Three HIV-Related RNA Genomes. PLoS Comput Biol. 2015a; 11(5):e1004230. [PubMed: 25992893]

LAVENDER CA, LORENZ R, ZHANG G, TAMAYO R, HOFACKER IL, WEEKS KM. Model-Free RNA Sequence and Structure Alignment Informed by SHAPE Probing Reveals a Conserved Alternate Secondary Structure for 16S rRNA. PLoS Comput Biol. 2015b; 11(5):e1004126. [PubMed: 25992778]

LEVITT M. Detailed moecular model for transfer ribonucleic acid. Nature. 1969; 224:759–763. [PubMed: 5361649]

LI C, WEN A, SHEN B, LU J, HUANG Y, CHANG Y. FastCloning: a highly simplified, purification-free, sequence- and ligation-independent PCR cloning method. BMC Biotechnology. 2011; 11(92)

LI F, ZHENG Q, VANDIVIER LE, WILLMANN MR, CHEN Y, GREGORY BD. Regulatory impact of RNA secondary structure across the *Arabidopsis* transcriptome. The Plant Cell. 2012; 24:4346–4359. [PubMed: 23150631]

LICATALOSI DD, MELE A, FAK JJ, ULE J, KAYIKCI M, CHI SW, CLARK TA, SCHWEITZER AC, BLUME JE, WANG X, DARNELL JC, DARNELL RB. HITS-CLIP yields genome-wide insights into brain alternative RNA processing. Nature. 2008; 456(7221):464–469. [PubMed: 18978773]

LINDENBURG LH, VINKENBORG JL, OORTWIJN J, APER SJA, MERKX M. MagFRET: The First Genetically Encoded Fluorescent Mg(2+) Sensor. PLoS One. 2013; 8(12):e82009. [PubMed: 24312622]

LIU B, DIAMOND JM, MATHEWS DH, TURNER DH. Fluorescence competition and optical melting measurements of RNA three-way multibranch loops provide a revised model for thermodynamic parameters. Biochemistry. 2011; 50(5):640–653. [PubMed: 21133351]

LIU B, MATHEWS DH, TURNER DH. RNA pseudoknots: folding and finding. F1000 Biol Rep. 2010a; 2:8. [PubMed: 20495679]

LIU B, SHANKAR N, TURNER DH. Fluorescence competition assay measurements of free energy changes for RNA pseudoknots. Biochemistry. 2010b; 49(3):623–634. [PubMed: 19921809]

LONDON RE. Methods for measurement of intracellular magnesium: NMR and fluorescence. Annual Reviews of Physiology. 1991; 53:241–258.

LORENZ R, WOLFINGER MT, TANZER A, HOFACKER IL. Predicting RNA Secondary Structures from Sequence and Probing Data. Methods.

LU ZJ, GLOOR JW, MATHEWS DH. Improved RNA secondary structure prediction by maximizing expected pair accuracy. RNA. 2009; 15:1805–1813. [PubMed: 19703939]

LU ZJ, TURNER DH, MATHEWS DH. A set of nearest neighbor parameters for predicting the enthalpy change of RNA secondary formation. Nucleic Acids Research. 2006; 34(17):4912–4924. [PubMed: 16982646]

LUSK JE, WILLIAMS RJ, KENNEDY EP. Magnesium and the growth of *Escherichia Coli*. Journal of Biological Chemistry. 1968; 243:2618–2624. [PubMed: 4968384]

MAHEN EM, HARGER JW, CALDERON EM, FEDOR MJ. Kinetics and thermodynamics make different contributions to RNA folding in vitro and in yeast. Molecular Cell. 2005; 19(1):27–37. [PubMed: 15989962]

MATHEWS DH. Using an RNA secondary structure partition function to determine confidence in base pairs predicted by free energy minimization. RNA. 2004; 10:1178–1190. [PubMed: 15272118]

MATHEWS DH. Revolutions in RNA secondary structure prediction. Journal of Molecular Biology. 2006; 359:526–532. [PubMed: 16500677]

MATHEWS DH, DISNEY MD, CHILDS JL, SCHROEDER SJ, ZUKER M, TURNER DH. Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. Proceedings of the National Academy of Sciences. 2004; 101:7287–7292.

MATTEUCCI MD, CARUTHERS MH. Synthesis of deoxyoligonucleotides on a polymer support. Journal Of The American Chemical Society. 1981; 103(11):3185–3191.

MCCASKILL JS. The equilibrium partition function and base pair probabilities for RNA secondary structure. Biopolymers. 1990; 29:1105–1119. [PubMed: 1695107]

MERINO EJ, WILKINSON KA, COUGHLAN JL, WEEKS KM. RNA structure analysis at single nucleotide resolution by selective 2′-hydoxyl acylation and primer extension (SHAPE). Journal Of The American Chemical Society. 2005; 127(4223–4231):4223. [PubMed: 15783204]

MIAO Z, ADAMIAK RW, BLANCHET MF, BONIECKI M, BUJNICKI JM, CHEN SJ, CHENG C, CHOJNOWSKI G, CHOU FC, CORDERO P, CRUZ JA, FERRE-D'AMARÉ AR, DAS R, DING F, DOKHOLYAN NV, DUNIN-HORKAWICZ S, KLADWANG W, KROKHOTIN A, LACH G, MAGNUS M, MAJOR F, MANN TH, MASQUIDA B, MATELSKA D, MEYER M, PESELIS A, POPENDA M, PURZYCKA KJ, SERGANOV A, STASIEWICZ J, SZACHNIUK M, TANDON A, TIAN S, WANG J, XIAO Y, XU X, ZHANG J, ZHAO P, ZOK T, WESTHOF

E. RNA-Puzzles Round II: assessment of RNA structure prediction programs applied to three large RNA structures. RNA. 2015; 21(6):1066–1084. [PubMed: 25883046]

MILLIGAN JF, GROEBE DR, WITHERELL GW, UHLENBECK OC. Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. Nucleic Acids Research. 1987; 15(21): 8783–8798. [PubMed: 3684574]

MINTON AP. The influence of macromolecular crowding and macromolecular confinement on biochemical media. Journal of Biological Chemistry. 2001; 276:10577–10589. [PubMed: 11279227]

MITCHELL DI, JARMOSKAITE I, SEVAL N, SEIFERT S, RUSSELL R. The long-range P3 helix of the *Tetrahymena* ribozyme is disrupted during folding between the native and misfolded conformations. Journal of Molecular Biology. 2013; 425:2670–2686. [PubMed: 23702292]

