Published in final edited form as:

Methods Enzymol. 2015 ; 558: 3–37. doi:10.1016/bs.mie.2015.01.008.

Native Purification and Analysis of Long RNAs

Isabel Chillón^{*,†}, Marco Marcia^{*,1}, Michal Legiewicz^{*,2}, Fei Liu^{*}, Srinivas Somarowthu^{*}, and Anna Marie Pyle^{*,†,‡,3}

^{*}Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, Connecticut, USA

[†]Howard Hughes Medical Institute, Chevy Chase, Maryland, USA

[‡]Department of Chemistry, Yale University, New Haven, Connecticut, USA

Abstract

The purification and analysis of long noncoding RNAs (lncRNAs) *in vitro* is a challenge, particularly if one wants to preserve elements of functional structure. Here, we describe a method for purifying lncRNAs that preserves the cotranscriptionally derived structure. The protocol avoids the misfolding that can occur during denaturation–renaturation protocols, thus facilitating the folding of long RNAs to a native-like state. This method is simple and does not require addition of tags to the RNA or the use of affinity columns. lncRNAs purified using this type of native purification protocol are amenable to biochemical and biophysical analysis. Here, we describe how to study lncRNA global compaction in the presence of divalent ions at equilibrium using sedimentation velocity analytical ultracentrifugation and analytical size-exclusion chromatography as well as how to use these uniform RNA species to determine robust lncRNA secondary structure maps by chemical probing techniques like selective 2'-hydroxyl acylation analyzed by primer extension and dimethyl sulfate probing.

1. INTRODUCTION

Biochemical and biophysical studies on RNA molecules *in vitro* require that target molecules are synthesized and purified as homogeneous, functional species. For catalytic RNAs, enzymatic assays can confirm that the target molecules are correctly folded (Russell et al., 2002; Wan, Mitchell, & Russell, 2009). For noncatalytic RNAs, folding homogeneity needs to be assessed by nonenzymatic, biochemical, or biophysical assays (Woodson, 2011).

Traditionally, two folding procedures have been successfully used in biochemical and biophysical studies of large RNAs, such as group I and II introns, other ribozymes, riboswitches, signal recognition particle RNA, and tRNAs. The first approach involves heat denaturation and refolding of the RNA after *in vitro* transcription (Fedorova, Waldsich, & Pyle, 2007; Lomant & Fresco, 1975; Takamoto, He, Morris, Chance, & Brenowitz, 2002;

Walstrum & Uhlenbeck, 1990; Zhang & Ferre-D'Amare, 2013), whereas the second approach consists of cotranscriptional folding without any denaturation step (Batey, 2014; Toor, Keating, Taylor, & Pyle, 2008). The latter method preserves the secondary and/or tertiary structure adopted by the RNA during transcription, which can be important considering that functional RNA structures are formed cotranscriptionally *in vivo* (Frieda & Block, 2012; Heilman-Miller & Woodson, 2003; Lai, Proctor, & Meyer, 2013).

Long noncoding RNAs (lncRNAs) are involved in a staggering diversity of fundamental cellular functions and they represent important subjects for ongoing research (Gutschner & Diederichs, 2012). Most lncRNAs do not have ribozyme activity, but they are essential in development, transcription, and epigenetic processes (Necsulea et al., 2014). These RNAs, which can reach tens of kilobases in length, appear to have encountered weaker evolutionary selection constraints than protein-coding genes, thus accumulating repetitive sequences (Derrien et al., 2012). lncRNAs are therefore challenging molecules for biophysical analysis because they can form alternative conformations in the absence of any structural constraint (Huthoff & Berkhout, 2002). Given these properties, cotranscriptional native purification may be particularly useful for lncRNA targets.

Here, we describe a native purification protocol that results in pure, homogeneous lncRNA preparations amenable for biochemical and biophysical studies. While other purification methods make use of affinity tags suitable to extract RNA from *in vivo* sources (Batey & Kieft, 2007; Said et al., 2009), our protocol does not necessarily require tags. The use of tags, generally added to the 3' end of target RNAs, ensures capturing a homogeneous population of full-length molecules. In our method, we achieve a similar goal by using a T7 polymerase construct that rarely produces short abortive transcripts (Tang et al., 2014), and centrifugal filtration and size-exclusion chromatography (SEC) as final polishing steps in purification. Not using tags may simplify cloning design and avoid inclusion of nonnative sequences that may interfere with structure formation of the target RNA. However, if tags are useful for downstream applications, their inclusion is compatible with our protocol.

We additionally describe methods to study lncRNA folding based on sedimentation velocity analytical ultracentrifugation (SV-AUC), analytical SEC, and chemical probing. These analytical techniques are provided as examples, as many other techniques (such circular dichroism, small angle X-ray scattering, etc.) can be used to monitor homogeneity and oligomeric state of long RNAs (Behrouzi, Roh, Kilburn, Briber, & Woodson, 2012; Pan & Sosnick, 1997; Rambo & Tainer, 2010). SV-AUC and analytical SEC allow one to monitor global compaction of RNA preparations in the presence of divalent ions, under equilibrium conditions (Cole, Lary, Moody, & Laue, 2008; Mitra, 2014), and the equipment required is commonly available. Chemical probing facilitates the determination of lncRNA secondary structure (Athavale et al., 2013; Novikova, Hennelly, & Sanbonmatsu, 2012), and it also utilizes reagents that are available to most investigators. In this review, we describe protocols for selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE) and dimethyl sulfate (DMS) chemical probing, as they have been applied for studying lncRNAs (Novikova et al., 2012; Watts et al., 2009). Again, these represent only a subset of available techniques for mapping RNA structure in solution. The SHAPE and DMS methods employ an electrophilic reagent that reacts selectively with flexible, accessible sites on

ribonucleotides, facilitating detection of loops and other single-stranded regions within lncRNA molecules.

2. NATIVE PURIFICATION OF LONG NONCODING RNAs

In this section, we outline the procedures for transcription and purification of lncRNAs following nondenaturing methods. These protocols are based on those developed in our laboratory in recent years (Fedorova, Su, & Pyle, 2002; Fedorova et al., 2007; Marcia & Pyle, 2012; Toor et al., 2008), which have been modified and updated to fit the idiosyncrasies of large non-coding RNAs that range from one to several kilobases in length (up to 4 kb in our hands). The pipeline of the procedure includes the design of the construct, linearization of the DNA template and transcription of the RNA, digestion of DNA, enzyme proteolysis, buffer exchange, and finally, filtration and purification using SEC (Fig. 1A).

2.1 Construct design

The target lncRNA should be inserted into a suitable vector (e.g., pBluescript, Life Technologies) such that it is immediately flanked at the 5' end by T7 promoter sequence and at the 3' end by a unique and efficient restriction site (e.g., *Bam*HI). In addition, we recommend digesting the template DNA with a restriction enzyme that produces blunt or 5' overhanging ends to limit the production of heterogeneous RNA transcripts that vary in length (Schenborn & Mierendorf, 1985).

2.2 DNA plasmid linearization

1. Mix 50 µg of DNA plasmid with 80 U of the appropriate restriction enzyme in the presence of 1× restriction enzyme buffer in a final volume of 100 µl.
2. Incubate in a thermo block or mixer incubator at 37 °C overnight.
Note: It is recommended to check that the plasmid is completely digested by running an agarose gel before starting the transcription reaction.
3. Purify the plasmid through phenol extraction followed by ethanol precipitation or through a spin column (Illustra MicroSpin G-25 column, GE Healthcare).
4. Quantify the concentration of the DNA template in a Nanodrop spectrophotometer (Thermo Scientific).

2.3 *In vitro* transcription of RNA

1. Prepare a transcription reaction in a final volume of 200 µl (Tables 1 and 2).
 - a. We express our own batches of recombinant T7 RNA polymerase, which has a concentration of 2 mg/ml, approximately. We use the T7 RNA polymerase mutant P266L, which is known to produce short abortive transcripts less frequently than wild-type T7 polymerase (Tang et al., 2014). If using a commercial T7 RNA polymerase, we suggest following the manufacturer instructions.

