Published in final edited form as:

Methods Mol Biol. 2016 ; 1354: 91–117. doi:10.1007/978-1-4939-3046-3_7.

Novel Biochemical Tools for Probing HIV RNA Structure

Jason W. Rausch^{#1}, Joanna Sztuba-Solinska^{#1}, Sabrina Lusvarghi², and Stuart F.J. Le Grice³

¹Reverse Transcriptase Biochemistry Section, HIV Drug Resistance Program, Frederick National Laboratory for Cancer Research, Frederick, MD, 21702, USA.

²Laboratory of Bioorganic Chemistry, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD, 20892, USA.

³Reverse Transcriptase Biochemistry Section, HIV Drug Resistance Program, Frederick National Laboratory for Cancer Research, Building 535, Room 312, Frederick, MD, 21702, USA.
legrices@mail.nih.gov.

[#] These authors contributed equally to this work.

Abstract

Functional analysis of viral RNA requires knowledge of secondary structure arrangements and tertiary base interactions. Thus, high-throughput and comprehensive methods for assessing RNA structure are highly desirable. Selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE) has proven highly useful for modeling the secondary structures of HIV and other retroviral RNAs in recent years. This technology is not without its limitations however, as SHAPE data can be severely compromised when the RNA under study is structurally heterogeneous. In addition, the method reveals little information regarding the three-dimensional (3D) organization of an RNA. This chapter outlines four detailed SHAPE-related methodologies that circumvent these limitations. "Ensemble" and "in-gel" variations of SHAPE permit structural analysis of individual conformers within structurally heterogeneous mixtures of RNA, while probing strategies that utilize "through-space" cleavage reagents such as methidiumpropyl-EDTA (MPE) and peptides appended with an ATCUN (amino terminal copper/nickel binding motif) can provide insight into 3D organization. Combinational application of these techniques provides a formidable arsenal for exploring the structures of HIV RNAs and associated nucleoprotein complexes.

Keywords

Ensemble SHAPE; In-gel SHAPE; Methidiumpropyl-EDTA; ATCUN; Chemical probing; HIV RNA

1 Introduction

SHAPE, or selective 2'-OH acylation analyzed by primer extension, is a biochemical method that has been used to obtain secondary structural models for segments of HIV RNA including the 5' untranslated region (5' UTR; [1]) and the Rev response element (HIV-1 RRE; [2]), as well as the HIV-1 RNA genome in its entirety, both in vitro and in virio [3]. Unlike traditional chemical probing techniques, SHAPE reagents are non-hazardous and

exhibit minimal nucleotide reactivity bias. Moreover, SHAPE can be adapted for use with high-throughput capillary electrophoresis, making it the method of choice for biochemical RNA structure probing.

The defining feature of SHAPE is exposure of RNA to a probing reagent that selectively acylates unpaired ribonucleotides at the 2'-OH position. The sites and frequencies of RNA acylation are then recorded by reverse transcription, where cDNA synthesis pre-maturely terminates at RNA acylation sites. The resulting cDNA library is fractionated by denaturing polyacrylamide gel or capillary electrophoresis and detected/quantified as described herein and previously [4]. Raw cDNA quantities obtained by fractionating reverse transcription products are subjected to mathematical transformations such as signal decay correction, background subtraction and internal normalization to produce "reactivity values" attributable to individual template ribonucleotides. Collectively, a set of reactivity values obtained for a given RNA is referred to as the reactivity profile of that RNA. This metric may be inputted into RNAstructure software [5] to obtain the most energetically favorable secondary structural model(s) for the RNA in question. Alternatively, in the event that RNAstructure software is unavailable or does not meet the user's needs, a reactivity profile may be overlaid onto a secondary structural model generated by other methods, and the fit between the profile and model evaluated subjectively.

One shortcoming of SHAPE is that it has required the RNA under evaluation to be structurally homogeneous throughout the course of the acylation reaction. Failure to meet this requirement can produce an "averaged" structure model that fails to match any of the RNA conformers present in solution. However, two variations of the methodology have recently been developed that circumvent the requirement for RNA structural homogeneity: ensemble SHAPE and in-gel SHAPE. These variations enable assignment of RNA secondary structures to individual conformers present in a mixture, provided that these can be segregated by non-denaturing polyacrylamide gel electrophoresis. Although related, the SHAPE variations differ significantly in the manner in which reactivity data is obtained and processed.

The guiding concept of ensemble SHAPE is that experimentally determined reactivity values are actually composite values comprised of contributions from each of the individual RNA conformers present in a mixture. In addition, conformer contributions to the ensemble reactivity value are in proportion to their molar fractions in solution. As described herein, these principles can be expressed algebraically, thus allowing for mathematical extraction of theoretical reactivity values and profiles, and therefore structures, attributable to individual RNA conformers. Experimentally, ensemble SHAPE resembles conventional SHAPE except that the former additionally requires fractionation and quantitation of RNA conformers by non-denaturing polyacrylamide gel electrophoresis. SHAPE experiments must also be repeated under conditions that produce differing conformer distributions; specifically, at least one unique distribution is required per conformer in order for the resulting algebraic formulae to be solved. Ensemble SHAPE has been experimentally validated using in vitro-transcribed model RNAs derived from the HIV-1 5' UTR, and subsequently to identify intermediate structures assumed upon folding of the HIV-2 RRE [6].

Data acquisition and analysis are more straightforward for in-gel SHAPE, although this approach is more labor intensive and typically requires more RNA than ensemble SHAPE. During in-gel SHAPE, conformeric mixtures of RNA are fractionated by non-denaturing gel electrophoresis, bands comprised of individual conformers are excised from the gel, and acylating reagent is soaked into the slices. Modified RNAs are extracted by electroelution, precipitated, and subjected to the same reverse transcription and downstream procedures required in standard SHAPE. Like ensemble SHAPE, in-gel SHAPE is suitable for examining conformeric mixtures of RNA, particularly when the RNA is abundant and individual conformers can be completely separated by non-denaturing PAGE. In-gel probing is also theoretically applicable to footprinting of RNA-protein complexes, provided that these complexes are stable throughout the course of non-denaturing gel electrophoresis. In-gel SHAPE has been used to identify structural transitions that occur during dimerization of the HIV-1 5'UTR [7], as well as to define alternative conformations of the HIV-1 RRE (Rausch, J. W., unpublished observations).

While SHAPE is a simple and elegant tool for determining and refining RNA secondary structural models, none of its variations can be used to directly explore the 3D organization of an RNA. This information can be revealed using “through space” probing reagents, i.e., reagents that bind within defined motifs in an RNA and cleave or modify ribonucleotides in spatial proximity—including those not within or adjacent to the reagent binding site(s). Methidiumpropyl-EDTA (MPE) is one such reagent and has been used to validate the computer-generated three-dimensional structure of HIV-2 RRE RNA [6]. MPE comprises two functional groups: methidium, which has been reported to intercalate between adjacent ribonucleotides within defined RNA structural motifs such as CpG islands [8], and EDTA, which chelates Fe(II). When activated, tethered iron ions generate hydroxyl radicals that diffuse to and damage/cleave RNA near the MPE binding site(s). Nucleotides thus modified are identified by reverse transcription, and cDNAs are fractionated and quantified as in standard SHAPE. Such experiments, together with control experiments to identify specific methidium-binding sites, can provide useful insights into the 3D organization of an RNA.

Finally, in some instances, it is useful to tether a cross-linking or chemical cleavage agent to a protein or peptide component of a nucleoprotein complex in order to directly probe the ligand binding site(s) of an RNA. ATCUNs, or amino-terminal $\text{Cu}^{2+}/\text{Ni}^{2+}$ binding motifs, are 3-amino acid sequences (typically $\text{NH}_2\text{-Xaa-Xaa-His}$, where X is any amino acid) that have been shown to chelate these metal ions and consequently, when activated, damage or cleave proximal RNA. When incorporated into RNA-binding peptides, the modification patterns produced by these peptides correlate with and allow for definition of ligand-binding sites. We have generated peptides linking an ATCUN with the arginine-rich nucleic acid-binding motif of either HIV-1 or HIV-2 Rev (ARM1 or ARM2), and used these Rev-ATCUN peptides to probe Rev binding sites on their respective RREs. After being exposed to/modified by these peptides, RRE RNAs were subjected to reverse transcription and the subsequent processing of conventional SHAPE.