MITCHELL DI, RUSSELL R. Folding pathways of the Tetrahymena ribozyme. Journal of Molecular Biology. 2014; 426:2300–2312. [PubMed: 24747051]

MOAZED D, STERN S, NOLLER HF. Rapid chemical probing of conformation in 16 S ribosomal RNA and 30 S ribosomal subunits using primer extension. Journal of Molecular Biology. 1986a; 187(3):399–416. [PubMed: 2422386]

MOAZED D, STERN S, NOLLER HF. Rapid chemical probing of conformation in 16S ribosomal RNA and 30S ribosomal subunits using primer extension. Journal of Molecular Biology. 1986b; 187:399–416. [PubMed: 2422386]

MOORE M, SHARP P. Site-specific modification of pre-mRNA: the 2′-hydroxyl groups at the splice sites. Science. 1992; 256(5059):992–997. [PubMed: 1589782]

MULLIS KB. The unusual origin of the polymerase chain reaction. Scientific American. 1990; 262:56–61. 64–65. [PubMed: 2315679]

NADARAJAN SP, RAVIKUMAR Y, DEEPANKUMAR K, LEE CS, YUN H. Engineering lead-sensing GFP through rational designing. Chemical Communications. 2014; 50(100):15979–15982. [PubMed: 25383613]

NAKANO SI, KARIMATA HT, KITAGAWA Y, SUGIMOTO N. Facilitation of RNA enzyme activity in the molecular crowding media of cosolutes. Journal Of The American Chemical Society. 2009; 131:16881–16888. [PubMed: 19874030]

NAKANO S-I, KITAGAWA Y, YAMASHITA H, MIYOSHI D, SUGIMOTO N. Effects of cosolvents on the folding and catalytic activities of the hammerhead ribozyme. Chem Bio Chem. 2015

NAKANO SI, MIYOSHI D, SUGIMOTO N. Effects of Molecular Crowding on the Structures, Interactions, and Functions of Nucleic Acids. Chemical Reviews. 2014; 114(5):2733–2758. [PubMed: 24364729]

NALLAGATLA SR, HWANG J, TORONEY R, ZHENG X, CAMERON CE, BEVILACQUA PC. 5′-Triphosphate-Dependent Activation of PKR by RNAs with Short Stem-Loops. Science. 2007; 318(5855):1455–1458. [PubMed: 18048689]

NICK H, GILBERT W. Detection *in vivo* of protein-DNA interactions within the lac Operon of *Escherichia Coli*. Nature. 1985; 313:795–798. [PubMed: 3883194]

NISSEN P, HANSEN J, BAN N, MOORE PB, STEITZ TA. The structural basis of ribosome activity in peptide bond synthesis. Science. 2000; 289:920–930. [PubMed: 10937990]

NOVIKOVA IV, HENNELLY SP, SANBONMATSU KY. Structural architecture of the human long non-coding RNA, steroid receptor RNA activator. Nucleic Acids Research. 2012; 40(11):5034–5051. [PubMed: 22362738]

OSBORNE RJ, THORNTON CA. RNA-dominant diseases. Hum Mol Genet. 2006; 15(2):R162–R169. [PubMed: 16987879]

OUYANG Z, SNYDER MP, CHANG HY. SeqFold: genome-scale reconstruction of RNA secondary structure integrating high-throughput sequencing data. Genome Res. 2013; 23(2):377–387. [PubMed: 23064747]

PAIGE JS, DUC TN, SONG W, JAFFREY SR. Fluorescence imaging of cellular metabolites with RNA. Science (New York, Ny). 2012; 335(6073):1194–1194.

PAIGE JS, WU K, JAFFREY SR. RNA mimics of green fluorescent protein. Science (New York, Ny). 2011; 333(6042):642–646.

PAUDEL BP, RUEDA D. Molecular crowding accelerates ribozymes docking and catalysis. Journal Of The American Chemical Society. 2014; 136:16700–16703. [PubMed: 25399908]

POLLACK L. Time resolved SAXS and RNA folding. Biopolymers. 2011; 95(8):543–549. [PubMed: 21328311]

POUVREAU S. Genetically encoded reactive oxygen species (ROS) and redox indicators. Biotechnology Journal. 2014; 9(2):282–293. [PubMed: 24497389]

RANGAN P, MASUIDA B, WESTHOF E, WOODSON SA. Assembly of core helices and rapid tertiary folding of a small bacterial group I ribozyme. Proceedings of the National Academy of Sciences. 2003; 100(4):1574–1579.

RECORD MTJ, COURTENAY ES, CAYLEY SD, GUTTMAN HJ. Responses of *E. coli* to osmotic stress: large changes in amounts of cytoplasmic solutes and water. Trends in biochemical sciences. 1998; 23(4):143–148. [PubMed: 9584618]

REEDER J, GIEGERICH R. Consensus shapes: an alternative to the Sankoff algorithm for RNA consensus structure prediction. Bioinformatics. 2005; 21(17):3516–3523. [PubMed: 16020472]

REEDER J, HOCHSMANN M, REHMSMEIER M, VOSS B, GIEGERICH R. Beyond Mfold: recent advances in RNA bioinformatics. Journal of Biotechnology. 2006; 124(1):41–55. [PubMed: 16530285]

REYES, FE., GARST, AD., BATEY, RT. Methods in Enzymology. Vol. 469. Academic Press; 2009. Chapter 6 - Strategies in RNA Crystallography; p. 119-139.

RICHARDS EG, FLESSEL CP, FRESCO JR. Polynucleotides. IV. Molecular properties and conformations of polyribonucleic acids. Biopolymers. 1963; 1:431–446.