After proteolytic digestion, divalent ions are chelated through the addition of EDTA to the solution. This treatment disrupts the RNA tertiary structure, but preserves stable secondary structure elements formed during the transcription reaction.

1. Add an amount of EDTA to the reaction that is equivalent to the number of divalent cations present in the transcription buffer.
 - a. The reaction mixture described above contains 2400 nmol of MgCl_2 . Therefore, add:

$$2400 \text{ nmol} : 500 \text{ nmol}/\mu\text{l of EDTA} = 4.8 \mu\text{l of EDTA } 0.5 \text{ M}$$

2. Vortex briefly and spin down the tubes.

2.7 Buffer exchange and purification

Buffer exchange and separation of the RNA from enzymatic products is performed using commercially available Amicon® Ultra Centrifugal Filters (Millipore) (<http://www.emdmillipore.com/US/en/life-science-research/protein-sample-preparation/protein-concentration/amicon-ultra-centrifugal-filters/YKSb.qB.GXgAAAFBTrllvyxi.nav>). This treatment allows the RNA to be equilibrated in filtration buffer (Table 4), a buffer solution containing physiological amounts of monovalent ions that preserve the secondary structure, while washing out unreacted nucleotides, digested DNA, protein fragments, and other components of the previous reactions.

Short macromolecules, such as abortive transcription products and the proteinase K enzyme (18.5 kDa) are also generally filtered out, although the extent of the filtration depends on the cut-off of the Amicon filter selected. All the centrifugation steps should be carried out at room temperature. It is important to keep the RNA at room temperature at all times, since changes in temperature can trap RNA molecules in alternative conformations, decreasing the homogeneity of the sample. The size of the concentration device and the centrifugation conditions depend upon the final volume of the transcription reaction (Table 5).

The Amicon® Ultra Filters are manufactured in a range of molecular weight cut-offs (MWCO), which have been calibrated for globular proteins. We use those with a MWCO of 100 kDa for our lncRNAs, which range from 2 to 4 kb long, although the selected MWCO will vary according to the target of the study. In general, it is important to consider that selectivity of Amicon® Ultra Filters is different for proteins and for RNAs. For example, a filter that retains globular proteins of a given molecular weight will retain RNA molecules that are six times smaller (Kim, McKenna, Viani Puglisi, & Puglisi, 2007).

2.8 Size-exclusion chromatography

SEC or gel filtration chromatography is used for native purification of the RNA in a homogenous form. Alternative conformations of the RNA that are induced by temperature fluctuations, multimeric forms of the RNA formed by self-complementarity and aggregation and prematurely terminated transcripts can be resolved by SEC without altering the native secondary structure of the lncRNA (Fig. 1B).

We use different bead size resins that we pack into a Tricorn 10/300 Empty High Performance Column (GE Healthcare) (CV of approximately 25 ml). The media are selected according to the size of the RNA, following GE Healthcare recommendations (<http://www.gelifesciences.co.jp/catalog/pdf/18112419.pdf>) (Table 6).

All the procedures are carried out at room temperature, and empty columns should be packed at room temperature, following manufacturer recommendations (<http://ge-catalog.campaignhosting.se/Documents/Tricorn/TricornInstructionsAC.pdf>). We keep packed columns equilibrated with ddH₂O or 20% ethanol at 4 °C when they are not being used to avoid bacterial contamination. In addition, all filtration solutions and packed columns should be warmed up to room temperature before the run.

Use an appropriate fast protein liquid chromatography (FPLC) system (i.e., Akta FPLC, Akta Purifier, Akta Explorer, Akta Pure) equipped with wavelength detection at 260 nm. The system should be washed with an RNase decontamination solution like RNaseZap (Life Technologies) and the column washed with at least 3 column volumes (CV) of ddH₂O and equilibrated with 3 CV of filtration buffer.

The following steps should be performed during SEC of lncRNA:

1. Collect fractions of 0.5 ml volume. The homogeneity of the preparation will usually be apparent from the elution chromatogram (Fig. 1B). If the RNA of interest is homogenous and elutes as a single peak, this fraction should be collected.

Note: Self-packed columns and RNA preparations usually lead to broader peaks than those generated by protein preparations that are run in commercially packed and calibrated columns.

2. Measure the concentration of each fraction using a Nanodrop spectrophotometer (Thermo Scientific) and record absorption value at 260 nm.
3. The fraction with the highest absorption should be used for further studies. If more material is necessary, neighboring fractions should be combined and the absorption value of the mixture recorded again.

At this point, the lncRNA is used for subsequent applications, like folding studies or secondary structure probing. If long-term storage is necessary, the preservation process should be validated to be proven that it does not affect RNA biochemical and biophysical properties (e.g., formation of nonspecific interactions, multimerization, structural rearrangements, precipitation, or degradation). Based on our experience, storage of long RNAs in cold impaired the reproducibility of our assays. For that reason, we work exclusively on freshly purified RNA.

3. STUDY OF THE RNA TERTIARY FOLDING BY SEDIMENTATION VELOCITY ANALYTICAL ULTRACENTRIFUGATION

In this section, we describe the preparation of the sample for an equilibrium SV-AUC assay, and we provide a brief description of the assembly of the analytical cells and the instrument, the ProteomeLabTM XL-I analytical ultracentrifuge (Beckman Instruments). We also explain

how to process the raw data in order to obtain the hydrodynamic parameters for a given RNA species. There are two basic parameters to consider when preparing an SV-AUC experiment:

The first parameter is the concentration of the sample. We typically determine the concentration of the RNA in terms of absorbance at 260 nm in the Nanodrop (Thermo Scientific). The range for maintaining a linear response is from 0.2 to 1.2 absorbance units. High sample concentrations are more prone to aggregation and can lead to artifacts; in addition, different preparations of RNA samples can vary in yield and using high concentrations of sample in every experiment reduces the number of assays that can be performed with the same RNA preparation. For these reasons, we recommend using the lowest concentration of RNA that provides a good signal. Generally, an absorbance of 0.2 at 260 nm is sufficient.

The second parameter is the rotor speed. This parameter needs to be adjusted based on the molecular weight and the expected hydrodynamic radius of the target molecule. We typically apply a speed ranging between 20,000 and 25,000 rpm for our lncRNAs. For further reference, Cole et al. (2008) provides a quick guide for selecting the rotor speed based on the molecular weight of the molecule (although the values he provides refer to globular proteins), and the estimated sedimentation coefficient of the particles, which can be assessed in a preliminary test.

3.1 Preparation of samples for a study of RNA folding

1. Take the fraction with the highest absorbance at 260 nm after the SEC step and write down the absorbance value.
2. Determine the number of samples that you can run by SV-AUC by dividing the absorbance value of the sample by the final absorbance value that you will use in the experiment, for example, 0.2. Following these recommendation, we usually have material to perform at least 12 titration measurements. If the concentration of the RNA sample is not enough for making the desired number of measurements, then mix several fractions. If there is insufficient material in the main fraction for your experiment, we recommend taking only the two fractions eluting immediately before and immediately after the peak.
3. Set up the folding reaction in a volume of 420–500 μ l. Dilute the RNA up to the desired concentration (i.e., 0.2) and add 2 $\times$ AUC buffer (Table 7).
4. For each titration point, you should prepare a blank control, which does not contain RNA. The blank control contains the same volume of filtration buffer than RNA sample and the same concentration of the other components (Table 8).
5. Incubate the folding reaction at 37 °C on a heat block for a period of time that allows the RNA to reach its equilibrium structure. We usually let the RNA equilibrate for 30 min to 1 h. At this point, the samples are ready to be loaded in the optical cells of the analytical ultracentrifuge.

3.2 Assembly of the optical cells, sample loading, and instrument setup

Each analytical cell consists of a centerpiece, two window assemblies, and a cell housing. We use the charcoal-filled Epon double-sector centerpiece by Beckman and sapphire windows because they are harder than quartz windows and so they have longer life. The cells are loaded into an An-60 Ti analytical rotor (Beckman Instruments).