Together, ensemble SHAPE, in-gel SHAPE, MPE-Fe(II), and peptide-ATCUN probing constitute a formidable arsenal for exploring the structures of HIV RNAs and nucleoprotein complexes—even when the RNAs/complexes are not structurally homogeneous. These

methods are presented together here both because of their related applications and because they share a common backbone protocol of RNA modification, reverse transcription, cDNA fractionation and data processing/analysis. An additional method, i.e., three-dimensional modeling of RNAs using the recently developed RNAComposer software [9], is also briefly discussed.

2 Materials

Take precautions working with RNA to prevent non-specific degradation. Use molecular biology-grade chemicals and reagents and ultrapure water for preparation of all solutions.

1. RNA.

Extract RNA from biological samples or transcribe in vitro. It is often preferable to purify RNAs prior to probing, i.e., via denaturing polyacrylamide gel electrophoresis [10] or using commercially available kits, but this is not always essential. Shorter RNAs (<400 nt) are generally best suited for the applications presented here and may be engineered to contain a 3' structure cassette [4] to maximize the amount of native sequence that is readable. Store frozen, in low salt buffer such as TE light: 10 mM Tris-Cl, pH 7.6; 0.1 mM EDTA.

2. RNA folding.

10× RNA folding buffer: 500 mM Tris-HCl, pH 8.0, 50 mM MgCl₂, 1.0 M NaCl.

2× ND gel RNA folding buffer: 70 mM Tris-HCl, pH 8.0, 180 mM KCl, 0.3 mM EDTA, 8 mM MgCl₂, 10 % glycerol.

3. Non-denaturing polyacrylamide gel electrophoresis (ND PAGE) and RNA detection/quantitation.

10× TBE: 0.89 M Tris-borate, pH 8.3, 20 mM Na₂EDTA.

40 % acrylamide/bis-acrylamide (19:1).

10 % ammonium persulfate (store at 4 °C).

Tetramethylethylenediamine (TEMED).

10× ND gel loading buffer: 50 % glycerol, 0.025 % bromophenol blue.

Electrophoresis equipment (power supply, gel box, gel plates, combs, spacers).

SYBR Green II RNA gel stain (Life Technologies).

Fluorescent TLC plate.

Clear plastic wrap or plastic bag.

Fluorescence scanner.

4. RNA modification.

N-methylisatoic anhydride (NMIA) (store desiccated at ambient temperature).

1-Methyl-7-nitroisatoic anhydride (1M7) (*see* **Note 1**).

Dimethyl sulfoxide (DMSO).

5× TBE reaction buffer: 5× TBE, 25 mM MgCl₂.

5. Precipitation.

100 % ethanol (used to make 95 % ethanol).

3 M NaOAc, pH 5.5.

10 mg/mL glycogen.

6. MPE cleavage.

Methidiumpropyl-EDTA (*see* **Note 2**).

Methidium carboxylate TFA salt (*see* **Note 2**).

10 mM sodium ascorbate (prepare fresh with each use, protect from light).

0.3 % hydrogen peroxide (prepare fresh with each use).

Fe(SO₄)₂ (NH₄)₂ 6H₂O.

7. ATCUN-peptide RNA modification.

Peptides containing the ATCUN motif (N-terminus-aa1-aa2-His) can be synthesized in-house using solid-phase synthesis or purchased commercially.

Peptides should be HPLC purified and stored in water at −20 °C.

RNA buffer: 50 mM Tris–HCl, pH 8.0, 100 mM NaCl, 5 mM MgCl₂.

NiCl₂.

8. Reverse transcription.

SuperScript™ III Reverse Transcriptase (Life Technologies; *see* **Note 3**).

10 mM dNTP Mix, PCR grade.

Primers 5′ end-labeled with Cy5 or Cy5.5 (*see* **Note 4**).

9. Preparation of sequencing ladders.

Primers 5′ end-labeled with WellRed D2 or Licor IR800 (*see* **Note 4**).

USB Cycle Sequencing kit (*see* **Note 5**).

10. Capillary electrophoresis (*see* **Note 6**).

Deionized formamide.

Beckman-Coulter CEQ 8000 Genetic Analyzer.

96-Well sample microtiter plate.

Thermowell® 96 well PCR plate.

CEQ™ Separation Buffer (Beckman Coulter). LPA-1 gel (Beckman Coulter).

11. RNA modeling.

ShapeFinder software: To obtain this software, contact Morgan C. Giddings (<http://giddingslab.org/welcome>).

RNAstructure software: <http://rna.urmc.rochester.edu/RNAstructure.html>

VARNA RNA folding software: <http://varna.lri.fr/>

3 Methods

3.1 Ensemble SHAPE

Ensemble SHAPE is useful for studying RNAs that assume more than one conformational state in solution. While the differences between ensemble and conventional SHAPE relate primarily to data processing, two conditions must be met before an RNA may be analyzed using the former method. First, it must be possible to segregate RNA conformers by electrophoretic mobility on a non-denaturing polyacrylamide gel. This in turn requires that conformer interconversion occurs only slowly or not at all during electrophoresis. Second, at least one distinct mixture of conformers—each with a unique conformer distribution—must be evaluated for every conformer observed. This requirement relates to the algebraic relationships among the reactivity values measured for each mixture, the distribution of conformers in each mixture, and the theoretical reactivity values calculated for each conformer. This algebraic relationship is discussed in more detail below.

Ensemble SHAPE was used to determine the structures of transient folding intermediates of the HIV-2 RRE (Fig. 1) [6]. In this case, three distinct conformers could be segregated by non-denaturing gel electrophoresis, and the distribution of conformers in the mixture was found to vary with RNA folding time (i.e., the amount of time the RNA was incubated at 37 °C in the presence of 5 mM MgCl₂ after heating/cooling). As folding time increased, the less stable conformations were increasingly converted into the most stable form. We were able to evaluate five different conformer distributions, one each at 5, 10, 20, 40, and 100 min of RNA folding, thus satisfying both of the aforementioned requirements for using ensemble SHAPE.

3.1.1 RNA Folding—Parameters that must be considered in an RNA folding procedure include (1) RNA quantity and concentration, (2) reaction volume, (3) salt concentration, (4) Mg²⁺ ion concentration, and (5) folding duration. Although specific optimization goals may vary, it is generally desirable to establish conditions that are close to physiological and/or produce a structurally homogeneous RNA. Conditions used to fold HIV-2 RRE RNA are given below.

For ensemble SHAPE, RNA mixtures must be evaluated by non-denaturing gel electrophoresis (to determine the fractional contribution of each conformer) as well as by SHAPE. However, provided a given set of folding conditions consistently produces the same conformer distribution, it is not essential that the two experiments be conducted using portions of the same RNA sample. For the HIV-2 RRE, conformer distributions were the same regardless of RNA concentration. Hence, differing amounts of RNA (4 pmol or 40

pmol, respectively) were used for non-denaturing gel electrophoresis and for SHAPE. A typical HIV-2 RRE RNA folding procedure is described here:

1. Dissolve or dilute 4 (40) pmol RNA in 36 μ L of water.
2. Heat to 95 °C for 3 min and then immediately cool on ice for at least 5 min.
3. Add 4 μ L of 10 $\times$ RNA folding buffer. Mix.
4. Incubate at 37 °C.
5. At 5, 10, and 20 min, transfer 10 μ L folding mix to precooled 0.5 mL microfuge tubes on ice to arrest RNA folding.
6. Analytical ND gel samples: Add 1 μ L 10 $\times$ ND gel loading buffer and fractionate by ND PAGE (Subheading 3.1.2).
7. SHAPE samples: Dilute with 30 μ L 1 $\times$ RNA folding buffer (40 μ L total, each) and divide mixture into 20 μ L modified (+) and control (–) aliquots. Continue with RNA modification (Subheading 3.1.4).