ROBERTUS JD, LADNER JE, FINCH JT, RHODES D, BROWN RS, CLARK BFC, KLUG A. Structure of yeast phenylalanine tRNA at 3 Å resolution. Nature. 1974; 250:546–551. [PubMed: 4602655]

ROH JH, GUO L, KILBURN D, BRIBER R, IRVING T, WOODSON S. Multistage collapse of a bacterial ribozyme observed by time-resolved small-angle X-ray scattering. Journal Of The American Chemical Society. 2010; 132:10148–10154. [PubMed: 20597502]

ROMANI AM. Magnesium homeostasis in mammalian cells. Frontiers in Bioscience. 2007; 12:308–331. [PubMed: 17127301]

ROOK MS, TREIBER DK, WILLIAMSON JR. Fast folding mutants of the *Tetrahymena* group I ribozyme reveal a rugged folding energy landscape. Journal of Molecular Biology. 1998; 281:609–620. [PubMed: 9710534]

ROTH A, WEINBERG Z, CHEN AGY, KIM PB, AMES TD, BREAKER RR. A widespread self-cleaving ribozymes class is revealed by bioinformatics. Nature Chemical Biology. 2014; 10(1): 56–60. [PubMed: 24240507]

ROUSKIN S, ZUBRADT M, WASHIETL S, KELLIS M, WEISSMAN JS. Genome-wide probing of RNA structure reveals active unfolding of mRNA structures in vivo. Nature. 2014; 505:701–705. [PubMed: 24336214]

ROY R, HOHNG S, HA T. A practical guide to single-molecule FRET. Nature Methods. 2008; 5(6): 507–516. [PubMed: 18511918]

SALEHI-ASHTIANI K, LUPTAK A, LITOVCHICK A, SZOSTAK JW. A genomewide search for ribozymes reveals an HDV-like sequence in the Human CPEB3 gene. Science. 2006; 313:1788–1792. [PubMed: 16990549]

SANTANGELO P, NITIN N, BAO G. Nanostructured Probes for RNA Detection in Living Cells. Annals of Biomedical Engineering. 2006; 34(1):39–50. [PubMed: 16463087]

SCALVI B, WOODSON S, SULLIVAN M, CHANCE MR, BRENOWITZ M. Time-resolved synchrotron X-ray “footprinting”, a new approach to the study of nucleic acid structure and function: application to protein-DNA interactions and RNA folding. Journal of Molecular Biology. 1997; 266:144–159. [PubMed: 9054977]

SCARINGE S, WINCOTT FE, CARUTHERS MH. Novel RNA synthesis method using 5′-O-silyl-2′-orthoester protecting groups. Journal Of The American Chemical Society. 1998; 120:11820–11821.

SCHROEDER SJ, TURNER DH. Factors affecting the thermodynamic stability of small asymmetric internal loops in RNA. Biochemistry. 2000; 39:9257–9274. [PubMed: 10924119]

SCHROEDER SJ, TURNER DH. Opical melting measurements of nucleic acid thermodynamics. Methods Enzymology. 2009; 468:371–387.

SCLAVI B, SULLIVAN M, CHANGE MR, BRENOWITZ M, WOODSON S. RNA folding at millisecond intervals by synchrotron hydroxyl radical footprinting. Science. 1998; 279:1940–1943. [PubMed: 9506944]

SEETIN MG, MATHEWS DH. RNA structure prediction: an overview of methods. Methods in Molecular Biology. 2012a; 905:99–122. [PubMed: 22736001]

SEETIN MG, MATHEWS DH. TurboKnot: Rapid Prediction of Conserved RNA Secondary Structures Including Pseudoknots. Bioinformatics. 2012b; 28:792–798. [PubMed: 22285566]

SERGANOV A, NUDLER E. A decade of riboswitches. Cell. 2013; 152(1–2):17–24. [PubMed: 23332744]

SERGANOV A, PATEL D. Ribozymes, riboswitches and beyond: regulation of gene expression without proteins. Nature. 2007; 8:776–790.

SERRA MJ, BAIRD JD, DALE T, FEY BL, RETATAGOS K, WESTHOF E. Effects of magnesium ions on the stabilization of RNA oligomers of defined structures. RNA. 2002; 8(3):307–323. [PubMed: 12003491]

SHERPA C, RAUSCH JW, LE GRICE SFJ, HAMMARSKJOLD ML, REKOSH D. The HIV-1 Rev response element (RRE) adopts alternative conformations that promote different rates of virus replication. Nucleic Acids Research. 2015; 43(9):4676–4686. [PubMed: 25855816]

SIERZCHALA A, DELLINGER DJ, BETLEY JR, WYRZYKIEWICZ TK, YAMADA CM, CARUTHERS MH. Solid-phase oligodeoxynucleotide synthesis: A two-step cycle using peroxy anion deprotection. Journal Of The American Chemical Society. 2003; 125:13427–13441. [PubMed: 14583038]

SLOMA MF, MATHEWS DH. Improving RNA secondary structure prediction with structure mapping data. Methods in Enzymology. 2015; 553:91–114. [PubMed: 25726462]

SOLOMATIN SV, GREENFELD M, CHU S, HERSCHLAG D. Multiple native states reveal persistent ruggedness of an RNA folding landscape. Nature. 2010; 463:681–684. [PubMed: 20130651]

SOMAROWTHU S, LEGIEWICZ M, CHILLÓN I, MARCIA M, LIU F, PYLE ANNA M. HOTAIR Forms an Intricate and Modular Secondary Structure. Molecular Cell. 58(2):353–361.

SOTO AM, MISRA V, DRAPER DE. Tertiary structure of an RNA pseudoknot is stabilized by “diffuse” Mg2+ ions. Biochemistry. 2007; 46(11):2973–2983. [PubMed: 17315982]

SPITALE RC, FLYNN RA, ZHANG QC, CRISALLI P, LEE B, JUNG JW, KUCHELMEISTER HY, BATISTA PJ, TORRE EA, KOOL ET, CHANG HY. Structural imprints in vivo decode RNA regulatory mechanisms. Nature. 2015; 519(7544):486–490. [PubMed: 25799993]

STAEL S, WURZINGER B, MAIR A, MEHLMER N, VOTHKNECHT UC, TEIGE M. Plant organellar calcium signalling: an emerging field. Journal of Experimental Botany. 2011; 63:1525–1542. [PubMed: 22200666]

STEIN A, CROTHERS DM. Conformational changes of transfer RNA. The role of magnesium(II). Biochemistry. 1976; 15(1):160–168. [PubMed: 764858]

STRULSON CA, BOYER JA, WHITMAN EE, BEVILACQUA PC. Molecular crowders and cosolutes promote folding cooperativity of RNA under physiological ionic conditions. RNA. 2014; 20(3):331–347. [PubMed: 24442612]

STRULSON CA, MOLDEN RC, KEATING CD, BEVILACQUA PC. RNA catalysis through compartmentalization. Nature Chemistry. 2012; 4:941–946.