1. Clean all components with RNaseZAP and distilled water, especially the centerpieces and the windows, which are in direct contact with the RNA sample.
2. Assemble the cells and the counterbalance following the manufacturer instructions (<https://www.beckmancoulter.com/wsrportal/techdocs?docname=LXLA-TB-003>).
3. Place the cell horizontally, facing the screw ring, and load 400 μ l of blank sample in the left cavity and 380 μ l of RNA-containing sample in the right cavity with a 500 μ l Hamilton syringe. Repeat this procedure for the other two cells.

Note: It is important that the meniscus of the blank sample is higher than that of the RNA sample so that it does not interfere with the data analysis (see Section 3.4).

4. Weigh the filled cells and make sure that opposing cells weigh within 0.5 g of each other.
5. Seal the filling holes with the red plug gaskets and screw the housing plugs in with a screwdriver.
6. Switch on the AUC instrument and place the rotor inside. Place the monochromator inside the ultracentrifuge as described by the manufacturer (<https://www.beckmancoulter.com/wsrportal/techdocs?docname=LXLAI-IM-10>). Activate the chamber vacuum by pressing the vacuum key on the instrument control panel.

Note: The monochromator contains a blocking filter that can be used to eliminate wavelengths below 400 nm. For absorbance at 260 nm, position the monochromator blocking filter lever perpendicular to the stem.

3.3 Setting up a sedimentation velocity experiment

Here, we describe the parameters for a sedimentation velocity run of an lncRNA in the range of 2–3 kb in length. These parameters should be adapted for different targets under study, following sample concentration and rotor speed recommendations above to ensure complete sedimentation of the RNA molecules.

1. Open the ProteomeLab XL-A/XL-I User Interface Software (Beckman Instruments).
2. Click File>New File.
3. In the setup area, select the following values and options:
 - a. Rotor: 4 hole
 - b. Speed: 25,000 rpm
 - c. Time: hold

- d. Temperature: 20 °C
 - e. Type of scan: Velocity
 - f. Optical system: Absorbance
 - g. Use default values of $R_{\min}$ and $R_{\max}$, and select 260 nm in the wavelength field.
 - h. If you want to use the same scan settings for each cell, enter settings for Cell 1 and then click the All Settings Identical to Cell 1 box.
4. Click the Detail button and select the following options in the Velocity Detail dialog box:
 - a. Replicates: 2
 - b. Mode: Continuous
 - c. Centerpiece: 2
 - d. Accept the default values for the rest of the fields
5. In the setup area, click on Scan Options and select the following parameters:
 - a. Leave all boxes unselected except for Stop XL after last scan.
 - b. Select Overlay last 1 scan(s).
6. Click the Method button. Accept the default values in each field and select 100 scans in the number of scans field.
7. We usually perform a short- and low-speed method for the radial calibration of the instrument before performing the actual experiment. The following modifications to the setup parameters described above should be selected for the radial calibration run:
 - a. In the setup area, change the following options:
 - i. Speed: 3000 rpm
 - b. Under Details, select the following options:
 - i. Replicates: 1
 - c. In Scan options, select the following box in addition to the previous ones:
 - i. Radial calibration before first scan.
 - d. In Method, select 1 scan.
8. Once the radial calibration is finished (around 7–10 min), let the sample stand inside the instrument for around 30 min to 1 h, to allow for temperature equilibration. After that, start the data acquisition with the Start Method Scan button.
9. When the run is complete, turn off the vacuum and take the monochromator and the rotor out. Remove the sample from the cells with a pipette and gel-loading tips and

disassemble the optical cells. Clean them well with Beckman soap and distilled water, and place them back.

3.4 Data analysis

We use the software SEDFIT for data analysis (<http://www.analyticalultracentrifugation.com/download.htm>) (Schuck, 2000). SEDFIT is a method for size-distribution analysis that uses numerical solutions to the Lamm equations (Lebowitz, Lewis, & Schuck, 2002; Fig. 2A and B). SED-FIT also calculates hydrodynamic parameters like the Stokes radius (hydration radius) of the species sedimented in the presence of different concentrations of magnesium. The values of Stokes radius are then fit to the Hill equation and the value of the K_d or midpoint of the equilibrium folding transition is determined (Fig. 2C; Mitra, 2009, 2014).

1. Click on Data and Load new files. Select all the data corresponding to one optical cell (RA1, RA2, or RA3). In the pop-up dialog, enter the loading interval, preferably 1, if the number of scans is not very high (around 100 is good).
2. In the main software interface, drag and drop the red line in the absorbance display to indicate tentatively the position of the sample meniscus, generally located at about 6 cm and the blue line to indicate the bottom of the cell, which is usually located at around 7.2 cm.
3. In the same view, drag and drop the green lines, which indicate the limits of the analysis. We recommend placing the left green line as close as possible to the red line, whereas we usually place the right green line at around 7 cm, to avoid including convection artifacts in the analysis.
4. In the Model dialog box, select Continuous c(s) with other prior knowledge and in the pull down dialog select Continuous c(s) Conformational Change Model. This model is useful when the molar mass of the main species is known and when this species undergoes a conformational change under the different ionic conditions. In this case, the main species will sediment with different s-values under the different magnesium concentrations.
5. Click on Parameters and enter the following data:
 - a. The molar mass of the RNA under study.
 - b. A tentative range of s-min and s-max to display the results.
 - c. The resolution or number of data points of the analysis, which is generally left as default (100).
 - d. The partial specific volume, which is approximately 0.53 ml/g for RNA molecules (Mitra, 2009).
 - e. The buffer density and the buffer viscosity for the experimental conditions. We use the program SEDNTERP (<http://sednterp.unh.edu/>) to determine the density and viscosity of each buffer used in the magnesium titration. Select 20 °C under the Experiment tab and introduce the components of the buffer

$$Y(R_H) = Y(R_{H,0}) + (Y(R_{H,f}) - Y(R_{H,0})) \times \left(\frac{[Mg]^n}{K_{Mg}^n + [Mg]^n} \right),$$

where R_H is the Stokes radius (in Å), $R_{H,f}$ and $R_{H,0}$ are the Stokes radii for unfolded and folded RNA, which correspond to the titration point with the lower and the highest concentration of magnesium, respectively. K_{Mg} value is the concentration of magnesium at which 50% of the RNA is folded and n is the Hill coefficient, which indicates the cooperativity of the folding transition (Mitra, 2009; Su, Brenowitz, & Pyle, 2003; Fig. 2C).

4. ANALYSIS OF THE RNA TERTIARY FOLDING BY ANALYTICAL SIZE-EXCLUSION CHROMATOGRAPHY

SEC can also be used as an analytical tool to observe global compaction of RNA molecules in the presence of divalent ions. The advantage of this approach lies in its ease of use and in the higher availability of necessary equipment. Specifically, we use the same chromatography system (Akta FPLC) (see Section 2.8.). In analytical SEC there is no need to collect fractions, although it is mandatory that the FPLC system is equipped with a UV cell to monitor absorbance of the RNA sample. A description of this technique is summarized below:

1. Purify the RNA following the procedures described in Section 2.
2. Set up a folding reaction as explained in Section 3.1. Test a broad range of magnesium concentrations to determine the upper limit at which the RNA is no longer compacting, but rather aggregating. This phenomenon is especially easy to detect in analytical SEC, since large aggregates of molecules elute in the void volume of the column (Fig. 2D).
3. For every magnesium concentration tested, the FPLC system must be equilibrated in the same buffer used for the folding of RNA. The equilibration is carried out with 1 CV of buffer, provided that the experiment is starting from the lower magnesium concentration, so that no residual magnesium interferes with the assay. In other cases, equilibration should be performed with 3 CV of buffer.
4. Inject the sample and run the same protocol as that used for the purification of the RNA (Section 2.8). Note that the UV cell of the Akta FPLC system is very sensitive and very low sample volumes can be used for an analytical SEC experiment. In particular, we load 3 µg of RNA per titration condition.
5. A typical elution profile that is representative of a titration experiment is shown in Fig. 2D. A shift in the elution volume at each titration point is indicative of the compaction of the RNA at increasing magnesium concentrations. Compared to an SV-AUC profile (Fig. 2B), the resolution of the shift and that of the decrease in absorbance due to the hypochromicity effect are lower. Nevertheless, analytical SEC offers a reliable method to evaluate homogeneity of the RNA at different magnesium concentrations and to detect aggregation and degradation events.