3.1.2 RNA Fractionation by ND PAGE—Prior to conducting a SHAPE experiment, the structural homo- or heterogeneity of the RNA being studied should first be assessed on an analytical scale by non-denaturing polyacrylamide gel electrophoresis (ND PAGE) [11]. In-gel detection of RNA by SYBR Green staining typically requires loading of 25–100 ng RNA per well, depending on the length of the RNA in question, the number and distribution of conformers in the mixture, well width, and other parameters that must be empirically optimized. If the RNA is found to be heterogeneous, ensemble or in-gel SHAPE may then be used to determine the structures of individual conformers. However, these methods require additional analytical scale ND gel electrophoresis followed by quantification of RNA conformers or preparative-scale ND PAGE, respectively.

Because the conditions required for ND PAGE vary considerably according to the RNA probing methodology being used, the RNA being fractionated, electrophoresis equipment and a number of other factors, we are unable to provide a definitive protocol for this technique that is suitable for all of the applications presented here. However, the following procedure provides general guidelines for ND-PAGE that may be modified to suit a specific application:

1. Pouring the gel: We typically pour gels containing 5–8 % polyacrylamide (19:1 acrylamide:bis-acrylamide), 1 $\times$ TBE $\pm$ 4–5 mM MgCl_2 to fractionate RNAs from 100 to 400 nt in length. Gel plate dimensions are approximately 15 cm $\times$ 15 cm and spacers/combs are typically 0.8 or 1.6 mm thick. Polymerization of gel solution is activated prior to pouring by mixing in of 1/100th volume 10 % ammonium persulfate and 1/1000th volume tetramethylethylenediamine (TEMED).
2. Pre-running the gel: Clamp gel into the gel tank and add buffer. Remove comb and clean out the wells out with a syringe/needle. Gels are typically pre-run for

30–60 min in 1× TBE (± 5 mM MgCl₂) at 150–200 V and 4 °C prior to sample loading.

3. Sample loading: Sample volumes and RNA concentrations vary widely by method. Typically, 5–20 μ L samples containing a total of 25–100 ng RNA are loaded in each well for analytical ND PAGE and ensemble SHAPE, while 40–200 μ L samples of 10 μ g RNA or more may be required for in-gel SHAPE. Unless glycerol was already present in the RNA folding buffer, add 1/10th volume of 10× ND gel loading buffer to the samples, so they layer evenly at the bottoms of the wells.
4. Running the gel: Gels are run under conditions matching the pre-run, i.e., typically at 150–200 V and 4 °C. It is important that the RNA remains cool both immediately prior to and during the gel run to minimize interconversion among conformers and smearing of the associated bands.

3.1.3 Detection and Quantitation of RNA—Small quantities of RNAs that have been fractionated by ND PAGE are detected using the fluorescent dye SYBR Green, and quantified as necessary for ensemble SHAPE.

1. Remove gel sandwich from gel box and place flat on workbench. Carefully remove top plate using spacers and/or spatulas. Pour or wick away with a paper towel any pooled buffer on the surface of the gel.
2. Dilute 1 μ L SYBR Green dye in 10 mL water and pour solution evenly over gel surface. Use roller where necessary to ensure complete coverage. Wait for 5 min to allow dye to soak into gel and bind RNA.
3. Pour off or wick away SYBR Green solution from the gel. Carefully place clear plastic wrap over the gel surface and around the bottom plate to seal in residual liquid.
4. Scan gel in fluorescence scanner at 520 nm and quantify fractionated conformers using appropriate software.

3.1.4 RNA Modification and Recovery—1M7 and NMIA are the best-characterized SHAPE reagents, with half-lives in aqueous solution of ~70 s and ~8 min, respectively [12]. However, in order to minimize the opportunity for RNA conformers to interconvert during the course of the acylation reaction, the more reactive 1M7 reagent is generally preferred for use in ensemble SHAPE experiments.

1. Prepare a 30 mM solution of 1M7 by adding small amounts of reagent to a 1.5 mL microfuge tube and then adding an appropriate volume of DMSO (*see Note 7*).
2. Add 2 μ L 1M7 solution or DMSO, respectively, to the 20 μ L (+) and (–) aliquots produced in the RNA folding step (Subheading 3.1.1).
3. Incubate reactions at 37 °C for 5 min (*see Note 8*).

4. Precipitate RNA. Add 52.5 μL cold 100 % ethanol, 2.1 μL 3 M NaOAc (pH 5.5) and 0.2 μL glycogen (10 mg/mL). Mix gently. Incubate at $-20\text{ }^{\circ}\text{C}$ for 2 h.
5. Collect precipitated RNA by centrifugation at $14,000 \times g$ for 30 min, carefully remove supernatant, and air-dry the pellet for 1–5 min (*see Note 9*).
6. Dissolve RNA in 10 μL of TE buffer. Proceed to reverse transcription (Subheading 3.1.5), or store at $-20\text{ }^{\circ}\text{C}$.

3.1.5 Reverse Transcription—This step generates cDNA products that indirectly identify the degree to which RNA nucleotides have been modified by 1M7 or NMIA. If reverse transcription reactions are primed with fluorescently tagged oligonucleotides, the relative abundance of each cDNA fragment can be assessed via capillary electrophoresis (CE). Oligonucleotides labeled with Cy5 and Cy5.5 are used to prime the (+) and (–) reactions, respectively. These primers should be stored in small aliquots in amber microfuge tubes at $-20\text{ }^{\circ}\text{C}$. Note that 5' ^{32}P -labeled primers (5'-RP) may be used as an alternative to fluorescent primers in the reverse transcription reactions, in which case lower concentrations of both primer and template may be used. Additional comments addressing the use of radiolabeled primers are designated 5'-RP below.

1. Prepare (+) and (–) RNA samples for reverse transcription in 0.5 mL microfuge tubes. Transfer 5 μL of (+) and (–) samples from modification step (~2.5 pmol RNA each) into 0.5 mL microfuge tubes. Store the unused portion at $-20\text{ }^{\circ}\text{C}$. For the (+) RT reaction, add 6 μL of water and 1 μL Cy5-labeled primer (10 μM); for the (–) RT reaction, add 6 μL water and 1 μL Cy5.5-labeled primer (10 μM).
2. Thaw the RT mix reagents (5 $\times$ RT buffer, 100 mM DTT, RNase-free water, and 10 mM dNTPs) and place on ice.
3. Place (+) and (–) reaction tubes in a thermal cycler and initiate the following program: 85 $^{\circ}\text{C}$, 1 min; 60 $^{\circ}\text{C}$, 5 min; 35 $^{\circ}\text{C}$, 5 min and 50 $^{\circ}\text{C}$, 50 min (*see Note 10*).
4. During annealing, prepare 2.5 $\times$ RT mix by combining 4 μL 5 $\times$ RT buffer, 1 μL 100 mM DTT, 1.5 μL water, 1 μL 10 mM dNTPs and 0.5 μL SuperScript™ III reverse transcriptase for each RT reaction. Mix well and incubate at 37 $^{\circ}\text{C}$ for 5 min.
5. Once the temperature of the annealing reactions reaches 50 $^{\circ}\text{C}$, add 8 μL of 2.5 $\times$ RT mix to each and continue incubation at 50 $^{\circ}\text{C}$ for 50 min (*see Note 11*).
6. Add 1 μL 4 M NaOH to each RT reaction and incubate at 95 $^{\circ}\text{C}$ for 3 min to hydrolyze the RNA. Cool the tubes on ice and neutralize with 2 μL of 2 M HCl (*see Note 12*).
7. Combine (+) and (–) reactions in a single tube and precipitate by adding 54 μL water, 10 μL 3 M NaOAc, 0.8 μL glycogen (10 mg/mL), and 300 μL 100 % ethanol. Incubate at $-20\text{ }^{\circ}\text{C}$ for 2 h. Pellet cDNA by centrifugation at $13,000 \times g$ for 30 min, 4 $^{\circ}\text{C}$.

