Target-specific DMS-MaPseq for in vivo RNA structure determination

Jonathan S. Weissman (jonathan.weissman@ucsf.edu)
UCSF and Howard Hughes Medical Institute
Silvi Rouskin (srouskin@wi.mit.edu)
Whitehead Institute and MIT
Meghan Zubradt
Weissman Lab
Paromita Gupta
Rouskin Lab
Sitara Persad
Rouskin Lab
Alan M. Lambowitz

Method Article

Keywords: RNA structure, RNA folding

Posted Date: November 7th, 2016

DOI: https://doi.org/10.1038/protex.2016.069

Abstract

Target-specific dimethyl sulfate mutational profiling with sequencing \(DMS-MaPseq) is a method suitable for the _in vivo_ investigation of RNA folding on specific or low abundance RNA species. Based on the RNA structure-specific _in vivo_ chemical modification by DMS and the encoding of resultant modifications as mismatches during reverse transcription, it is possible to easily amplify an RNA of interest using this simple and cost-effective RT-PCR strategy. Additionally, this approach can be used to investigate RNA structure heterogeneity under the same physiological conditions, including allele-specific RNA structure differences and structural variation between pre-mRNA and its mature spliced counterparts. Here we describe the target-specific DMS-MaPseq protocol, which produces libraries suitable for Illumina sequencing and yields excellent RNA structure data for multiple model systems.

Introduction

This protocol is used for the _in vivo_ DMS-based chemical probing of RNA structure using target-specific enrichment of RNA species from _S. cerevisiae_, human cells in culture, or _D. melanogaster_ ovaries, providing quantitative RNA structure data encoded as mismatches. In the first step, cells/tissues are treated with DMS, followed by total RNA extraction, DNase treatment, and rRNA depletion. The remaining RNA is used for reverse transcription with gene-specific primers and the thermostable group II intron reverse transcriptase \(TGIRT), which encodes DMS modifications as mismatches in the cDNA. Using gene-specific primers in a PCR step, the RNA region of interest is amplified with an optimal amplicon length of ~500 nucleotides. The dsDNA product is then subject to NexteraXT tagmentation to randomly fragment and add adaptors for Illumina sequencing. An alternative method that includes a unique molecular index \(UMI) at the RT step is prepared without tagmentation and used for longer-read sequencing.

Reagents

With the exception of standard laboratory reagents \(such as RNase-free or molecular-biology grade water, ethanol, buffers, salts, etc.), the following items are used this protocol: Dimethyl sulfate \(D186309-5ML, Sigma Aldrich) Beta-mercaptoethanol \(M3148, Sigma Aldrich) Isoamyl alcohol \(W205710, Sigma Aldrich) Acid phenol or Trizol \(AM9720 or 15596026, Ambion) Chloroform \(BP1145-1 Fisher Scientific) Ribo-Zero Gold rRNA Removal Kit \(species-specific product numbers, Epicentre) Hybridase Thermostable RNaseH \(H39100, Epicentre) TURBO DNase \(AM2238, Ambion) RNA Clean & Concentrator-5 columns \(R1015, Zymo Research) DNase I \(AM2222, Ambion) SUPERase Inhibitor \(AM2696, Ambion) TGIRT-III \(Ingex) 0.1 M Dithiothreitol i.e. DTT \(D9163, Sigma Aldrich), prepared the day of reverse transcription RNase H \(Y9220F, Enzymatics) DNA Clean & Concentrator-5 columns \(D4003, Zymo Research) Advantage HF 2 PCR Kit \(639123, Clontech) Novex TBE Gels \(varying percentages available, Invitrogen) SYBR Gold Nucleic Acid Gel Stain \(S11494, Invitrogen) Costar Spin-X columns \(8162, Corning) GlycoBlue Coprecipitant \(AM9516, Ambion) NexteraXT DNA Library Preparation Kit \(FC-131-1024, Illumina) PCRClean DX beads \(C-1003, Aline Biosciences) Phusion High Fidelity DNA polymerase \

(M0530S, New England Biolabs) MiSeq v2 kit \(varying lengths, Illumina) HiSeq 2500 v2 kit or HiSeq 3000/4000 kit \(varying options, Illumina)

Equipment

With the exception of standard laboratory equipment \(such as nuclease-free tubes, centrifuges, thermocyclers, incubators, water baths, etc.), the following items are used this protocol: Eppendorf Thermomixer, for use with 1.5 ml tubes Blue light \(ex. Invitrogen) Magnetic rack for 1.5 ml tubes \(ex. DynaMag-2, 12321D, Thermo Fisher Scientific) Bioanalyzer \(Agilent) or Fragment Analyzer \(Advanced Analytical Technologies) HiSeq 2500, HiSeq 4000, or MiSeq sequencer \(Illumina)