6. To compare elution profiles between different columns, we calculate the chromatographic partition coefficient, K_{av} , or distribution constant (Fig. 2D inset; Dai, Dubin, & Andersson, 1998):

$$K_{av} = \frac{(V_e - V_0)}{(V_t - V_0)}$$

where V_e is the RNA elution volume, V_0 is the void volume of the column, and V_t is the total volume of the column. In addition, the partition coefficient also provides a quantitative determination of the highest concentration of magnesium at which the RNA ceases to continue compacting. This value also represents the concentration of magnesium that is optimal to include in the folding step of subsequent biophysical and biochemical applications. In the case of HOTAIR lncRNA, we use 25 mM of magnesium (Fig. 2D inset).

5. DETERMINATION OF THE SECONDARY STRUCTURE OF LNCRNAS BY CHEMICAL PROBING

In this section, we will provide an approach for determining the secondary structure of an lncRNA by using two chemical probing techniques: SHAPE and DMS. It is important to note that chemical probing should be performed immediately after the purification step and the samples have to be kept at room temperature at all times, avoiding any cooling down or freezing step, as this promotes the rearrangement of RNA structures. We make use of automated, high-throughput capillary electrophoresis for the analysis of fragments generated after the primer extension reactions, which are carried out with fluorescent-coupled primers. The runs are performed in a campus-located facility (<http://dna-analysis.research.yale.edu/>), and the raw data are analyzed by either one the following software programs: ShapeFinder (Vasa, Guex, Wilkinson, Weeks, & Giddings, 2008) and QuShape (Karabiber, McGinnis, Favorov, & Weeks, 2013).

5.1 Designing and coupling of primers

Fluorescent-coupled primers are used for primer extension detection of modifications in the RNA. Primers should be designed following standard guidelines as for DNA sequencing (approximate length of 20 nucleotides, starting and ending with purine residues, and with an approximate GC content of 50%). In addition, G should be avoided at the 5' end if possible, as it can quench fluorescence (Behlke, Huang, Bogh, Rose, & Devor, 2005). The primers should cover the full sequence of the target gene, spaced evenly every 200–220 nucleotides, approximately. Oligonucleotides are synthesized in 50 nmol scale and modified at the 5' end with an amine group.

We use four spectrally different fluorescent dyes in the form of succinimidyl esters: 5-carboxy-X-rhodamine (5-ROX), 5-carboxyfluorescein (5-FAM), 6-carboxy-4',5'-dichloro-2',7'-dimethoxyfluorescein (6-FAM), and 6-carboxytetramethylrhodamine (6-TAMRA) (Anaspec). Succinimidyl esters (SE) are recommended because they form a very stable amide bond between the dye and the amine-modified oligonucleotide primer. SE dyes

are hydrophobic and they are freshly dissolved in DMSO prior to the labeling reaction. We have adapted the protocol for labeling amine-modified oligonucleotides from Life Technologies (<http://tools.lifetechnologies.com/content/sfs/manuals/mp00143.pdf>). A summary of the dye conjugation step is provided below:

1. Prepare a labeling reaction by mixing the amine-modified oligonucleotide with the dye (Table 10).
2. Incubate the reaction in the dark for 6 h to overnight at room temperature.
3. Ethanol precipitate the oligonucleotide to ensure removal of the free dye.
4. Purify the oligonucleotide in a denaturing 20% polyacrylamide gel.
5. Locate the labeled and unlabeled oligonucleotides by UV shadowing and cut out the labeled oligonucleotide.
6. Purify the oligonucleotide by the “crush-and-soak” method using 10 mM Tris (pH 8.0) and 300 mM NaCl and precipitate with ethanol. Oligonucleotides are finally dissolved in ddH₂O at a concentration of 20 μM (stock) and 2 μM (working concentration).

5.2 Generation of sequencing ladders

Two sequencing ladder reactions are generated for assays to be analyzed by ShapeFinder and one or two for those to be analyzed by QuShape. In the first case, primers labeled with ROX and TAMRA are reserved for the sequencing ladders. In the second case, we usually prepare a single sequencing ladder with primers labeled with FAM. The nucleotide selected for the ladder varies with every target composition, and it is recommended to make ladder(s) complementary to the most frequent nucleotide(s). Here, we present an example for a ladder performed to detect the adenosines in the target sequence.

1. Combine the following amounts in an amber PCR tube:
28.8 μl Master mix (Table 11)
4 μl Labeled primer 2 μM
2. Place each PCR tube, corresponding to each sequencing ladder, into a thermocycler and run a sequencing-ladder PCR program (Table 13).
3. Stop the reaction by adding 6.4 μl of stop mixture (Table 14).
4. At this point, mix both sequencing ladders if using two for the analysis. Add 196 μl of 100% ethanol and incubate at 4 °C for 5 min.
5. Spin at maximum speed for 30 min and wash two times with 70% ethanol.
6. Resuspend the pellet in 2 μl of water and 38 μl of formamide.
7. Store at −20 °C.

5.3 SHAPE reaction

The RNA purified following the procedure described in Section 2 is subjected to chemical modification using 1-methyl-7-nitroisatoic anhydride (1M7). We synthesize 1M7 in our laboratory following the protocol by Kevin Weeks and coworkers (Mortimer & Weeks, 2007), but there are other approaches and SHAPE reagents that can be employed, like NMIA and 1M6 (Rice, Leonard, & Weeks, 2014). The optimal amount of 1M7 should be titrated for every RNA target under study, since factors like the length of the molecule and the purity of each batch of reagent will influence the extent of the modification. As a starting point, use a final concentration of 1 mM 1M7 dissolved in DMSO.

1. Prepare the folding reaction (Table 15). Reactions are carried out in triplicate. For each replicate, prepare two tubes that will contain the folded RNA sample and one of the following: the reagent (1M7) for the (+) reagent sample or the solvent of the reagent (DMSO) for the (–) reagent sample. The latter sample will be used as a reference for the data analysis because it indicates reverse transcriptase (RT) stops that are not attributable to specific modifications of the 1M7 reagent in the molecule, but rather to RNA degradation or to premature termination of reverse transcription.
2. Fold the RNA by incubating at 37 °C for 45 min. The incubation time will vary depending on the target RNA.
3. Add 1M7 or DMSO to the tubes containing the folded RNA (Table 17).
4. Incubate the reaction at 37 °C for 5 min.
5. Add 320 µl of quench mixture (Table 18) and 1 volume of isopropanol.
6. After quenching the reaction mixture, samples are incubated on ice for an hour or overnight.
7. Spin the samples at maximum speed for 30 min and wash two times with 70% ethanol.
8. Dissolve the pellet in 50 µl of RNA storage buffer (Table 19) and proceed immediately with the primer extension analysis (Section 5.5) without freezing the samples.

5.4 DMS reaction

The second chemical modifier that we use in our laboratory for probing is DMS (Sigma-Aldrich, 99.8% purity). The stock of DMS is prepared by dissolving 5 µl of DMS in 645 µl of cold 100% ethanol (81.3 mM). The final concentration of DMS in reaction has to be titrated for each target RNA molecule, usually ranging from 1 to 15 mM. Here, we describe a DMS probing reaction optimized for HOTAIR lncRNA.

1. The folding reaction takes place using the same volumes as the SHAPE reaction. However, HEPES buffer reacts with DMS and so we use cacodylate buffer in the 5× monovalent ion mixture (Table 20).
2. Prepare a DMS reaction (Table 21).

3. Incubate the reaction at room temperature for 10 min.
4. Stop the reaction by adding 54.4 µl of quench mixture (Table 22).
5. Precipitate the RNA (Table 23).
6. Spin at maximum speed for 30 min and wash two times with 70% ethanol.
7. Resuspend the pellet in 50 µl of RNA storage buffer (Table 19).
8. Store at −80 °C or use in the primer extension reaction.