8. Wash pellet(s) twice with 400 μL of cold 70 % ethanol and centrifuge immediately at $13,000 \times g$ for 5 min (*see Note 13*).
9. Remove supernatant and dry pellet for 3 min by centrifugal vacuum evaporation (*see Note 14*).
10. Dissolve DNA in 40 μL of deionized formamide (*see Note 15*).

3.1.6 Preparation Sequencing Ladder—5'-end-labeled sequencing ladders generated by cycle sequencing of (USB cycle sequencing kit, product #78500) are co-fractionated with (+) and (–) reverse transcription products by CE in order to map the 3' termini of modified cDNAs. PCR-generated transcription templates can serve as sequencing templates, provided that they are purified from PCR primers and dNTPs prior to use. Due to the relatively low fluorescence of WellRed D2 and IRDye800, sequencing reactions must be scaled up severalfold to ensure sufficient yield for detection by capillary electrophoresis. In the example below, ddA and ddC are used as chain terminating nucleotides. However, any combination of dideoxynucleotide triphosphates may be used to generate sequencing ladders. 5'-RP: Use the USB cycle sequencing kit protocols for 5' end-labeling and cycle sequencing.

1. Mix 16 μL ddA termination mix, 0.5 pmol of DNA template, 1.8 μL 10 $\times$ Sequenase buffer, 2.0 μL WellRed D2 labeled primer (5 μM) in a 0.5 mL thin-walled microfuge tube. Add water to bring the total volume to 30 μL , mix, add 2.0 μL Sequenase, and mix again. Store on ice until cycling.
2. Prepare a second, parallel sequencing reaction including primer labeled with IRDye800 and ddC termination mix.
3. Cycle both sequencing mixes according to the program recommended by the kit manufacturer.
4. Combine the ddA and ddC sequencing reactions into one 1.5 mL microfuge tube (64 μL)
5. Precipitate sequencing products by adding 6.4 μL 3 M NaOAc (pH 5.5), 6.4 μL 100 mM EDTA, 0.5 μL 10 mg/mL glycogen, and 192 μL 95 % ethanol. Mix well, incubate at 4 $^{\circ}\text{C}$ for 2 h, and centrifuge at $13,000 \times g$ for 30 min at 4 $^{\circ}\text{C}$.
6. Remove supernatant by pipetting and dry pellet by vacuum centrifugation for 5 min.
7. Resuspend pelleted DNA in 100 μL of deionized formamide by heating to 65 $^{\circ}\text{C}$ for 10 min, followed by vortexing for at least 30 min.
8. Mix 10 μL pooled sequencing reactions with 40 μL pooled (+) and (–) RT reactions. Store the remaining sequencing reaction mix at –20 $^{\circ}\text{C}$.

3.1.7 Capillary Electrophoresis—High-throughput capillary electrophoresis with fluorescence detection allows simultaneous fractionation and detection of four differently labeled DNA products in a single sample. The Beckman-Coulter CEQ 8000 Genetic

Analyzer can fractionate as many as eight sets of pooled reactions simultaneously, while as many as 96 samples can be fractionated consecutively in a single program.

1. Load 40 μ L of the pooled reactions (i.e., (+) Cy5-labeled cDNAs, (–) Cy5.5-labeled cDNAs, and two sequencing ladders labeled with WellRed D2 and IRDye800) into a single well in a 96-well sample plate. Repeat as necessary for additional pooled reactions, making sure to load them into the same row.
2. Program and prepare capillary electrophoresis instrument according to the previously published capillary automated footprinting analysis (CAFA) method parameters [13] and initiate run as per the manufacturer's instructions.
3. Export raw fluorescence traces into ShapeFinder.

5'-RP: Radioactive reverse transcription and cycle sequencing products can be fractionated over a denaturing slab gel (5–8 % 19:1 polyacrylamide:bis-acrylamide, 1 $\times$ TBE, 7 M urea). Typically, samples are loaded in adjacent wells, and loading is staggered to maximize the amount of readable sequence. Gels are dried and exposed to a phosphorimaging cassette.

3.1.8 Data Analysis by ShapeFinder—ShapeFinder is free available software that allows the user to mathematically transform and graphically visualize CE migration data. To acquire ShapeFinder software, contact the Giddings lab (<http://www.giddingslab.org>). We provide here an overview of how SHAPE data are transformed to obtain an RNA reactivity profile. Detailed instructions for data handling in ShapeFinder are provided with the software documentation.

1. Adjust raw data to account for (1) fluorescent background, (2) spectral overlap between fluorescent traces, (3) differential mobility of equivalent DNAs labeled with different fluorophores, and (4) signal decay resulting from the imperfect processivity of reverse transcriptase (see **Note 16**).
2. Correlate peaks in the adjusted electropherogram to specific ribonucleotides in the RNA sequence using the “Setup” function of the “Align and Integrate” tool. Errors in initial assignments can be corrected manually using the “Modify” function of the same tool.
3. Calculate the areas under the aligned (+) and (–) reaction peaks using the “Fit” function. These values are automatically tabulated in a tab-delimited text file readable in Microsoft Excel or other spreadsheet software.

5'-RP: Bands corresponding to individual cDNAs are detected by phosphorimaging and quantified using software such as ImageQuant TL (GE Healthcare Life Sciences). Band volumes for (–) reaction products are subtracted from the corresponding (+) reaction products to generate raw reactivity values. Note that while cDNAs co-migrate with chain-terminated sequencing products of the same length, reactivity values should be attributed to ribonucleotides in the RNA sequence one position 5' to the complementary cDNA stop position. Also, unlike ShapeFinder, ImageQuant TL cannot systematically correct for signal decay or band overlap as slower migrating products become less intense and closer together. For this, SAFA, or semiautomated footprinting analysis software, has proven useful [14].

This software is free and can be downloaded from http://ribosnitch.bio.unc.edu/The_Laederach_Lab/Software.html.

3.1.9 Data Normalization—Before raw reactivity values can be used for secondary structure modeling, SHAPE data must be normalized according to the “2–8 % rule” described previously [15, 16]. After this transformation, reactivity values of <0.2 and >0.8 (for example) are deemed to reflect highly or minimally constrained nucleotide bases that are most likely base-paired or single-stranded, respectively. Extreme values (i.e., those <-0.2 or >1.5) must be interpreted with caution, as they are often indicative of natural RT pause sites (“stops”) or are artificially produced by imperfect signal decay correction and/or data scaling. Because they typically represent $<1\%$ of the data, excluding outliers from the reactivity profile inputted into the RNAstructure software is often the most prudent course of action. A general summary of the data normalization process follows:

1. In a new column in Microsoft Excel, order raw reactivity values and identify quartiles (Q1, Q2, and Q3).
2. Identify “outlier” values; i.e., values greater than $Q3 + 1.5 \times (Q3 - Q1)$.
3. Excluding outliers, calculate the average of the highest 8 % of the remaining values. This average is defined as the “effective maximum” of the data.
4. Normalize the original, unsorted reactivity data by dividing all values (including outliers) by the effective maximum, thereby producing normalized reactivity values for all nucleotides for which SHAPE data was obtained.
5. For ensemble SHAPE, proceed to “Mathematical Deconvolution” below (Subheading 3.1.10), after which **steps 6 and 7** here should be followed to eliminate bad values and to format conformer-specific reactivity profiles for use with RNAstructure. For the other three methods, proceed directly to **steps 6 and 7** below.
6. Eliminate values less than -0.09 and those associated with unusually strong natural pause sites (i.e., “stops”). Replace omitted values with “-999”, the preferred skipped-value indicator in the RNAstructure software.
7. Save the paired nucleotide numbers and associated, normalized reactivity data as tab-delimited pairs arranged in adjacent columns as a text file with the suffix “.shape”. This file provides the “SHAPE reactivity values” used to generate pseudoenergy constraints in RNAstructure software.