Procedure

1. In vivo dimethyl sulfate \(DMS) modification *Cautions before working with DMS: DMS is highly toxic and should only be used in a well-ventilated fume hood. We recommend using multiple layers of nitrile gloves such that the exterior layer can be removed and replaced after handling DMS directly. DMS should have a faint yellow color, that will become darker due to oxidation over time. Open a fresh bottle of DMS every six months to ensure freshness and optimal reactivity. DMS will be quenched during the following procedure by BME and resulting liquid waste should be disposed of per your local hazardous waste regulations. With the exception of the centrifugation steps and final steps of total RNA extraction, all steps should be done in a fume hood. a. _S. cerevisiae_ ● Dilute saturated overnight yeast cultures to OD600 of ~0.09, and grow at 30°C to a final log phase OD600 of 0.5-0.7 at the time of DMS treatment. ● Pre-chill a 30 ml stop solution on ice for > 1 hour before DMS treatment. Stop solution contains 15 ml 100% isoamyl alcohol, 9 ml BME, and 6 ml ddH2O. Also prepare and pre-chill a 10 ml wash solution containing 3 ml BME and 7 ml ddH2O. \(Solutions will phase separate, invert to mix directly prior to use.) ● Move 15 ml of log-phase yeast culture into a 50 ml conical and place in a rack submerged in a 30°C water bath, such that the entire 15 ml culture is just submerged. ● Pipet 750 µl of DMS \(5% v/v) into the yeast culture, stirring immediately with a serological pipet. Treat the yeast for 4 min, with consistent stirring or swirling of the culture. Culture should remain submerged in the water bath during this time. ● Pour ice cold 30 ml stop solution into yeast culture with DMS, cap, invert to mix, and place on ice. The DMS will be fully hydrolyzed by the BME at this point. ● Centrifuge 50 ml conical for 4 min at 3,500 x g at 4°C. Optional: parafilm the lids to minimize the smell of DMS if centrifuge is located outside of a fume hood. ● Decant supernatant into DMS/BME waste, will have a small pellet of yeast at the bottom. Pour in ice-cold wash solution, and shake/pipet to resuspend yeast pellet. ● Parafilm lid \(optional) and spin again for 4 min at 3,500 x g at 4°C. ● Decant supernatant into waste. Residual wash solution remaining in tube is fine, and traces of remaining BME in tube will quench any remaining DMS released upon cell lysis. ● Add 0.6 ml total RNA lysis buffer \(6 mM EDTA, 45 mM NaOac pH 5.5) to tube and resuspend yeast pellet by pipet. Move sample to 1.5 ml Eppendorf tube that contains 40 µl 20% SDS. ● Prewarm 0.65 ml of acid phenol in thermomixer at 65°C. ● Move tube with yeast/SDS mixture to 65°C thermomixer and shake at 1,400 rpm for 30 sec to ensure mixing of SDS. ● Add yeast sample to tube