5.5 Primer extension reaction

The primer extension reaction is performed using the Superscript III RT (Life Technologies) to detect the specific stops of the enzyme due to the modifications in the RNA. The primers used in the reaction are labeled with fluorescent dyes. For experiments that will be analyzed with ShapeFinder, two different dyes, JOE and FAM, are used for the (+) and (−) reagent samples, respectively. For experiments that will be analyzed with QuShape, we use the same labeled primer for the reagent and the blank sample, in this case JOE because the samples will be loaded in different wells (see Section 5.8.). Here, we show an example where the (+) and the (−) reagent samples will be reverse transcribed using primers labeled with JOE and FAM.

1. Prepare an annealing reaction in amber Eppendorf tubes, protected from light (Table 24).
2. Heat the mixture at 95 °C for 2 min and then place the tubes on ice for 5 min. Incubate at 48 °C for 2 min.
3. Add 8 µl of RT mixture (Table 25) and incubate at 48 °C for 45 min.
4. Stop the reaction by adding 30 µl of the stop mixture (Table 26).
5. Precipitate by adding 1 volume (50 µl) of isopropanol.
6. At this moment, the reagent and the blank samples that have been reverse transcribed with different fluorophore-labeled primers are combined together (only for the data analysis in ShapeFinder). If the reagent and the blank samples are reverse transcribed with the same fluorophore-labeled primer (for the data analysis in QuShape), they are not combined.
7. Spin at maximum speed for 30 min and wash two times with ethanol 70%.
8. Resuspend the pellet in 2 µl of deionized water and 48 µl of deionized formamide. Dissolve pellet by heating at 65 °C for 2 min.

5.6 Reactions for mobility shift correction

A mobility shift correction must be performed manually during data analysis with ShapeFinder for each primer set because the dyes alter the electrophoretic migration rates, and therefore cDNAs of the same length may have slightly different elution times. To generate a file for mobility shift correction, we prepare sequencing ladders as described in Section 5.2. The sequencing ladder reactions for mobility shift correction are performed

with all the different labeled primers on the same DNA template and with the same type of ddNTP. Once primer extension reactions are complete, all the reactions with the different fluorophores are mixed and precipitated together, following the protocol described in Section 5.2. The pellet obtained is resuspended in 2 µl water and 38 µl formamide and dissolved at 65 °C in a heat block for 2 min. Note that if you change capillary electrophoresis conditions, a new mobility shift correction will be necessary. To avoid it, we recommend keeping the settings of the capillary instruments fixed.

5.7 Spectral calibration of the instrument

Before the first run on the instrument, a spectral calibration should be carried out with the dye set. A spectral calibration creates a matrix that corrects for the overlapping fluorescence emission spectra of the dyes used for coupling the primers. We calibrate an Applied Biosystems 3730xl DNA Genetic Analyzer for using it with the Applied Biosystems four-dye chemistry DS-20 Dye Set (5-FAM, JOE, TAMRA, ROX) (A filter set). The calibration standard is prepared as previously described (Watts et al., 2009). Briefly, four primer extension reactions are performed from primers labeled with the chosen dye set and located in different regions of a linearized control plasmid, so that four fluorescent cDNA fragments of different length are generated. The resultant primer extensions are then mixed together to obtain similar fluorescence intensity for all four dyes and loaded in the plate:

1. Linearize a pUC18 plasmid with restriction enzyme *HindIII*.
2. Label four primers complementary to pUC18 plasmid (Table 27) with the set of dyes used to calibrate the instrument, following the instructions in Section 5.1.
3. Prepare a PCR reaction for each labeled primer (Table 28).
4. Place the tubes in a thermocycler and run the spectral-calibration PCR program (Table 29).
5. Stop the primer extension reaction by adding 8 µl of stop mixture per reaction (Table 30).
6. Add 120 µl of 100% ethanol per primer extension reaction.
7. Spin at maximum speed for 30 min and wash with 70% ethanol.
8. Dissolve each pellet in 20 µl of deionized formamide and mix the four reactions together.

5.8 Preparation of samples for capillary electrophoresis

Samples are mixed with deionized formamide and loaded in a 96-well plate according to the following guidelines, which can be modified to adjust the output signal:

1. Plates prepared for data analysis with ShapeFinder contain the (+) and the (–) reagent sample mix, together with the sequencing ladder mix, in a single well for every primer extension reaction that is carried out (Table 31).
2. Plates prepared for data analysis on QuShape contain the (+) reagent sample and a single sequencing ladder in one well whereas the (–) reagent sample and the same

sequencing ladder are loaded in a different well. The advantage of this setup is that it allows the use of only two dyes (Table 32).

3. Mobility shift reactions are loaded in the same plate (Table 33).
4. Capillary electrophoresis is performed using an Applied Biosystems 3730x1 DNA Genetic Analyzer with 96 capillaries equipped with a 50-cm capillary array (DNA Analysis Facility at Science Hill, Yale University). Runs were performed while maintaining an oven temperature of 63 °C and a run voltage of 15 kV. The instrument was prerun for 180 s before injecting the sample at 1.2 kV for 16 s. Note that very high signal should be avoided as signal cross-talk can occur despite proper spectral precalibration of the instrument.

5.9 Data analysis

The analysis of chemical probing data consists of the following steps: (i) processing of the capillary electrophoresis electropherograms to determine the absolute reactivities for each reaction, (ii) normalization of the reactivities, and (iii) use of the reactivity profiles obtained as constraints in a software program to deduce a secondary structure of the molecule (Fig. 3).

5.9.1 Determination of chemical probing reactivity profiles—The determination of chemical probing reactivity profiles can be carried out using either ShapeFinder (Vasa et al., 2008) or QuShape (Karabiber et al., 2013) software. We have used both of them successfully in our laboratory. The first one has been more extensively used, although it is no longer available online for download. The second one has been released more recently; it offers enhanced flexibility for utilizing multiple different fluorophores and it does not require import of mobility shift correction files. Regardless of the software used, the output containing the absolute nucleotide reactivity of the RNA sequence is presented as a file containing the calculated peak areas as a function of nucleotide position for the entire analyzed trace.

5.9.2 Normalization of SHAPE and DMS reactivity profiles—We normalize SHAPE and DMS data to a scale in which values ranging from 0 to 0.3 are considered not reactive; from 0.3 to 0.8, moderately reactive; and higher than 0.8, very reactive (Fig. 3A). It is important to note that the DMS reagent has different affinities for adenosine and cytosine residues and so the normalization of their reactivity profiles has to be done separately. The method described below has been modified from that previously reported by McGinnis, Duncan, and Weeks (2009):

1. Open the tab-delimited text file generated after alignment and integration in Microsoft Excel.
2. Calculate the third quartile value for the RX.area-BG.area column using the Excel function “=QUARTILE(array,quart),” where *array* is the cell range of numeric values for which you want the quartile value and *quart* is a number (3 represents the third quartile).

3. Copy the RX.area-BG.area column to a new spreadsheet and order the values in increasing order. Copy to a new column all values that fall in the range of $\pm 10\%$ of the third quartile value.
4. Calculate the average “=AVERAGE(number1, [number2]...)” of this new range of values.
5. Calculate and store the effective maximum reactivity value using the formula “=1.5*(AVERAGE(number1, [number2]...)),” where the range is represented by the values calculated in the point 4.
6. Divide the unsorted (RX.area-BG.area) of each nucleotide by the effective maximum reactivity value to obtain the normalized reactivity values. Values greater than this value are considered outliers and are excluded.
7. Create a new text file comprising two columns: the nucleotide numbers at the left and the normalized reactivity values at the right.
8. Combine all the text files with the normalized reactivity values from all the fragments corresponding to the same RNA.

In our laboratory, we have developed a modification of this method when performing side-by-side comparisons of a wild-type lncRNA and several mutants derived of the same lncRNA. In those cases, we have observed several consistent RT stops in all molecules, independently of the mutant, in both the (+) and (−) reagent samples. This observation is likely due to the inherent flexibility of some positions for each particular type of target RNA. In those cases, our normalization protocol excludes the (−) reagent sample and compares exclusively the (+) reagent samples, thus significantly reducing the observed error in the replicates.