3.1.10 Mathematical Deconvolution (Ensemble SHAPE Only)—Normalized reactivity values obtained in ensemble SHAPE are comprised of contributions from all RNA conformers in a mixture. Because these are averaged values, they cannot be used to direct the secondary structural prediction of any single conformational variation of the RNA being probed. In order to obtain a reactivity profile attributable to an individual RNA conformer, theoretical reactivity values must be mathematically extracted from the experimentally obtained ensemble values, as described below. The resulting profile(s) may be inputted into RNAstructure software in the same manner as a profile generated by conventional SHAPE.

In this case, however, the set of theoretical reactivity values will be used by the software to generate pseudoenergy constraints that reflect the structure of the respective RNA conformation exclusively.

For an RNA that can assume any of the conformations A, B, or C (and perhaps others), and for each nucleotide position probed by SHAPE, the following mathematical relationship applies:

$$R_T = p_A R_A + p_B R_B + p_C R_C \dots$$

In this equation, R_T is the experimentally determined, normalized ensemble reactivity value for a specific nucleotide position, and p_A , p_B , and p_C are the fractional contributions of individual conformers to total RNA, as determined by quantitative ND gel electrophoresis. R_A , R_B , and R_C are the theoretical reactivity values for the same nucleotide in the respective conformers; these values are calculated, and should remain constant regardless of the conformer distribution established in a given set of experimental conditions.

In the simplest possible example for which ensemble SHAPE may be applied, two RNA conformers are differently distributed under two distinct folding/solution conditions 1 and 2. Given the experimentally established mathematical relationship presented above, a two-variable, two-equation system may be established for this scenario as follows:

$$R_{T1} = p_{A1} R_A + p_{B1} R_B \quad \text{Conditions—1}$$

$$R_{T2} = p_{A2} R_A + p_{B2} R_B \quad \text{Conditions—2}$$

Variable assignments are the same as in the general case above except that the subscript “1” or “2” indicates the value in question was measured using RNA folded under the specified set of conditions. Also, for the system of equations to be soluble in all but the trivial instance in which $R_{T1} = R_{T2}$, conformer distributions must differ under the two sets of conditions (i.e., so that $p_{A1} \neq p_{A2}$ and $p_{B1} \neq p_{B2}$). Note that the theoretical reactivity values R_A and R_B are inherent to the respective conformers and do not vary with differing folding conditions. The system may be solved for R_A and R_B using conventional algebraic techniques and/or matrix mathematics, and the process repeated for all nucleotides for which normalized ensemble reactivity values have been determined. This is most easily accomplished using Microsoft Excel or other spreadsheet software. The theoretical reactivity profiles obtained in this manner may be used to generate secondary structural models for individual RNA conformers as in conventional SHAPE.

3.1.11 RNA 2D Modeling—Prediction of experimentally supported RNA secondary structure requires RNAstructure, a stand-alone application that converts SHAPE-derived reactivity values into pseudo-free energy constraints which are then incorporated into the RNA folding algorithm [5]. The software generates 2D representations of RNA structures

both in graphical format and dot-bracket notation. The latter can be inputted into RNA structure viewers such as PseudoViewer [17] or Varna [18] to produce publication-quality images.

1. RNAstructure software is available at: <http://rna.urmc.rochester.edu/RNAstructure.html>.
2. Input the RNA sequence and SHAPE data as formatted text files containing “.seq” and “.shape” suffixes, respectively. Sequence data may also be inputted manually.
3. The user may alter “slope” and “intercept” parameters in the linear function that defines the relationship between reactivity and pseudo-free energy [15]. We generally do not find it necessary to adjust the default settings of 2.6 kcal/mol and –0.8 kcal/mol for slope and intercept, respectively.
4. RNAstructure ranks the models it produces by pseudo free energy, where the structure having the lowest energy is the one that best matches (1) the empirically determined thermodynamics rules for base-pairing embedded in the folding algorithm and (2) the pseudoenergy constraints imposed by the SHAPE reactivity profile. To facilitate qualitative analysis of the match between reactivity values and base pairing predictions, the software permits addition of a color-coded overlay of the reactivity profile onto the graphical representations. In a high-quality structure, weakly or unreactive and highly reactive nucleotides should generally be found in base-paired and unpaired regions, respectively.

One drawback to using RNAstructure as the exclusive means of predicting base pairing interactions is the inability of the software to predict the existence of pseudoknots, kissing loops, or other complex tertiary interactions. These must either be detected manually or through use of alternative RNA structure prediction software that, while useful, cannot optimize their models to match experimentally determined reactivity values. To address this shortcoming, a modification of the conventional SHAPE methodology, designated antisense-interfered or aiSHAPE, was developed to interrogate long-range tertiary interactions. In brief, a short oligonucleotide containing a combination of locked nucleic acid and 2'-O-methyl substitutions is site-specifically hybridized to a site within such a putative structural motif. Hybridization of the interfering oligonucleotide disrupts native base-pairing, with the consequence that the displaced complement is rendered susceptible to acylation. Reactivity profiles obtained in the presence and absence of the oligonucleotide are then compared, and a determination is made as to whether the differences between the two reactivity profiles are consistent with the existence of the predicted tertiary interaction. aiSHAPE has been used to verify the existence of pseudoknot and kissing loop interactions within the MusD RNA transport element MTE [19], as well as the Dengue virus [20] and HIV RNA genomes [21], and should be appropriate for use with the ensemble and in-gel SHAPE methods presented here. A more detailed summary of the aiSHAPE methodology has been presented elsewhere [22].

3.1.12 3D Modeling: RNAComposer—RNAComposer is a server-based modeling program that generates 3D models of RNA structure from inputted RNA sequences of up to

500 nts in length, the corresponding secondary structure(s) and, when available, atomic distance constraints [9]. Sequence and secondary structure information is provided as strings of text—the latter formatted into dot-bracket notation. Other formatting requirements are described in detail on the hosting website. In brief 3D RNA structure models are generated as follows: (1) the secondary structure is fragmented into small pieces; (2) the algorithm searches FRABASE, a comprehensive database of RNA structures that have been solved in three dimensions, for motifs matching or near-matching the secondary structure of each fragment; (3) the 3D structures of the best-matching motifs are selected for assembly; (4) a second algorithm assembles the motifs into a complete, intact structural model; and (5) the model is subjected to multiple rounds of energy minimization and, where possible, fit into conserved folding patterns. RNAComposer is available at <http://rnacomposer.cs.put.poznan.pl/>.

3.2. In-Gel SHAPE

3.2.1 RNA Folding—Because in-gel SHAPE requires recovery of RNA fractionated by ND PAGE, more starting material is required than for any of the other techniques presented here. This preparative scale folding protocol was optimized for analysis of a truncated HIV-1 RRE RNA construct by in-gel SHAPE (Fig. 2) (Rausch, J. W., unpublished observations). In general, it is assumed that prior to preparative scale RNA folding, folding and fractionation conditions will have been optimized on an analytical scale.

1. Mix RNA (100 pmol), water, and 2 μ L 10 $\times$ RNA folding buffer to a final volume of 20 μ L on ice.
2. Heat to 85 $^{\circ}$ C for 3 min, and then cool to 25 $^{\circ}$ C at a rate of 0.1 $^{\circ}$ C/s.
3. Add 20 μ L of 2 $\times$ ND gel RNA folding buffer. Mix.
4. Incubate at 37 $^{\circ}$ C for 20 min, then immediately place on ice prior to fractionation by ND PAGE.

For in-gel SHAPE, RNA samples should be evenly distributed over several lanes to facilitate even division into modified (+) and control (–) samples during gel excision. It is assumed that conformers of the RNA in question migrate differently and are distinguishable following electrophoresis. Other ND PAGE guidelines can be found in Subheading 3.1.2.

3.2.3 Detection, Quantitation, and Isolation of RNA by UV Shadowing (Preparative Scale)—It is impractical to detect large quantities of fractionated RNA using SYBR Green as background fluorescence produced during preparative scale ND PAGE is too high. Moreover, intercalation of SYBR Green into the RNA may alter its reactivity to NMIA or other probing reagents. As a consequence, an indirect detection method [7] or UV shadowing is preferable. The latter approach was used to detect HIV-1 RRE conformers (Rausch, J. W., unpublished observations) and is presented here.