STRULSON CA, YENNAWAR NH, RAMBO RP, BEVILACQUA PC. Molecular crowding favors reactivity of a human ribozyme under physiological ionic conditions. Biochemistry. 2013; 52:8187–8197. [PubMed: 24187989]

SÜKÖSD Z, ANDERSEN ES, SEEMANN SE, JENSEN MK, HANSEN M, GORODKIN J, KJEMS J. Full-length RNA structure prediction of the HIV-1 genome reveals a conserved core domain. Nucleic Acids Research. 2015; 43(21):10168–10179. [PubMed: 26476446]

SÜKÖSD Z, KNUDSEN B, KJEMS J, PEDERSEN CNS. PPfold 3.0: fast RNA secondary structure prediction using phylogeny and auxiliary data. Bioinformatics. 2012; 28(20):2691–2692. [PubMed: 22877864]

SUSSMAN JL, HOLBROOK SR, WARRANT W, CHURCH GM, KIM SH. Crystal structure of yeast phenylalanine transfer RNA. 1. Crystallographic refinement. Journal of Molecular Biology. 1978; 123:607–630. [PubMed: 357742]

SUURKUUSK J, ALVAREZ J, FREIRE E, BILTONEN R. Calorimetric determination of the heat capacity changes associated with the conformational transitions of polyriboadenylic acid and polyribouridylic acid. Biopolymers. 1977; 16(12):2641–2652. [PubMed: 597574]

SWANSON SJ, CHOI WG, CHANOCA A, GILROY S. In Vivo Imaging of Ca2+, pH, and Reactive Oxygen Species Using Fluorescent Probes in Plants. Annual Review of Plant Biology. 2011; 62(1):273–297.

SWISHER JF, SU LJ, BRENOWITZ M, ANDERSON VE, PYLE AM. Productive folding to the native state by a Group II intron ribozyme. Journal of Molecular Biology. 2002; 315:297–310. [PubMed: 11786013]

TALKISH J, MAY G, LIN Y, WOOLFORD JLJ, MCMANUS CJ. Mod-seq: high-throughput sequencing for chemical probind of RNA structure. RNA. 2014; 20:713–720. [PubMed: 24664469]

TANG S, REDDISH F, ZHUO Y, YANG JJ. Fast kinetics of calcium signaling and sensor design. Current Opinion in Chemical Biology. 2015a; 27:90–97. [PubMed: 26151819]

TANG Y, BOUVIER E, KWOK CK, DING Y, NEKRUTENKO A, BEVILACQUA PC, ASSMANN SM. StructureFold: genome-wide RNA secondary structure mapping and reconstruction *in vivo*. Bioinformatics. 2015b; 31:2668–2675. [PubMed: 25886980]

TANNER MA, CECH TR. Activity and thermostability of the small self-splicing group I intron in the pre-tRNAIle of the purple bacterium *Azoarcus*. RNA. 1996; 2:74–83. [PubMed: 8846298]

TANTAMA M, HUNG YP, YELLEN G. Imaging Intracellular pH in Live Cells with a Genetically Encoded Red Fluorescent Protein Sensor. Journal Of The American Chemical Society. 2011; 133(26):10034–10037. [PubMed: 21631110]

TINOCO IJ, BUSTAMANTE C. How RNA folds. Journal of Molecular Biology. 1999; 293:271–281. [PubMed: 10550208]

TORARINSSON E, HAVGAARD JH, GORODKIN J. Multiple structural alignment and clustering of RNA sequences. Bioinformatics. 2007; 23(8):926–932. [PubMed: 17324941]

TREIBER DK, ROOK MS, ZARRINKAR PP, WILLIAMSON JR. Kinetic intermediates trapped by native interactions in RNA folding. Science. 1998; 279:1943–1946. [PubMed: 9506945]

TRUONG DM, SIDOTE DJ, RUSSELL R, LAMBOWITZ AM. Enhanced group II intron retrohoming in magnesium-deficient *Escherichia coli* via selection of mutations in the ribozyme core. Proceedings of the National Academy of Sciences. 2013; 110:E3800–E3809.

TSIEN RY. The 2009 Lindau Nobel Laureate Meeting: Roger Y. Tsien, Chemistry 2008. Journal of Visualized Experiments : JoVE. 2010; (35):1575. [PubMed: 20072108]

TURNER DH, MATHEWS DH. NNDB: the nearest neighbor parameter database for predicting stability of nucleic acidsecondary structure. Nucleic Acids Research. 2010; 38:D280–D282. [PubMed: 19880381]

TYRRELL J, MCGINNIS JL, WEEKS KM, PIELAK GJ. The cellular envirnment stabilized adenine riboswitch RNA structure. Biochemistry. 2013; 52:8777–8785. [PubMed: 24215455]

TYRRELL J, WEEKS KM, PIELAK GJ. Challenge of mimicking the influence of the cellular environment on RNA structure by PEG-induced macromolecular crowding. Biochemistry. 2015; 54:6447–6453. [PubMed: 26430778]

UNDERWOOD JG, UZILOV AV, KATZMAN S, ONODERA CS, MAINZER JE, MATHEWS DH, LOWE TM, SALAMA SR, HAUSSLER D. FragSeq: transcriptome-wide RNA structure probing using high-throughput sequencing. Nature Methods. 2010; 7(12):995–1001. [PubMed: 21057495]

WALTER NG. Structural Dynamics of Catalytic RNA Highlighted by Fluorescence Resonance Energy Transfer. Methods. 2001; 25:19–30. [PubMed: 11558994]

WAN Y, KERTESZ M, SPITALE RC, SEGAL E, CHANG HY. Understanding the transcriptome through RNA structure. Nature Reviews Genetics. 2011; 12(9):641–655.