with hot acid phenol. Incubate in thermomixer at 65°C, shaking for 2 min at 1,400 rpm. Note, that DMS-treated RNA can be brittle and high temperature or vortexing steps should be minimized to prevent fragmentation. ● Immediately move samples into a dry ice / EtOH bath \($_{1:1 composition) until frozen,}$ 3 min. ● Spin tubes 3 min at 20,000 x g at room temp and immediately remove top aqueous phase into new 1.5 ml tube, avoiding material at the phase interface. ● Add 0.65 ml acid phenol \(room temperature) to aqueous phase. Incubate at room temperature in thermomixer for 2-3 min, shaking at 1400 rpm. ● Spin tubes 3 min at 20,000 x g at room temp and immediately remove top aqueous phase into new 1.5 ml tube, avoiding material at the phase interface. ● Add 0.65 ml RNase-free chloroform and vortex at room temperature in thermomixer for 30 sec. Spin tubes 2 min at 20,000 x g at room temp and immediately remove top aqueous phase into new 1.5 ml tube, avoiding material at the phase interface. ● Add >1 volume isopropanol \(ex. 700 µl) to final aqueous phase. Invert to mix well and chill >30min at -30°C. ● Spin >45 min at 20,000 x g at 4°C to pellet nucleic acids. ● Remove supernatant and wash pellet in 0.75 ml 70-80% EtOH. ● Pulse spin tube to collect residual EtOH at the bottom, and remove all liquid with small pipet tip. ● Air-dry 5 min and resuspend in 50 µl 10 mM Tris pH 7.0. May need to add additional volume if the pellet is large. ● Nanodrop a 1:10 or 1:100 dilution of RNA sample. Procedure should yield >50-100 µg RNA. b. Adherent human cells ● Begin with a 15 cm2 plate of fully adherent HEK 293T cells in 15 ml media \(other adherent cell lines will work with the procedure too). ● Working a fume hood, remove 10 ml media from the plate into a 50 ml conical and add 200-300 µl DMS. Shake vigorously to ensure mixing. ● Slowly pipette the DMS/media mixture back on the plate by tilting the plate and pipetting into a corner. Note that this step has to be done very carefully to avoid lifting cells from the plate. ● Parafilm the plate, and move it to a 37°C incubator / hybridization oven for 5 min, keeping the plate in a fume hood if possible and swirling occasionally. Some cells may become detached during treatment. ● Slowly pipet out the DMS/media into waste. ● Add 10 ml of 30% v/v BME \(diluted in 1 x PBS) and collect the cells into a 15 ml conical using a scraper to scrape the cells from the plate. ● Centrifuge cells at 1000 x g at 4°C for 3 min; decant the BME solution. ● Wash the cells by adding 10 ml of 1 x PBS and repeating the centrifugation step. Decant the PBS. ● Add 1 ml Trizol to plate, lysing cells, and extracting total RNA according to the Trizol protocol \(Optional: if using higher amounts of DMS you may add 10 µl 100% BME to the Trizol). c. _D. melanogaster_ ovaries ● Begin with 100 D. melanogaster ovaries, removed via microdissection and collected in 250 µl 1 x PBS in an Eppendorf tube. ● Add 100-250 µl DMS to ovaries \(depending on desired level of modification), shaking at 500 rpm in a thermomixer for 5 min at 26°C. ● Stop the reaction with the addition of 1 ml of 30% v/v BME solution. ● Transfer suspension into a sieve, removing DMS/BME liquid. ● With ovaries still in the sieve, wash 3 x \(in a 6 well cell culture plate) with 1 ml of 30% v/v BME, followed by 2 washes with dH2O. ● After final wash, collect ovaries from sieve into an Eppendorf tube using a pipette. ● Remove residual dH2O, and resuspend in 1 ml Trizol ● Grind ovaries in Trizol. ● Extract total RNA according to Trizol protocol. **2. DNase Treatment** ● Starting with 5 µg of total RNA in 17 µl 10 mM Tris 7.0 or RNase-free H2O, add 2 µl of 10x TURBO DNase Buffer and 1 µl of TURBO DNase to produce at 20 µl reaction. Mix well. ● Incubate 30 min at 37°C. ● Use RNA Clean & Concentrator-5 columns to clean-up the reaction, following the manufacturer small RNA elimination procedure. Elute sample from the column with 6 µl RNase-free water. Add another 6 µl of water to the column and elute again into the same tube, collecting ~11 µl of final