5.9.3 RNA secondary structure prediction and analysis—The normalized reactivity values obtained are used as constraints for the secondary structure prediction of the RNA target using the RNAstructure software (Mathews, 2014; Reuter & Mathews, 2010). The software is available as a Web server or as a program with a graphical user interface, although we prefer the second option to avoid the sequence length restraints of using the Web server. Subsequently, RNA secondary structure coordinates can be visualized and modified using the Java applet VARNA (Darty, Denise, & Ponty, 2009). Additional editing can be also performed with Adobe Illustrator (Adobe Systems) or similar editor for vector graphics.

Acknowledgments

Projects that led to develop the methods described in this chapter were supported by the National Institute of Health (R01GM50313). A. M. P. is an Investigator and I. C. is a Postdoctoral Fellow of the Howard Hughes Medical Institute. We are thankful to all members of the Pyle lab for valuable suggestions.

References

- Athavale SS, Gossett JJ, Bowman JC, Hud NV, Williams LD, Harvey SC. In vitro secondary structure of the genomic RNA of satellite tobacco mosaic virus. *PLoS One*. 2013; 8:e54384. [PubMed: 23349871]

- Batey RT. Advances in methods for native expression and purification of RNA for structural studies. *Current Opinion in Structural Biology*. 2014; 26:1–8. [PubMed: 24607442]
- Batey RT, Kieft JS. Improved native affinity purification of RNA. *RNA*. 2007; 13:1384–1389. [PubMed: 17548432]
- Behlke, MA.; Huang, L.; Bogh, L.; Rose, S.; Devor, EJ. Fluorescence quenching by proximal G-bases. *Integrated DNA Technologies*; 2005.
- Behrouzi R, Roh JH, Kilburn D, Briber RM, Woodson SA. Cooperative tertiary interaction network guides RNA folding. *Cell*. 2012; 149:348–357. [PubMed: 22500801]
- Cole JL, Lary JW, Moody TP, Laue TM. Analytical ultracentrifugation: Sedimentation velocity and sedimentation equilibrium. *Methods in Cell Biology*. 2008; 84:143–179. [PubMed: 17964931]
- Dai H, Dubin P, Andersson T. Permeation of small molecules in aqueous size-exclusion chromatography vis-a-vis models for separation. *Analytical Chemistry*. 1998; 70:1576–1580.
- Darty K, Denise A, Ponty Y. VARNA: Interactive drawing and editing of the RNA secondary structure. *Bioinformatics*. 2009; 25:1974–1975. [PubMed: 19398448]
- Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, et al. The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression. *Genome Research*. 2012; 22:1775–1789. [PubMed: 22955988]
- Fedorova O, Su LJ, Pyle AM. Group II introns: Highly specific endonucleases with modular structures and diverse catalytic functions. *Methods*. 2002; 28:323–335. [PubMed: 12431436]
- Fedorova O, Waldsich C, Pyle AM. Group II intron folding under near-physiological conditions: Collapsing to the near-native state. *Journal of Molecular Biology*. 2007; 366:1099–1114. [PubMed: 17196976]
- Frieda KL, Block SM. Direct observation of cotranscriptional folding in an adenine riboswitch. *Science*. 2012; 338:397–400. [PubMed: 23087247]
- Gutschner T, Diederichs S. The hallmarks of cancer: A long non-coding RNA point of view. *RNA Biology*. 2012; 9:703–719. [PubMed: 22664915]
- Heilman-Miller SL, Woodson SA. Effect of transcription on folding of the tetrahymena ribozyme. *RNA*. 2003; 9:722–733. [PubMed: 12756330]
- Huthoff H, Berkhout B. Multiple secondary structure rearrangements during HIV-1 RNA dimerization. *Biochemistry*. 2002; 41:10439–10445. [PubMed: 12173930]
- Karabiber F, McGinnis JL, Favorov OV, Weeks KM. QuShape: Rapid, accurate, and best-practices quantification of nucleic acid probing information, resolved by capillary electrophoresis. *RNA*. 2013; 19:63–73. [PubMed: 23188808]
- Kim I, McKenna SA, Viani Puglisi E, Puglisi JD. Rapid purification of RNAs using fast performance liquid chromatography (FPLC). *RNA*. 2007; 13:289–294. [PubMed: 17179067]
- Lai D, Proctor JR, Meyer IM. On the importance of cotranscriptional RNA structure formation. *RNA*. 2013; 19:1461–1473. [PubMed: 24131802]
- Lebowitz J, Lewis MS, Schuck P. Modern analytical ultracentrifugation in protein science: A tutorial review. *Protein Science*. 2002; 11:2067–2079. [PubMed: 12192063]
- Lomant AJ, Fresco JR. Structural and energetic consequences of non-complementary base oppositions in nucleic acid helices. *Progress in Nucleic Acid Research and Molecular Biology*. 1975; 15:185–218. [PubMed: 1094495]
- Marcia M, Pyle AM. Visualizing group II intron catalysis through the stages of splicing. *Cell*. 2012; 151:497–507. [PubMed: 23101623]
- Mathews DH. RNA secondary structure analysis using RNAstructure. *Current Protocols in Bioinformatics*. 2014; 46:12.16.1–12.16.25.
- McGinnis JL, Duncan CD, Weeks KM. High-throughput SHAPE and hydroxyl radical analysis of RNA structure and ribonucleoprotein assembly. *Methods in Enzymology*. 2009; 468:67–89. [PubMed: 20946765]
- Mitra S. Using analytical ultracentrifugation (AUC) to measure global conformational changes accompanying equilibrium tertiary folding of RNA molecules. *Methods in Enzymology*. 2009; 469:209–236. [PubMed: 20946791]