WAN Y, QU K, OUYANG Z, KERTESZ M, LI J, TIBSHIRANI R, MAKINO DL, NUTTER RC, SEGAL E, CHANG HY. Genome-wide measurement of RNA folding energies. Molecular Cell. 2012; 48:169–181. [PubMed: 22981864]

WAN Y, QU K, ZHANG QC, FLYNN RA, MANOR O, OUYANG Z, ZHANG J, SPITALE RC, SNYDER MP, SEGAL E, CHANG HY. Landscape and variation of RNA secondary structure across the human transcriptome. Nature. 2014; 505:706–709. [PubMed: 24476892]

WAN Y, SUH H, RUSSELL R, HERSCHLAG D. Multiple unfolding events during native folding of the *Tetrahymena* group I ribozyme. Journal of Molecular Biology. 2010; 400:1067–1077. [PubMed: 20541557]

WASHIETL S, HOFACKER IL, STADLER PF, KELLIS M. RNA folding with soft constraints: reconciliation of probing data and thermodynamic secondary structure prediction. Nucleic Acids Research. 2012; 40(10):4261–4272. [PubMed: 22287623]

WEEKS KM. Advances in RNA Secondary and Tertiary Structure Analysis by Chemical Probing. Current Opinion in Structural Biology. 2010; 20(3):295–304. [PubMed: 20447823]

WEYN-VANHENTENRYCK SEBASTIENM, MELE A, YAN Q, SUN S, FARNY N, ZHANG Z, XUE C, HERRE M, SILVER PAMELAA, ZHANG MICHAELQ, KRAINER ADRIANR, DARNELL ROBERTB, ZHANG C. HITS-CLIP and Integrative Modeling Define the Rbfox Splicing-Regulatory Network Linked to Brain Development and Autism. Cell Reports. 2014; 6(6):1139–1152. [PubMed: 24613350]

WICKISER JK, CHEAH MT, BREAKER RR, CROTHERS DM. The Kinetics of Ligand Binding by an Adenine-Sensing Riboswitch. Biochemistry. 2005; 44(40):13404–13414. [PubMed: 16201765]

WICKISER JK, WINKLER WC, BREAKER RR, CROTHERS DM. The Speed of RNA Transcription and Metabolite Binding Kinetics Operate an FMN Riboswitch. Molecular Cell. 18(1):49–60.

WILL S, REICHE K, HOFACKER IL, STADLER PF, BACKOFEN R. Inferring noncoding RNA families and classes by means of genome-scale structure-based clustering. PLoS Comput Biol. 2007; 3(4):e65. [PubMed: 17432929]

WU Y, SHI B, DING X, LIU T, HU X, YIP KY, YANG ZR, MATHEWS DH, LU ZJ. Improved prediction of RNA secondary structure by integrating the free energy model with restraints derived from experimental probing data. Nucleic Acids Research. 2015; 43(15):7247–7259. [PubMed: 26170232]

WUCHTY S, FONTANA W, HOFACKER IL, SCHUSTER P. Complete suboptimal folding of RNA and the stability of secondary structures. Biopolymers. 1999; 49:145–165. [PubMed: 10070264]

XIA T, SANTALUCIA J, BURKARD ME, KIERZEK R, SCHROEDER SJ, JIAO X, COX C, TURNER DH. Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs. Biochemistry. 1998; 37:14719–14735. [PubMed: 9778347]

XU Z, MATHEWS DH. Multilign: an algorithm to predict secondary structures conserved in multiple RNA sequences. Bioinformatics. 2011; 27(5):626–632. [PubMed: 21193521]

YANCEY PH, CLARK ME, HAND SC, BOWLUS RD, SOMERO GN. Living with water stress: evolution of osmolyte systems. Science. 1982; 217(4566):1214–1222. [PubMed: 7112124]

YANG S. Methods for SAXS-based Structure Determination of Biomolecular Complexes. Advanced Materials. 2014; 26:7902–7910. [PubMed: 24888261]

YANG Z, CAO J, HE Y, YANG JH, KIM T, PENG X, KIM JS. Macro-/micro-environment-sensitive chemosensing and biological imaging. Chemical Society Reviews. 2014; 43(13):4563–4601. [PubMed: 24723011]

YOU M, JAFFREY SR. Structure and Mechanism of RNA Mimics of Green Fluorescent Protein. Annual review of biophysics. 2015; 44(1):187–206.

YOU M, LITKE JL, JAFFREY SR. Imaging metabolite dynamics in living cells using a Spinach-based riboswitch. Proceedings of the National Academy of Sciences of the United States of America. 2015; 112(21):E2756–E2765. [PubMed: 25964329]

ZARRINGHALAM K, MEYER MM, DOTU I, CHUANG JH, CLOTE P. Integrating Chemical Footprinting Data into RNA Secondary Structure Prediction. PLoS One. 2012; 7(10):e45160. [PubMed: 23091593]

ZARRINKAR PP, WANG J, WILLIAMSON JR. Slow folding kinetics of RNase P RNA. RNA. 1996; 2(6):564–573. [PubMed: 8718685]

ZAUG A, CECH TR. Analysis of the structure of *Tetrahymena* nuclear RNAs in vivo: Telomerase RNA, the self-splicing rRNA intron, and U2 snRNA. RNA. 1995; 1:363–374. [PubMed: 7493315]

ZHENG Q, RYVKIN P, LI F, DRAGOMIR I, VALLADARES O, YANG J, CAO K, WANG LS, GREGORY BD. Genome-Wide Double-Stranded RNA Sequencing Reveals the Functional Significance of Base-Paired RNAs in *Arabidopsis*. PLoS Genet. 2010; 6(9):e1001141. [PubMed: 20941385]

ZHUANG X, BARTLEY LE, BABCOCK HP, RUSSELL R, HA T, HERSCHLAG D, CHU S. A Single-Molecule Study of RNA Catalysis and Folding. Science. 2000; 288(2048–2051):2048. [PubMed: 10856219]

ZIMMERMAN SB, TRACH SO. Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of Escherichia coli. Journal of Molecular Biology. 1991; 222(3):599–620. [PubMed: 1748995]

ZUKER M. On finding all suboptimal foldings of an RNA molecule. Science. 1989; 244:48–52. [PubMed: 2468181]