1. Remove gel plate or gel sandwich from gel box and place flat on workbench. Remove top plate using spacers and/or spatulas. Pour off or blot any pooled buffer from the gel with a paper towel.

2. Wrap an appropriately sized fluorescent TLC plate in clear plastic wrap, minimizing wrinkles over the silica-coated surface. Alternatively, the TLC plate may be placed inside a smooth, clear plastic bag. Place wrapped or bagged TLC plate silica-side down on the gel surface. Invert gel sandwich and use a spatula to carefully remove the glass plate.
3. Wearing eye protection and covering exposed skin, use a handheld short-wave UV lamp (254–300 nm) to illuminate the TLC plate. RNA bands will be evident as “shadows” of UV absorption amidst a background of white fluorescence.
4. Using a clean spatula, excise gel slices containing the fractionated conformers, segregating equivalent modified (+) and control (–) RNA conformer samples. Work quickly to minimize UV damage to RNA.

3.2.4 RNA Modification with NMIA—For in-gel SHAPE, it is essential that the acylating reagent have a half-life in solution sufficient to allow it to soak into the gel before being hydrolyzed. Consequently, a relatively slow-acting reagent such as NMIA is best suited for these experiments. The protocol below was used to chemically modify two distinct HIV-1 RRE RNA conformers.

1. Prepare 50 mL each of 5× TBE reaction buffer and 1× TBE. Store at 4 °C.
2. For each conformer to be probed, mix 1050 µL water and 300 µL 5× TBE reaction buffer in each of two 2 mL microfuge tubes labeled NMIA(+) and DMSO(–). Heat to 37 °C.
3. Prepare 100 mM NMIA solution in DMSO. Keep desiccated and protected from light until use.
4. Submerge the (+) and (–) gel slices obtained from UV shadowing and gel excision in the corresponding labeled tube.
5. Immediately add 150 µL NMIA solution to the NMIA(+) reaction. Add 150 µL DMSO to the control reaction. Mix both reactions well.
6. Incubate all reactions at 37 °C for 50 min, then cool to 4 °C, and remove reaction solution.
7. Wash gel slices by adding 1.5 mL wash buffer (1× TBE), inverting 3–4 times, and then centrifuging briefly at ~2500 × *g* on a tabletop centrifuge. Repeat twice to completely remove NMIA hydrolysis byproducts.
8. Recover RNA by electroelution in 1× TBE buffer according to the manufacturer’s recommendations.
9. Add 0.1 volumes 3 M NaOAc (pH 5.5), 3 volumes 95 % EtOH, and 0.5 µL 10 mg/mL glycogen to recovered RNA solutions.
10. Centrifuge 20 min at 13,000 × *g* in refrigerated tabletop centrifuge. Decant supernatant, and wash with ice-cold 70 % EtOH.

11. Centrifuge an additional 5 min to recover RNA pellet. Decant supernatant, and air-dry for 10 min. Resuspend in 12 μ L TE light.
12. Determine concentration of recovered RNA by UV absorption.
13. Determine the molar concentration of each RNA, and dilute in TE light to a concentration suitable for reverse transcription (e.g., 0.5 pmol/ μ L). Proceed to reverse transcription, or store samples at -20°C .

3.2.5 Further Processing and Data Analysis—Once RNA has been modified and electroeluted from the gel, subsequent processing and data analysis closely resembles that required for ensemble SHAPE. Accordingly, proceed with reverse transcription (Subheading 3.1.5), preparation of sequencing ladder (Subheading 3.1.6), capillary electrophoresis (Subheading 3.1.7), ShapeFinder data analysis (Subheading 3.1.8), Data Normalization (Subheading 3.1.9), RNA 2D Modeling (Subheading 3.1.11), and RNA 3D Modeling (Subheading 3.1.12) as appropriate.

3.3 Probing RNA with MPE-Fe(II)

MPE is a bifunctional reagent composed of methidium, an intercalator reported to site-bind RNA at select sites, and EDTA, a chelator of metal ions such as Fe^{2+} (Fig. 3). Under oxidizing conditions, the MPE-Fe(II) complex generates highly reactive hydroxyl radical species inducing strand scission at the MPE-binding site(s) as well as in neighboring regions of RNA. MPE cleavage can therefore provide information on RNA nucleotides that are proximal to the reagent binding site, even when these nucleotides and the binding site are separated in primary sequence.

MPE-Fe(II) has been used to probe the 3D structure of tRNA containing engineered CpG islands [23]. In a separate study, however, binding of MPE-Fe(II) to RNA was considered to be non-specific, and the reagent was used to map solvent accessibility throughout the RNAs of interest [8]. The following protocol describes the application of MPE-Fe(II) to validate the two-dimensional models generated for RRE2 and provide clues regarding the 3D organization of RRE2 substructures [6].

3.3.1 RNA Folding: MPE—The following protocol was optimized for MPE and Rev-ATCUN probing of the HIV-2 RRE [6]. As with all protocols presented here, RNA heterogeneity should be verified by analytical scale folding and ND PAGE.

1. Working on ice, bring RNA (20 pmol) to a total volume of 36 μ L in water.
2. Heat RNA to 85°C for 3 min and then cool to 25°C at a rate of 0.1°C/s .
3. Add 4 μ L of $10\times$ RNA folding buffer. Mix.
4. Incubate at 37°C for 20 min, and then immediately place on ice.

3.3.2 MPE Binding and Cleavage—The activated metal ion chelated by MPE (Fe^{2+}) is responsible for RNA damage.

1. Transfer 18 μL folded RNA (Subheading 3.3.1) into each of the two 0.5 mL microfuge tubes labeled (+) and (–), respectively.
2. Prepare fresh solutions of 10 mM sodium ascorbate and 0.03 % hydrogen peroxide (H_2O_2).
3. Form the MPE–Fe(II) complex by mixing equal volumes of 20 μM MPE and 8 μM $\text{Fe}(\text{SO}_4)_2 (\text{NH}_4)_2 \cdot 6\text{H}_2\text{O}$ solutions to achieve final concentrations of 10 μM MPE and 4 μM Fe(II) (1:0.4 ratio). Incubate for 10 min at room temperature. This solution is 10 $\times$ relative to the binding and cleavage reaction, so 2 μL is required per 20 μL reaction. Prepare enough for all experiments.
4. For use in the (–) control reactions, mix equal volumes of 20 μM MPE and water to generate the 10 $\times$ MPE control solution. Again, prepare enough for all (–) control reactions (2 μL required per reaction).
5. Add 2 μL 10 $\times$ MPE–Fe(II) complex and 10 $\times$ MPE control to the (+) and (–) RNA solutions, respectively. Incubate for 1 min at room temperature to allow methidium to bind RNA.
6. Initiate cleavage by adding sodium ascorbate and hydrogen peroxide solutions (final concentrations of 1 mM and 0.003 %, respectively) to both the (+) and (–) reactions. Incubate at room temperature for 15 s.
7. Quench reactions by adding 2 μL 3 M NaOAc, 1 μL of 10 mg/ mL glycogen, and 50 μL ice-cold 95 % ethanol.
8. Pellet RNA by centrifugation at $14,000 \times g$ for 30 min. Carefully remove supernatant and air dry the pellet for 1–5 min.
9. Dissolve the RNA in 10 μL of H_2O and proceed to reverse transcription (Subheadings 3.3.3 and 3.3.4).

3.3.3 Determining Methidium-Binding Site by Hydroxyl Radical Cleavage—

Hydroxyl radical footprinting can be used to indirectly identify methidium-binding sites. In brief, RNA is incubated with methidiumpropyl compound devoid of the EDTA moiety and the resulting complex probed by exogenous Fe^{2+} -EDTA. In contrast to MPE probing, hydroxyl radical footprinting target nucleotides that are not protected by methidium binding, thus revealing the methidium footprint(s). Once methidium-binding sites have been identified, comparing these sites with those targeted by MPE reveals which cleavages are produced by diffusion of hydroxyl radicals to nearby sites in the RNA that do not directly contact MPE.