- Mitra S. Detecting RNA tertiary folding by sedimentation velocity analytical ultracentrifugation. *Methods in Molecular Biology*. 2014; 1086:265–288. [PubMed: 24136610]
- Mortimer SA, Weeks KM. A fast-acting reagent for accurate analysis of RNA secondary and tertiary structure by SHAPE chemistry. *Journal of the American Chemical Society*. 2007; 129:4144–4145. [PubMed: 17367143]
- Necsulea A, Soumillon M, Warnefors M, Liechti A, Daish T, Zeller U, et al. The evolution of lncRNA repertoires and expression patterns in tetrapods. *Nature*. 2014; 505:635–640. [PubMed: 24463510]
- Novikova IV, Hennelly SP, Sanbonmatsu KY. Structural architecture of the human long non-coding RNA, steroid receptor RNA activator. *Nucleic Acids Research*. 2012; 40:5034–5051. [PubMed: 22362738]
- Pan T, Sosnick TR. Intermediates and kinetic traps in the folding of a large ribozyme revealed by circular dichroism and UV absorbance spectroscopies and catalytic activity. *Nature Structural Biology*. 1997; 4:931–938.
- Rambo RP, Tainer JA. Bridging the solution divide: Comprehensive structural analyses of dynamic RNA, DNA, and protein assemblies by small-angle X-ray scattering. *Current Opinion in Structural Biology*. 2010; 20:128–137. [PubMed: 20097063]
- Reuter JS, Mathews DH. RNAstructure: Software for RNA secondary structure prediction and analysis. *BMC Bioinformatics*. 2010; 11:129. [PubMed: 20230624]
- Rice GM, Leonard CW, Weeks KM. RNA secondary structure modeling at consistent high accuracy using differential SHAPE. *RNA*. 2014; 20:846–854. [PubMed: 24742934]
- Russell R, Zhuang X, Babcock HP, Millett IS, Doniach S, Chu S, et al. Exploring the folding landscape of a structured RNA. *Proceedings of the National Academy of Sciences of the United States of America*. 2002; 99:155–160. [PubMed: 11756689]
- Said N, Rieder R, Hurwitz R, Deckert J, Urlaub H, Vogel J. In vivo expression and purification of aptamer-tagged small RNA regulators. *Nucleic Acids Research*. 2009; 37:e133. [PubMed: 19726584]
- Schenborn ET, Mierendorf RC Jr. A novel transcription property of SP6 and T7 RNA polymerases: Dependence on template structure. *Nucleic Acids Research*. 1985; 13:6223–6236. [PubMed: 2995921]
- Shuck P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophysical Journal*. 2000; 78:1606–1619. [PubMed: 10692345]
- Su LJ, Brenowitz M, Pyle AM. An alternative route for the folding of large RNAs: Apparent two-state folding by a group II intron ribozyme. *Journal of Molecular Biology*. 2003; 334:639–652. [PubMed: 14636593]
- Takamoto K, He Q, Morris S, Chance MR, Brenowitz M. Monovalent cations mediate formation of native tertiary structure of the *Tetrahymena thermophila* ribozyme. *Nature Structural Biology*. 2002; 9:928–933.
- Tang GQ, Nandakumar D, Bandwar RP, Lee KS, Roy R, Ha T, et al. Relaxed rotational and scrunching changes in P266L mutant of T7 RNA polymerase reduce short abortive RNAs while delaying transition into elongation. *PLoS One*. 2014; 9:e91859. [PubMed: 24651161]
- Toor N, Keating KS, Taylor SD, Pyle AM. Crystal structure of a self-spliced group II intron. *Science*. 2008; 320:77–82. [PubMed: 18388288]
- Vasa SM, Guex N, Wilkinson KA, Weeks KM, Giddings MC. ShapeFinder: A software system for high-throughput quantitative analysis of nucleic acid reactivity information resolved by capillary electrophoresis. *RNA*. 2008; 14:1979–1990. [PubMed: 18772246]
- Walstrum SA, Uhlenbeck OC. The self-splicing RNA of *Tetrahymena* is trapped in a less active conformation by gel purification. *Biochemistry*. 1990; 29:10573–10576. [PubMed: 2271667]
- Wan Y, Mitchell D 3rd, Russell R. Catalytic activity as a probe of native RNA folding. *Methods in Enzymology*. 2009; 468:195–218. [PubMed: 20946771]
- Watts JM, Dang KK, Gorelick RJ, Leonard CW, Bess JW Jr, Swanstrom R, et al. Architecture and secondary structure of an entire HIV-1 RNA genome. *Nature*. 2009; 460:711–716. [PubMed: 19661910]
- Woodson SA. RNA folding pathways and the self-assembly of ribosomes. *Accounts of Chemical Research*. 2011; 44:1312–1319. [PubMed: 21714483]

Zhang J, Ferre-D'Amare AR. Co-crystal structure of a T-box riboswitch stem I domain in complex with its cognate tRNA. *Nature*. 2013; 500:363–366. [PubMed: 23892783]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 1.

Enzymatic synthesis and purification of long noncoding RNAs. (A) The native purification pipeline of lncRNAs comprises the following steps (from left to right and top to bottom): the RNA is synthesized with the T7 RNA polymerase system; when the transcription reaction is finished, the DNA template is digested with DNase enzyme; enzymes in the reaction are then proteolyzed with the action of the proteinase K enzyme; RNA is captured from reaction components and buffer exchanged by ultra-filtration with Amicon centrifugal devices (Millipore); finally, the RNA is subjected to size-exclusion chromatography. (B) The resulting chromatogram represents the absorbance of the filtered RNA as a function of the elution volume. Material in the void volume of the column (10 ml elution) and shorter RNA molecules (22 ml elution) are excluded. The fraction(s) corresponding to the main peak are then selected for downstream analysis.

Figure 2.

Study of the RNA tertiary folding by sedimentation velocity analytical ultracentrifugation (SV-AUC) and analytical size-exclusion chromatography (SEC). (A) Raw data from SV-AUC are analyzed using Sedfit program and the resulting fit is visualized using the GUSII applet. (B) The sedimentation coefficient distribution is obtained for each magnesium titration point. The sedimentation coefficient of the main species increases as RNA is compacting in the presence of magnesium. The observed decrease in the absorbance at increasing magnesium concentrations is due to the hypochromicity effect, which occurs when RNA bases stack in the tertiary structure, thus diminishing its capacity to absorb ultraviolet light. (C) Stokes radii are calculated, represented as a function of magnesium concentration and fitted to the Hill equation. (D) Elution volumes in an analytical SEC experiment are plotted as a function of magnesium concentration. The hypochromicity effect is also observed as a decrease in the absorbance of the eluted RNA. At 50 mM magnesium chloride, the sample is mainly eluting in the void volume, which indicates that a nonspecific aggregation is occurring.

Figure 3.

Determination of the secondary structure of lncRNAs by chemical probing. (A) After quantifying the absolute reactivity at each nucleotide position, SHAPE and DMS data is normalized on a scale spanning 0 to 2, approximately. Values lower than 0.3 represent constrained positions, which are part of double-stranded regions of tertiary interactions; values higher than 0.8 represent single-stranded regions; and values between 0.3 and 0.8 are likely single stranded. (B) Normalized reactivity values are used as constraints for RNA secondary structure prediction in RNAstructure software. In the figure, resulting secondary structure values are represented using VARNA software.

Table 1*In vitro* transcription reaction

Component	Final concentration	Stock	Amount
Transcription buffer (Table 2)	1×	10×	20 μ l
DTT	10 mM	1 M	2 μ l
ATP	3.6 mM	30 mM	24 μ l
CTP	3.6 mM	30 mM	24 μ l
GTP	3.6 mM	30 mM	24 μ l
UTP	3.6 mM	30 mM	24 μ l
RNaseOUT™	80 U	40 U/ μ l	2 μ l
Pyrophosphatase	4 U	2 U/ μ l	2 μ l
T7 polymerase	0.1 mg/ml	2 mg/ml	10 μ l
Linearized plasmid	6.25 μ g		31.25 ng/ μ l

Add deionized H₂O to 200 μ l.

Table 2

Transcription buffer (10×)

Component	Final concentration	Stock (<i>M</i>)	Amount (μl)
MgCl ₂	120 mM	1	120
Tris-HCl (pH 8.0)	400 mM	1	400
Spermidine	20 mM	2	10
NaCl	100 mM	5	20
Triton X-100	0.1%		1

Add deionized H₂O to 1 ml.

Table 3

Proteinase K storage buffer (1×)

Component	Final concentration	Stock	Amount
Tris-HCl (pH 7.5)	10 mM	1 M	100 µl
CaCl ₂ (1 M)	1 mM	1 M	10 µl
Glycerol	40%	100%	4 ml

Add deionized H₂O to 10 ml.

Table 4

Filtration buffer (1×)

Component	Final concentration (mM)	Stock (M)	Amount
MOPS buffer (pH 6.5)	8	1	4 ml
KCl	100	2	25 ml
EDTA (pH 8.5)	0.1	0.5	100 µl

Add deionized H₂O to 500 ml.

Table 5

Amicon-device centrifugation guide

Amicon filter	Txn volume (ml)	Rotor	Speed (rcf)	Time (min)
Amicon Ultra 0.5 ml	0.2–1	Fixed-angle rotor suitable for 2-ml tubes	4000	5×5 times
Amicon Ultra 15 ml	1	Swinging bucket rotor suitable for 50-ml tubes	3000	10–15

Table 6

Selection of chromatography media and running conditions

RNA size (nt)	Media	Max. operating flow rate approx. (ml/min)	Max. operating back pressure (MPa)
1–200	Superdex 75	1	0.2
200–600	Superdex 200	1	0.2
600–2000	Sephacryl 400	1	0.2
1000–2000	Sephacryl 500	0.83	0.2
>2000	Sephacryl 1000	0.67	Not determined

Table 7

AUC buffer (2×)

Component	Final concentration (mM)	Stock (M)	Amount
HEPES buffer (pH 7.4)	100	1	15 ml
KCl	400	2	30 ml
EDTA (pH 8.5)	0.2	0.5	60 µl

Add deionized H₂O to 150 ml.