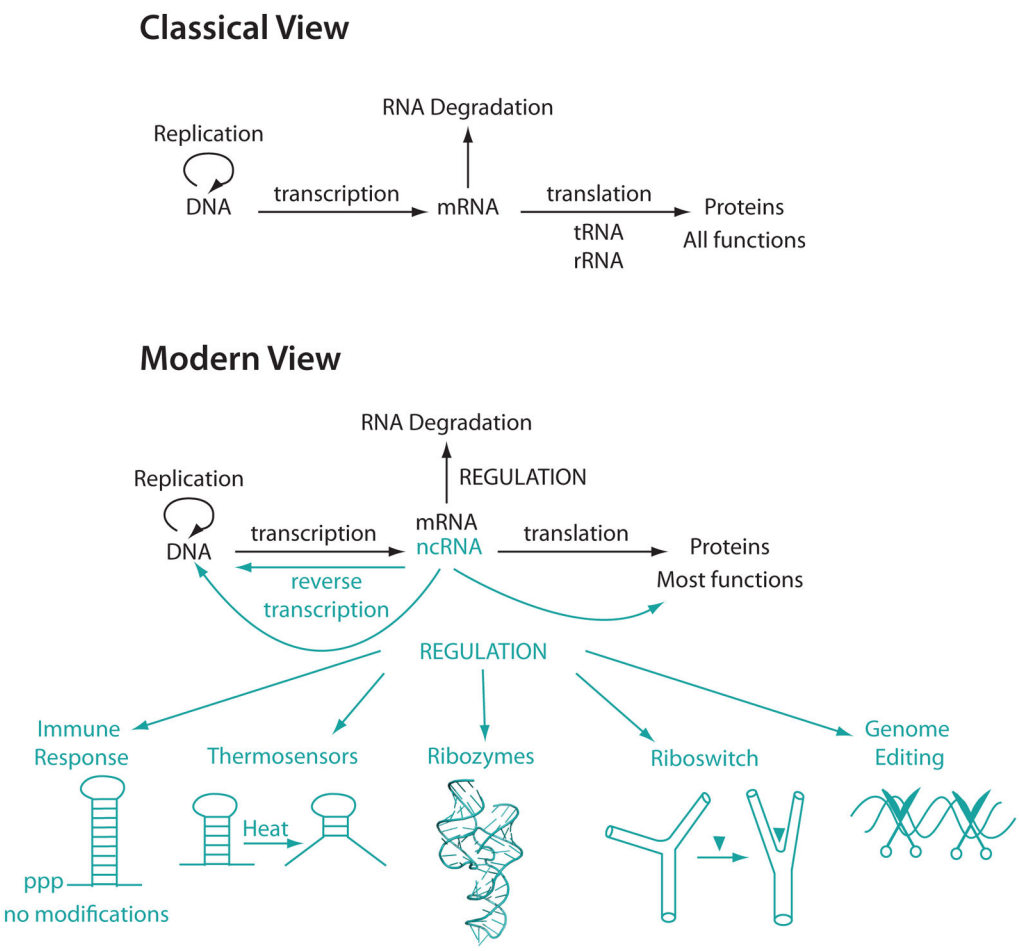


Figure 1.

The Classical View (top) and the Modern View (bottom) of RNA's role in biology. In the classical view of biology, RNA (top) serves as a messenger molecule between DNA and proteins and proteins have all the main functions in cells. Messenger RNA serves to translate information from DNA to proteins. The modern view of biology (bottom) has emerged in the last 25 years as the field learns more about the many functions of RNA. Noncoding RNA (ncRNA) has vast regulatory functions, some of which include immune responses ('ppp'=5′-triphosphate, which activates PKR)(Nallagatla et al., 2007), thermosensors, ribozymes, riboswitches, and genome editing. In the modern view of biology, proteins still have most cellular functions, but RNA plays essential roles in the cell beyond its classical functions.

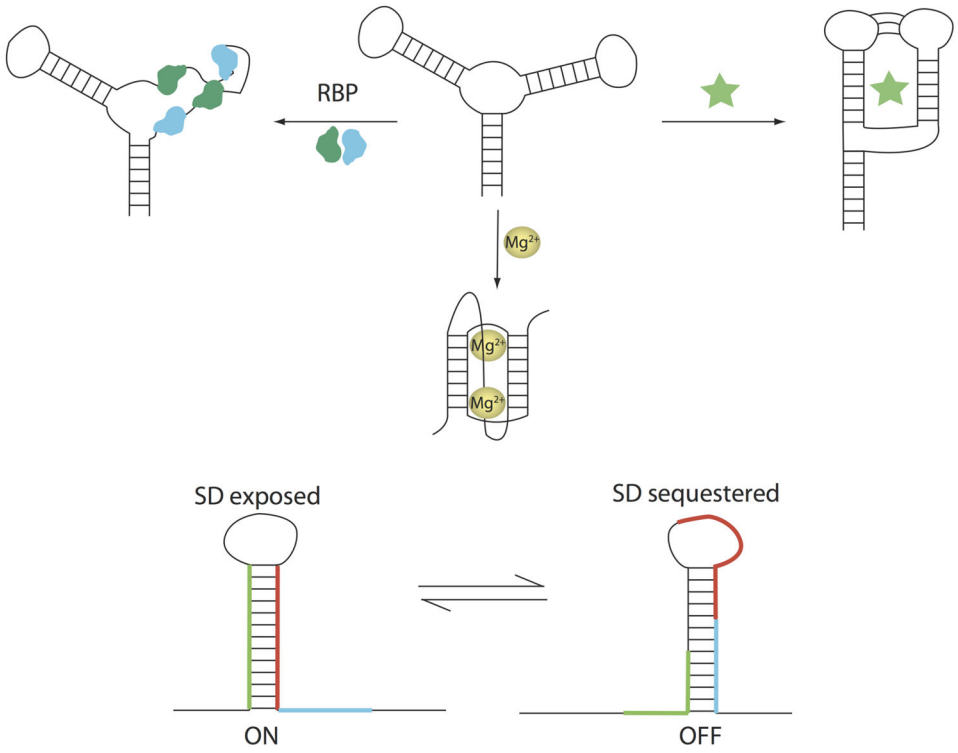


Figure 2.

RNA interactions with RNA binding proteins (RBP, left), metal ions (central), and ligands (star, right) can result in structure changes. Unlike typical *in vitro* conditions, there are other molecules and complex solutions conditions *in vivo* that can interact with RNA and change its structure. These structure changes can result in an RNA with less structure (top left) more structure (top right), or an alternate conformation than the structure that is prevalent *in vitro* (bottom). Also shown (bottom) are the bacterial expression platforms of riboswitches that switch between two mutually exclusive structures that turn a gene ON (left) or OFF (right) by exposing or sequestering the Shine-Dalgarno sequence (blue).