1. An additional 20 pmol RNA must be folded in the same manner as described in Subheading 3.3.1.
2. Transfer 9 μL folded RNA (Subheading 3.3.1) into each of the two 0.5 mL microfuge tubes labeled (+) and (–).
3. Prepare fresh 10 μM methidium carboxylate solution in double-distilled water.

4. Add 1 μL of 10 μM methidium carboxylate solution to the (+) reaction and 1 μL of H_2O to the (–) reaction. Incubate for 10 min at room temperature.
5. Perform hydroxyl radical footprinting according to previously published procedures [24]. Briefly, prepare Fe(II)-EDTA solution (10 mM Fe(II), 20 mM EDTA), 0.6 % H_2O_2 and 10 mM sodium ascorbate. Carefully place 1 μL of each reagent in separate droplets on the inside wall of the (+)-labeled microtube containing folded RNA. Centrifuge the tube to mix the reagents and initiate the footprinting reaction. Incubate the reaction mixture for 30 s at room temperature. A control experiment should be performed in the same manner, except add water rather than methidium carboxylate to the reaction.
6. Quench as described above for the MPE experiment (**step 6**).
7. Proceed to reverse transcription (Subheading 3.3.4), or store at -20°C .

3.3.4 Further Processing and Data Analysis—Once MPE and methidium-binding/cleavage reactions have been performed, subsequent processing and data analysis closely resembles that required for ensemble SHAPE. Accordingly, proceed with reverse transcription (Subheading 3.1.5), preparation of sequencing ladder (Subheading 3.1.6), capillary electrophoresis (Subheading 3.1.7), ShapeFinder data analysis (Subheading 3.1.8), and Data Normalization (Subheading 3.1.9).

3.3.5 RNA 2D Modeling—After modification data have been normalized, values can be overlaid onto RNA secondary structure maps in RNAstructure by color-coding. This means of viewing the data can be helpful for discerning patterns or groupings among affected nucleotides. However, the secondary structure maps themselves must be generated using SHAPE or other methods. MPE or methidium footprinting data cannot be used directly to generate RNA secondary structural models.

3.3.6 RNA 3D Modeling: RNAComposer—Values obtained by MPE probing or methidium footprinting indicate sites damaged by MPE and/or bound by methidium. As such, they may be used to help validate 3D RNA structural models such as those obtained for RRE2 folding intermediates [6]. It might also be possible to use MPE probing data to generate atomic distance constraints useful for modeling in RNAComposer [9]. More information on RNAComposer and imposing distance constraints during 3D modeling can be found in Subheading 3.1.12, the RNAComposer website (<http://rnacomposer.cs.put.poznan.pl/>), and the original RNAComposer ref. 9.

3.4 ATCUN Probing

Much like MPE, ATCUN-peptides can be used to damage or cleave RNA in the vicinity of the peptide binding site(s) (Fig. 4). The method presented here was adapted from previous work in which ATCUN-peptides containing the arginine rich motif (ARM) of Rev1 were incubated with RRE RNA and activated to cleave the HIV-1 RRE [25]. These peptides were developed as antiviral agents, and bound Cu^{2+} rather than Ni^{2+} as described here. In our experiments, ATCUN-ARM peptides derived from HIV-1 or HIV-2 Rev were complexed with Ni^{2+} , activated with the oxidizing reagent MMPP and used to map the binding sites of

Rev1 and Rev2 on their respective RREs (Lusvarghi, S., unpublished observations). This approach also offers the potential to map Tat, Gag, NC, and/or other HIV-related protein-binding sites on HIV RNAs.

3.4.1 RNA Folding ATCUN: RNA is folded during setup for the ATCUN-binding/cleavage reaction (Subheading 3.4.2).

3.4.2 ATCUN Peptide Binding/RNA Modification

1. Prepare experimental (+) (400 pmol peptide, 360 pmol NiCl_2) and control (–) (400 pmol peptide only) peptide solutions in 20 μL of 1 $\times$ RNA folding buffer (50 mM Tris–HCl, pH 8.0, 100 mM NaCl, 5 mM MgCl_2). Incubate peptide solutions for 20 min at 37 °C prior to mixing with RNA. Solutions may be stored at –20 °C for later use.
2. Heat 10 μL aliquots of RNA (4 μM in water) to 95 °C for 3 min, and then immediately place on ice for 3 min.
3. To complete RNA folding, add 2 μL of 10 $\times$ RNA folding buffer and 8 μL water to the RNA sample, mix, and incubate at 37 °C for 20 min.
4. Divide RNA solutions evenly between experimental (+) and control (–) reactions (10 μL each). Mix RNA solutions with ATCUN peptide solutions with (+) or without (–) Ni^{2+} , respectively, at the different peptide:RNA ratios (e.g., 0.5:1, 1:1, or 2:1). The final volume of each reaction should be adjusted to 20 μL with 1 $\times$ RNA folding buffer. Incubate samples for 20 min at 37 °C to allow peptide binding. The peptide:RNA ratio yielding optimal levels of RNA modification must be empirically determined.
5. Add 1 μL freshly prepared 20 mM magnesium monoperoxyphthalate (MMPP) solution to each sample. Incubate an additional 20 min at 37 °C.
6. Terminate reaction and desalt RNA samples by spin-chromatography through Sephadex-G50 columns or ethanol precipitation.
7. Use 2–10 pmol of the recovered RNA for reverse transcription.

3.4.3 Further Processing and Data Analysis—Once exposed to the activated ATCUN peptide and recovered, treated RNA is ready for subsequent processing as required for ensemble SHAPE. Therefore, proceed with reverse transcription (Subheading 3.1.5), preparation of sequencing ladder (Subheading 3.1.6), capillary electrophoresis (Subheading 3.1.7), ShapeFinder data analysis (Subheading 3.1.8), and Data Normalization (Subheading 3.1.9).

3.4.4 RNA 2D Modeling—As with MPE-Fe(II) probing data, normalized ATCUN-peptide data can be overlaid onto RNA secondary structure maps in RNAstructure. This means of viewing the data can be helpful for identifying ATCUN-peptide-binding sites. Secondary structure maps must be generated using SHAPE or other methods.

4 Notes

1. 1M7 was synthesized according to previously published procedure [12].
2. These reagents were synthesized according to previously published procedure [26].
3. Of the RT variants we have tested, SuperScript III works best for SHAPE-related applications.
4. 5' fluorescently labeled primers can be either purchased or synthesized in-house. Primers 5'-labeled with Cy5, Cy5.5, WellRedD2 (Beckman Coulter), and IRDye800 (LI-COR) (Beckman Coulter) are best suited for the Beckman Coulter 8000 CEQ. Primers should be designed to hybridize near the RNA 3' terminus or within the structure cassette in order to maximize the number of nucleotides for which SHAPE data can be obtained. Store in amber test tubes in water at -20 °C in 10 µM aliquots.
5. Of the commercial kits we have tested, the USB Cycle Sequencing kit works best for this application.
6. All equipment, accessories and reagents are supported by Beckman Coulter for use with their CEQ 8000 Genetic Analyzer. We have tried using other accessories and reagents with equipment, but have generally found the results to be inferior.
7. 1M7 should be stored in a desiccator at 4 °C and allowed to warm gradually to room temperature prior to use.
8. Precision in reaction time duration is not absolutely essential here, as 1M7 is over 90 % reacted with RNA or hydrolyzed after 5 min.
9. Do not dry the RNA by vacuum centrifugation. Over-drying the pellet may make it difficult to resuspend.
10. The first three steps of the program serve to anneal the labeled primers to the RNA template.
11. It is our experience that incubation for longer than 50 min can result in aberrant cDNA products.
12. Omitting this step results in poor separation of cDNA products.
13. Centrifugation at a higher speed or for a longer period may result in difficulties in resuspending the pellet(s) and/or coprecipitation of salt, which can adversely affect capillary electrophoresis.
14. Always vacuum dry DNA pellets and avoid over-drying.
15. Pellets may require prolonged vortexing and/or gentle heating (e.g., 65 °C for 5 min) to dissolve. Failure to adequately dissolve the pellet at this stage may result in a lack of signal or weak signal during capillary electrophoresis.