Table 8

Preparation of an SV-AUC experiment

B1	x μ l filtration buffer 250 μ l of $2 \times$ AU buffer 100 μ l of $5 \times$ MgCl ₂ 50 mM x μ l water
S1	x μ l RNA fraction num. X 250 μ l of $2 \times$ AU buffer 100 μ l of $5 \times$ MgCl ₂ 50 mM x μ l water

Example: 10 mM MgCl₂.

VolF = 500 μ l.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 9

Density and viscosity values of selected SV-AUC buffers as calculated by SEDNTERP

MgCl ₂ (mM)	Density (g/l)	Viscosity (P)
0	1.01157	0.01031
0.01	1.01167	0.01030
1	1.01168	0.01031
5	1.0119	0.01033
10	1.01223	0.01035
25	1.0133	0.01042

Table 10

Primer labeling reaction

Component	Final concentration	Stock	Amount (μl)
Oligonucleotide	440–660 ng/μl	4–6 mg/ml	11
SE-dye	2.52 μg/μl	18 mg/ml	14
Sodium tetraborate buffer (pH 8.5) ^a	75 mM	0.1 M	75

^aPrepare freshly before the coupling reaction or keep frozen aliquots at –20 °C.

Table 11

Sequencing-ladder master mix

Component	Final concentration	Stock	Amount (μl)
Plasmid template	1 ng/μl		
Thermo sequenase buffer			1.84
ddTTP mixture (Table 12)			16
Thermo sequenase	0.25 U/μl	4 U/μl	1.84

Add deionized H₂O to 28.8 μl.

Table 12

ddTTP mixture

Component	Final concentration (μM)	Stock (mM)	Amount (μl)
dATP	150	100	1.5
dCTP	150	100	1.5
dGTP	150	100	1.5
dTTP	150	100	1.5
ddTTP	1.5	1	1.5

Add deionized H₂O to 1 ml.

Table 13

Sequencing-ladder PCR program

Number of cycles	Temperature (°C)	Time
1	96	1 min
25	96	20 s
	55	20 s
	72	1 min
1	72	5 min

Table 14

Stop mixture

Component	Final concentration	Stock	Amount (μl)
Sodium acetate (pH 5.2)	1.5 <i>M</i>	3 <i>M</i>	200
Na-EDTA (pH 8.5)	0.05 <i>M</i>	0.1 <i>M</i>	200
Glycogen	0.8 mg/ml	20 mg/ml	16

Add deionized H₂O to 416 μl.

Table 15

Folding reaction

Component	Final amount/conc.	(+) Reagent sample (μl)	(–) Reagent sample (μl)
RNA	20 pmol		
5 × monovalent mixture (Table 16)	1×	98	98
MgCl ₂ 250 mM	25 mM	49	49

Add deionized H₂O to 490 μl.

Table 16

Monovalent mixture (5×)

Component	Final concentration	Stock	Amount (ml)
KCl	1 <i>M</i>	2 <i>M</i>	50
HEPES (pH 7.4)	0.25 <i>M</i>	1 <i>M</i>	25
Na-EDTA (pH 8.5)	0.5 mM	100 mM	0.5

Add deionized H₂O to 100 ml.

Table 17

SHAPE reaction

Component	Final amount/conc.	(+) Reagent sample	(-) Reagent sample
1M7 10 mM	1 mM	54.4 µl	–
DMSO		–	54.4 µl

Table 18

Quench mixture

Component	Final concentration	Stock	Amount (μl)
Sodium acetate (pH 5.2)	1.5 <i>M</i>	3 <i>M</i>	200
Na-EDTA (pH 8.5)	0.05 <i>M</i>	0.1 <i>M</i>	200
Glycogen	0.8 mg/ml	20 mg/ml	16

Add deionized H₂O to 100 ml.

Table 19

RNA storage buffer

Component	Final concentration (mM)	Stock (M)	Amount (μl)
MOPS (pH 6.5)	10	1	100
Na-EDTA (pH 8.5)	0.1	0.5	2

Add deionized H₂O to 10 ml.

Table 20

Monovalent ion mixture (5×)

Component	Final concentration	Stock	Amount (ml)
KCl	1 <i>M</i>	2 <i>M</i>	50
Cacodylic buffer (pH 7.0)	0.125 <i>M</i>	0.25 <i>M</i>	25
Na-EDTA (pH 8.5)	0.5 mM	100 mM	0.5

Add deionized H₂O to 100 ml.

Table 21

DMS reaction

Component	Final amount/conc.	(+) Reagent sample	(-) Reagent sample
DMS 81.3 mM	8.13 mM	54.4 µl	–
Ethanol		–	54.4 µl

Table 22

Quench mixture

Component	Final concentration	Stock	Amount
2-Mercaptoethanol	715 mM	14.3 M	50 µl

Add deionized H₂O to 1 ml.

Table 23

Precipitation reaction

Component	Final concentration	Stock	Amount (μl)
Sodium acetate (pH 5.2)	50.6 mM	3 M	20
Na-EDTA (pH 8.5)	8.4 mM	0.5 M	20
Glycogen	8.4 μg/μl	20 mg/ml	0.5
Isopropanol			600

Table 24

Annealing reaction

Component	Final amount/ conc.	(+) Reagent sample (μl)	(–) Reagent sample (μl)
RNA	1 pmol	<i>x</i>	<i>x</i>
2 μM labeled primer (JOE)	0.1 μM	1	3
2 μM labeled primer (FAM)	0.1 μM	–	1
2 mM Na-EDTA (pH 8.5)	0.1 mM	1	1
5 M betaine	1 M	4	4

Add deionized H₂O to 12 μl.

Table 25

Reverse transcriptase (RT) mixture

Component	Final concentration	Stock	Amount (μl)
RT buffer	2.5×	5×	96
DTT	12.6 mM	100 mM	24
Betaine	944 mM	5 M	36
dNTP mix	1.26 mM	10 mM	24
Superscript III RT	10 U/μl	200 U/μl	9.6
Deionized H ₂ O			1

Table 26

Stop mixture

Component	Final concentration	Stock	Amount (μl)
Sodium acetate (pH 5.2)	1.3 <i>M</i>	3 <i>M</i>	800
Na-EDTA (pH 8.5)	11 mM	100 mM	200
Glycogen	110 μg/μl	20 mg/ml	10
Deionized H ₂ O			800

Table 27

Primers used for spectral calibration

Primer ^a	Dye	Fragment (nt)
CAGAGCAGATTGTACTGAGAG	5-FAM	242
GTGTGAAATACCGCACAGAT	6-JOE	206
GCGTAAGGAGAAAATACCGCATC	6-TAMRA	188
CGCCATTGCGCATTGAGGCTGCGCAACTG	5-ROX	155

^aWatts et al. (2009).

Table 28

Spectral-calibration PCR reaction

Component	Final concentration	Stock	Amount (μl)
Linearized pUC18 plasmid	0.1 ng/μl	0.2 ng/μl	0.5
Thermo sequenase buffer	1×	10×	4
dNTP mixture	75 μM	0.75 mM	4
Labeled primer	0.0125 pmol/μl	2 μM	5
Thermo sequenase	0.16 U/μl	4 U/μl	1.6

Add deionized H₂O to 40 μl.

Table 29

Spectral-calibration PCR program

Number of cycles	Temperature (°C)	Time
1	96	2 min
25	96	20 s
	55	20 s
	72	1 min
1	72	10 min

Table 30

Stop mixture

Component	Final concentration	Stock	Amount (μl)
Sodium acetate (pH 5.2)	1.5 <i>M</i>	3 <i>M</i>	100
Na-EDTA (pH 8.5)	0.25 mM	0.5 <i>M</i>	100
Glycogen	10 mg/ml	20 mg/ml	10

Table 31

Capillary electrophoresis sample preparation (ShapeFinder)

Component	Amount (µl)
Reagent/blank mix	7
Sequencing ladder mix	2
Deionized formamide	12

Table 32

Capillary electrophoresis sample preparation (QuShape)

Component	Amount (μl)
Reagent or blank sample	7
Sequencing ladder	2
Deionized formamide	12

Table 33

Mobility shift sample preparation

Component	Amount (μl)
Mobility shift reaction	5
Deionized formamide	15