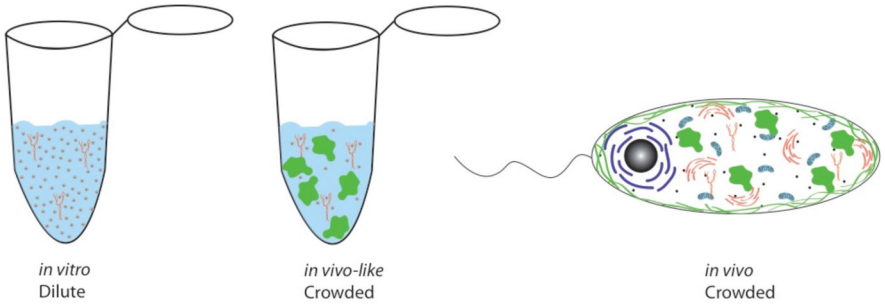


Figure 3.

Artist's rendition of *in vitro* conditions (left), *in vivo* conditions (right) and *in vivo-like* conditions (center). Typical *in vitro* solutions are dilute with high monovalent ion concentrations that are very different from cellular conditions. The cellular environment is complex with monovalent and divalent salts, macromolecules, cosolutes, and organelles. *In vivo-like* conditions (center) bridge *in vitro* and *in vivo* conditions and are more complex than *in vitro* conditions with added synthetic crowding agents and proteins and physiological ion concentrations. However *in vivo-like* conditions are still much less complex than those prevailing *in vivo*.

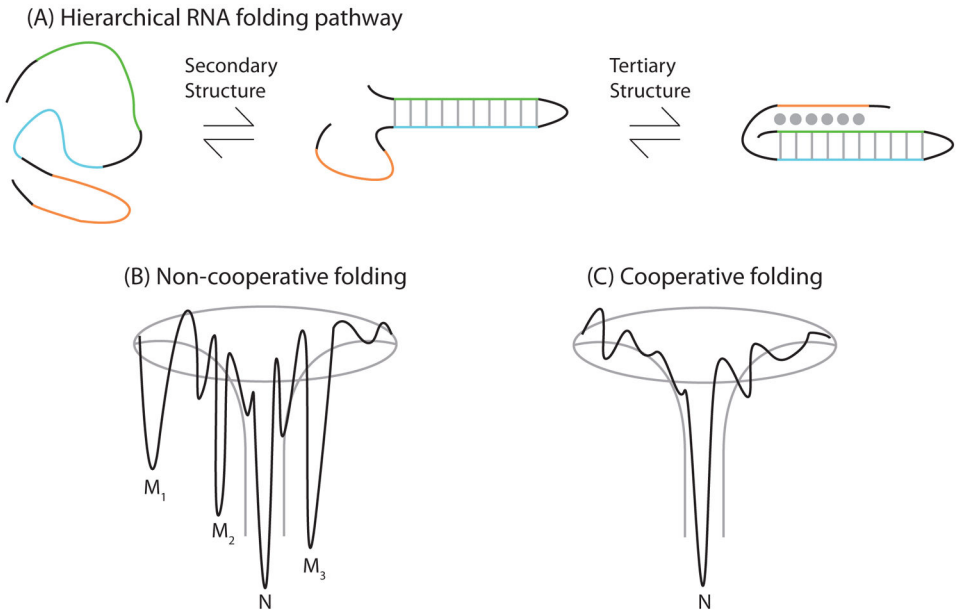


Figure 4.

Depiction of the hierarchical RNA folding pathway, and folding funnels for non-cooperative and cooperative folding. (A) RNA folds in a hierarchical manner in which secondary structures form followed by tertiary structure. Hierarchical folding can be (B) rugged and non-cooperative in which the pathway intermediates are populated and the RNA can form misfolds (M_i) before populating the native state (N), or folding can occur in a (C) cooperative manner in which the intermediates do not populate and the RNA folds in a single transition.

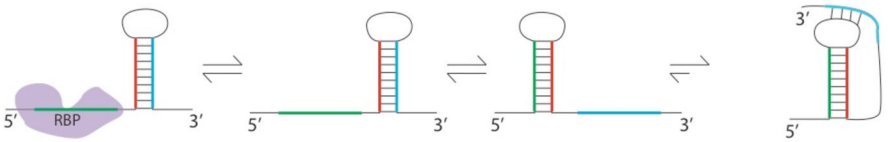
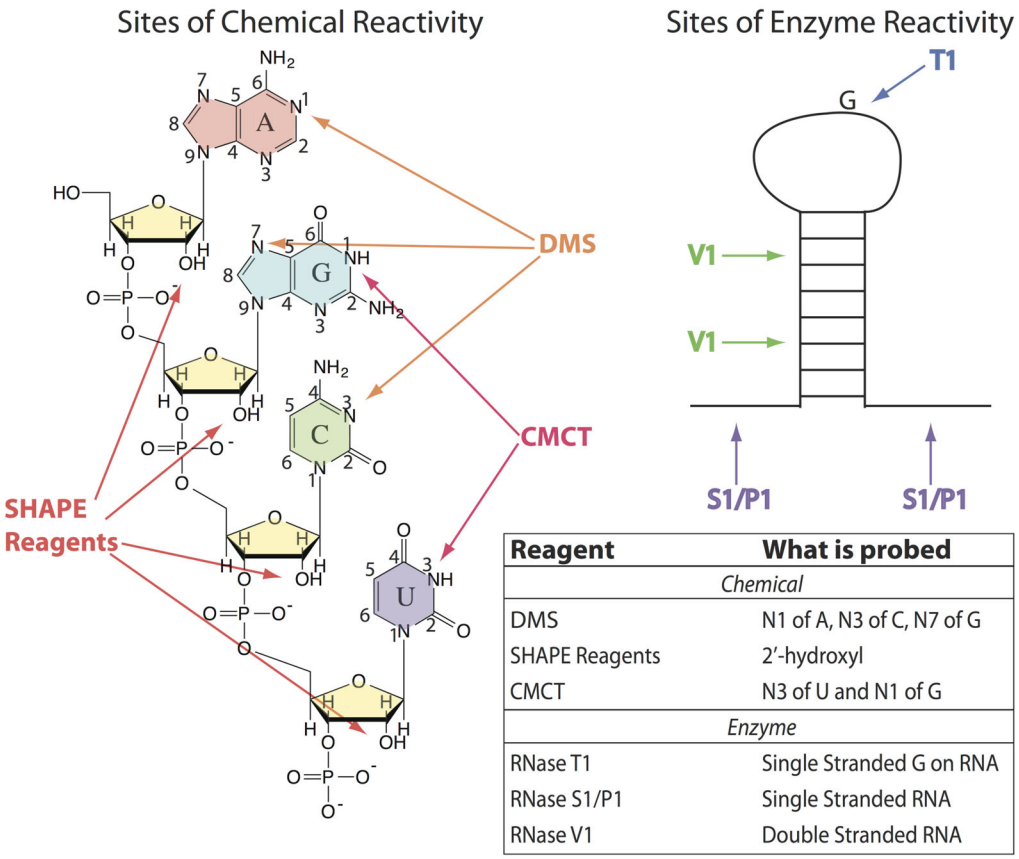


Figure 5.