- 16.** In order to analyze the data from SHAPE experiment, two initial calibration reactions are necessary: (1) matrixing—to correct for fluorescent overlaps in multifluor runs, necessary only if you use older Beckman capillary electrophoresis instruments; (2) mobility shift—to correct for small differences in the electrophoretic mobility between fluorescent dyes.

Acknowledgements

This work was funded by the Intramural Research Program of the National Cancer Institute, National Institutes of Health, Department of Health and Human Services. The authors would like to thank Jennifer Miller for critical reading of the manuscript.

Abbreviations

ATCUN	Amino terminal copper- and nickel-binding motif
MPE	Methidiumpropyl-EDTA
ND	Non-denaturing
PAGE	Polyacrylamide gel electrophoresis
RRE	Rev response element
SHAPE	Selective 2'-hydroxyl acylation analyzed by primer extension
TAR	Trans-activation response element

References

- Deforges J, Chamond N, Sargueil B (2012) Structural investigation of HIV-1 genomic RNA dimerization process reveals a role for the Major Splice-site Donor stem loop. *Biochimie* 94:1481–1489 [PubMed: 22365986]
- Legiewicz M et al. (2008) Resistance to RevM10 inhibition reflects a conformational switch in the HIV-1 Rev response element. *Proc Natl Acad Sci U S A* 105: 14365–14370 [PubMed: 18776047]
- Wilkinson KA et al. (2008) High-throughput SHAPE analysis reveals structures in HIV-1 genomic RNA strongly conserved across distinct biological states. *PLoS Biol* 6:e96 [PubMed: 18447581]
- Merino EJ, Wilkinson KA, Coughlan JL, Weeks KM (2005) RNA structure analysis at single nucleotide resolution by selective 2'-hydroxyl acylation and primer extension (SHAPE). *J Am Chem Soc* 127:4223–4231 [PubMed: 15783204]
- Reuter JS, Mathews DH (2010) RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinform* 11:129
- Lusvarghi S et al. (2013) The HIV-2 Rev-response element: determining secondary structure and defining folding intermediates. *Nucleic Acids Res* 41:6637–6649 [PubMed: 23640333]
- Kenyon JC, Prestwood LJ, Le Grice SF, Lever AM (2013) In-gel probing of individual RNA conformers within a mixed population reveals a dimerization structural switch in the HIV-1 leader. *Nucleic Acids Res* 41:e174. doi:10.1093/nar/gkt690 [PubMed: 23935074]
- Han H, Schepartz A, Pellegrini M, Dervan PB (1994) Mapping RNA regions in eukaryotic ribosomes that are accessible to methidiumpropyl-EDTA.Fe(II) and EDTA. Fe(II). *Biochemistry* 33:9831–9844 [PubMed: 8060991]
- Popenda M et al. (2012) Automated 3D structure composition for large RNAs. *Nucleic Acids Res* 40:e112 [PubMed: 22539264]

10. McGookin R (1988) Electrophoresis of DNA in nondenaturing polyacrylamide gels. *Methods Mol Biol* 4:75–79 [PubMed: 21424628]
11. Rio DC, Ares M, Jr, Hannon GJ, Nilsen TW (2010) Nondenaturing agarose gel electrophoresis of RNA. *Cold Spring Harb Protoc* 2010:pdb.prot5445
12. Mortimer SA, Weeks KM (2007) A fast-acting reagent for accurate analysis of RNA secondary and tertiary structure by SHAPE chemistry. *J Am Chem Soc* 129:4144–4145 [PubMed: 17367143]
13. Mitra S, Shcherbakova IV, Altman RB, Brenowitz M, Laederach A (2008) High-throughput single-nucleotide structural mapping by capillary automated footprinting analysis. *Nucleic Acids Res* 36:e63 [PubMed: 18477638]
14. Das R, Laederach A, Pearlman SM, Herschlag D, Altman RB (2005) SAFA: semi-automated footprinting analysis software for high-throughput quantification of nucleic acid foot-printing experiments. *RNA* 11:344–354 [PubMed: 15701734]
15. Deigan KE, Li TW, Mathews DH, Weeks KM (2009) Accurate SHAPE-directed RNA structure determination. *Proc Natl Acad Sci U S A* 106:97–102 [PubMed: 19109441]
16. McGinnis JL, Dunkle JA, Cate JH, Weeks KM (2012) The mechanisms of RNA SHAPE chemistry. *J Am Chem Soc* 134:6617–6624 [PubMed: 22475022]
17. Byun Y, Han K (2006) PseudoViewer: web application and web service for visualizing RNA pseudoknots and secondary structures. *Nucleic Acids Res* 34:W416–W422 [PubMed: 16845039]
18. Darty K, Denise A, Ponty Y (2009) VARNAs: interactive drawing and editing of the RNA secondary structure. *Bioinformatics* 25: 1974–1975 [PubMed: 19398448]
19. Legiewicz M et al. (2010) The RNA transport element of the murine musD retrotransposon requires long-range intramolecular interactions for function. *J Biol Chem* 285: 42097–42104 [PubMed: 20978285]
20. Sztuba-Solinska J et al. (2013) Structural complexity of Dengue virus untranslated regions: cis-acting RNA motifs and pseudoknot interactions modulating functionality of the viral genome. *Nucleic Acids Res* 41:5075–5089 [PubMed: 23531545]
21. Stephenson JD et al. (2013) Three-dimensional RNA structure of the major HIV-1 packaging signal region. *Structure* 21:951–962 [PubMed: 23685210]
22. Sztuba-Solinska J, Le Grice SF (2014) Insights into secondary and tertiary interactions of dengue virus RNA by SHAPE. *Methods Mol Biol* 1138:225–239 [PubMed: 24696340]
23. Gherghe CM, Leonard CW, Ding F, Dokholyan NV, Weeks KM (2009) Native-like RNA tertiary structures using a sequence-encoded cleavage agent and refinement by discrete molecular dynamics. *J Am Chem Soc* 131:2541–2546 [PubMed: 19193004]
24. Celander DW (2001) Probing RNA structures with hydroxyl radicals. *Curr Protoc Nucleic Acid Chem* Chapter 6: Unit 6.5
25. Jin Y, Cowan JA (2007) Cellular activity of Rev response element RNA targeting metalloproteins. *J Biol Inorg Chem* 12:637–644 [PubMed: 17356872]
26. Hertzberg RP, Dervan PB (1984) Cleavage of DNA with methidiumpropyl-EDTA-iron(II): reaction conditions and product analyses. *Biochemistry* 23(17):3934–45. [PubMed: 6435669]

Fig. 1. Non-denaturing gel migration and distribution of HIV-2 RRE conformers as a function of RNA folding time (artistic depiction). During the early stages of RNA folding, the HIV-2 RRE exists as a mixture of three conformers designated A, B, and C. The relative abundance of these conformers varies with RNA folding time, with the most stable conformer C ultimately predominating.

Fig. 2. In-gel RNA detection and probing. [1] Illuminate ND gel TLC plate overlay with UV lamp (254 or 300 nm). [2] Locate RNA bands and excise them with a clean scalpel. [3] Submerge (+) and (-) gel slices in NMIA and DMSO solutions, respectively.

Fig. 3. Examining RNA tertiary interactions by through-space hydroxyl radical cleavage ($\bullet\text{OH}$) with the threading intercalator MPE. Once the secondary structure of the studied RNA is experimentally determined, an MPE cleavage experiment can be performed to provide clues regarding its 3D organization.

Fig. 4.

Peptides containing the amino terminal copper- and nickel-binding motif (ATCUN) can be used to map specific sites of metallopeptide-RNA interactions.