Different RNA structures can be populated under *in vitro*, *in vivo*, and *in vivo-like* conditions. RNA structures induced by the cellular environment, including proteins and crowding, are shown in the two outermost structures. The conditions *in vitro* favor the population of a structure that may not always be the functional RNA structure (center two structures). Depending on the *in vivo-like* conditions chosen, specific RNA structures will be populated.



Reagent	What is probed
Chemical	
DMS	N1 of A, N3 of C, N7 of G
SHAPE Reagents	2′-hydroxyl
CMCT	N3 of U and N1 of G
Enzyme	
RNase T1	Single Stranded G on RNA
RNase S1/P1	Single Stranded RNA
RNase V1	Double Stranded RNA

Figure 6.

RNA modifications by DMS (dimethyl sulfate), SHAPE reagents (Selective 2′-hydroxyl acylation analyzed by primer extension), and CMCT (1-cyclohexyl-(2-morpholinoehyl)carbodiimide metho-p-toluene). SHAPE reagents modify the 2′-hydroxyl on the sugar of all four nucleobases. SHAPE reagents include 1M7 (1-methyl-7-nitroisatoic anhydride), NMIA (N-methylisotoic anhydride), and NAI (2-methylnicotinic acid imidazolide). DMS modifies the N1 of A and the N3 of C as well as the N7 of G. CMCT modifies N3 of U and N1 of G. The chemical modifications (except N7 of G) can be detected immediately by RT followed by gel electrophoresis or high-throughput sequencing, and the enzymatic cleavages, which modify single- and double-stranded RNA, can be read out through gel electrophoresis or high-throughput sequencing.

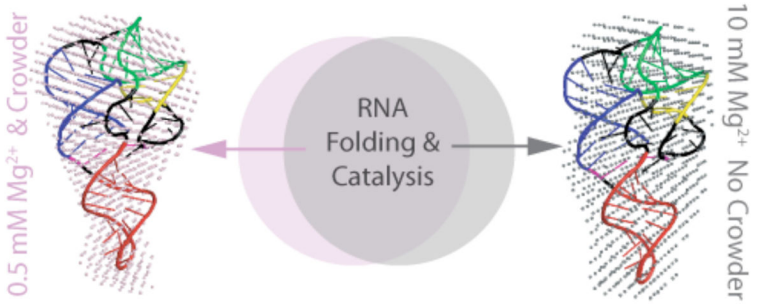


Figure 7.

Under both *in vitro* (right, grey) and *in vivo-like* conditions with molecular crowding (left, pink) RNA fold into their native state that is functional, indicated in this figure by catalysis. High concentrations of Mg^{2+} (10 mM or higher) is needed to achieve the folded state *in vitro* compared to *in vivo-like* crowded conditions where low physiological Mg^{2+} (0.5 mM) folds the RNA. Reprinted with permission from Strulson, C. A., Yennawar, N. H., Rambo, R. P. & Bevilacqua, P. C. (2013). Molecular crowding favors reactivity of a human ribozyme under physiological ionic conditions. Biochemistry, 52, 8187–8197. Copyright 2016 American Chemical Society.

Table 1

Comparison of solution conditions *in vitro* and *in vivo* solution conditions.

Condition	*in vitro* (Historical in the RNA field)	*in vivo*
Molecular Crowding	0%	20–40% (w/w)
Monovalent Salt	0–1M[a]	140 mM K^+[c]
Divalent Salt (free)	0–100 mM[b]	0.5–1.0 mM Mg^{2+} (eukaryotes)[d] 1.5–3.0 mM Mg^{2+} (prokaryotes)[e]
Divalent Salt (total)	0–100 mM[b]	20 mM
Ionic Strength	0–1 M (monovalent only) 0–0.3 M (divalent only)	0.142–0.143 M (eukaryotes) 0.145–0.149 M (prokaryotes)

Conditions *in vitro* are the conditions historically used to study RNA. Typical values in the literature are listed in the table, although actual values differ across various studies. *In vivo-like* conditions, not provided in this table, typically emulate at least one of the conditions missing during *in vitro* experiments.

[a] Typically Na^+ is used *in vitro* although K^+ is found. (Freier et al., 1986a; Xia et al., 1998)

[b] Typically Mg^{2+} is used. (Herschlag & Cech, 1990; Tanner & Cech, 1996)

[c] (Feig & Uhlenbeck, 1999)

[d] (Alberts et al., 1994; London, 1991; Romani, 2007)

[e] (Lusk et al., 1968; Truong et al., 2013)

Table 2

Common experimental techniques used to study RNA structure and folding.

Method	Perturbation	What is Probed	References
		Low Resolution Methods	
Small angle X-ray scattering	*in vitro*	Overall structure	(Yang, 2014) (Pollack, 2011)
Optical melting	*in vitro*	Thermodynamics/Thermostability	(Schroeder & Turner, 2009)
Stopped-flow, Hand mixing	*in vitro*	Kinetics	
Temperature-jump	*in vitro and in vivo*	Folding Kinetics/Thermodynamics	(Dyer & Brauns, 2009, Gao et al., 2016)
Single molecule FRET	*in vitro*	Overall Structure	(Klostermeier & Millar, 2001; Roy et al., 2008)
Hydroxyl radical footprinting	*in vitro and in vivo*	Structure context of nucleotides	
Enzymatic mapping	*in vitro*	Structure context of nucleotides	(Wan et al., 2011; Clatterbuck Soper, S. F. et al., 2013)
Chemical mapping	*in vitro* and *in vivo*	Structure context of nucleotides	(Kwok et al., 2015; Weeks, 2010)
		High Resolution Methods	
X-ray crystallography	*in vitro*	Å-Resolution Structure	(Reyes et al., 2009)
Nuclear magnetic resonance	*in vitro*	Å-Resolution Structure	