volume. **3. rRNA depletion** a. Ribo-Zero ● Using entire DNase-treated RNA sample \(<5 µg), follow protocol instructions for the depletion of rRNA. ● Instead of a final EtOH precipitation as described in the Ribo-Zero protocol, the final supernatant removed from beads was cleaned by RNA Clean & Concentrator-5 column, using the small RNA elimination procedure and eluted 2 x 6 µl RNase-free H2O as above. b. RNase H In the event that Ribo-Zero is not available for a particular species or is not cost-effective, the following RNase H protocol may be used. Some troubleshooting may be required when using new RNA subtraction oligos. ● Begin with <5 µg of DNase-treated RNA, which has been depleted of small RNA species by RNA Clean & Concentrator-5 columns. ● Combine RNA with 5 µg of published RNase H oligos[1] in a total volume of 30 µl 1x Hybridization buffer \(200 mM NaCl, 100 mM Tris pH 7.5). ● Incubate reaction at 68°C for 1 min in a thermocycler, then ramp down the temperature at a rate of 1°C / min down to 45°C. ● Add MgCl2 to a 10 mM final concentration, and add 3 µl of Hybridase Thermostable RNase H, followed by a 30 min incubation at 45°C. ● Purify the reaction by RNA Clean & Concentrator-5 column with the small RNA elimination procedure, eluting once in 6 µl RNase-free water. ● Use DNase I \(Ambion) by following manufacturer instructions and a final RNA Clean & Concentrator-5 column clean-up to remove excess RNase H subtraction oligos. **4. Reverse Transcription** ● Start with 20-100 ng of DNase-treated and rRNA-depleted RNA in 5 µl 10 mM Tris pH 7.0 for a 10 µl reaction volume in a 0.2 ml PCR tube. ● Add 2 µl 5x First Strand Synthesis buffer \(final reaction concentration of 50 mM Tris-HCl pH 8.3, 75 mM KCl, 3 mM MgCl2), 1 µl 10 M gene-specific RT primer \(Tm > 60.5°C), 1 µl 10 mM dNTPs, 0.5 µl 0.1 M DTT, 0.5 µl \(i.e. 100 U) TGIRT-III, and 0.5 µl SUPERase Inhibitor. Mix. *The 0.1M DTT stock is prepared freshly from powder the day of reverse transcription. Additionally, fresh aliquots of all RT components, including the primer, should be used when possible. *It is possible to combine multiple gene-specific primers in a single RT reaction. We have tested up to five in combination, adding 0.5 µl of each RT primer and excluding additional water. *Reverse transcription can also be completed with an anchored oligo dT\(42) primer, recommended in a 20 µl reaction. ● Incubate reaction at 57°C for 2 hours in a thermocycler with a heated lid. ● Move reaction to ice, and add 1 µl of RNase H \(Enzymatics) to release cDNA from RNA:DNA hybrids. Mix. ● Incubate RNaseH reaction for 20 min at 37°C. *RNA hydrolysis by the addition of 1 µl 5M NaOH and incubation at 95°C for 3 min can be used instead of RNase H. ● Optional: Purify cDNA using the ssDNA procedure for DNA Clean & Concentrator-5 columns, eluting 2 x 6 µl with RNase-free water **5. PCR** Reactions will be described for 50 µl total reaction volume, but this can be scaled down to 20 µl. Additional options/variations of the PCR protocol can be found under the troubleshooting section of this protocol. PCR primers should be designed with Tm > 62°C. The reverse PCR primer used here may be the same primer as used for reverse transcription above, and the ideal PCR product size will be 300-600 nucleotides in length and begin only slightly 5’ of the RT primer binding site. Longer amplicons \(~2 kb) can be amplified successfully if highly abundant, ex. rRNA. To assess the success of a PCR amplification, we recommended running two PCR reactions in parallel—one that is limited in cycle number according to the following instructions and another with a high number of PCR cycles that can be easily visualized on an agarose or TBE polyacrylamide gel. ● Start with 1 µl from the purified cDNA above \(~ 1/12 of the sample). ● Set up 50 µl PCR reaction using the Advantage HF 2 PCR kit and its included HF buffer, dNTPs, and 1 µl 50 x HF polymerase. Add 1 µl 10 M of each PCR primer, using fresh primer aliquots for optimal PCR amplification. PCR Program Initial

denaturation: 94°C, 1 min 10 cycles with the following settings: Denaturation: 94°C, 30 sec Extension: 65°C, 1.5 min 20-25 cycles with the following settings: Denaturation: 94°C, 30 sec Extension: 57°C, 1.5 min ● Run 5 µl of each PCR product on 1% Agarose gel to validate the expected product size. Will pool PCR products from remaining reaction volume, not used for the gel. ● Combine all desired PCR products \(in proportion to the band intensities on the agarose gel) to be sequenced into one tube. ● PCR clean up using DNA Clean & Concentrator-5 columns, proceed to step 6. Optional: If there are unexpected bands during the PCR or the bands are highly contaminated with smears, gel purify the PCR products using the following protocol: ● Run 3 µl of PCR product on 8% TBE polyacrylamide gel at 180 V for 50 min. \(Run time will depend on size of dsDNA product, and amount of input material will vary based on robustness of PCR reaction). May run samples from a low and high-number of PCR cycles side-by-side to observe clear band with high amplification, but cut the gel region of corresponding band size from the low amplification product. Note: it is also possible to purify from an agarose gel, particularly if the PCR reaction does not produce other contaminating species. ● Stain gel with SYBR Gold, and cut PCR bands using a blue light. ● To extract PCR fragment from TBE gel, nest a 0.5 ml Eppendorf tube in a 1.5 ml Eppendorf tube with two holes poked in to the bottom of the smaller tube using a 20-gauge needle. Place the cut gel piece into the small 0.5 ml tube, and spin the nested tubes at room temperature for 3 min at 20,000 x g. ● Move any remaining gel pieces into the bottom of the 1.5 ml tube, and discard the 0.5 ml tube. ● Add 300 µl 0.3 mM NaCl and shake in a thermomixer at 70°C for 10 min with vigorous shaking \(1400 rpm). Gel slurry can also be left rotating overnight at 4°C. ● Transfer the gel slurry to a Costar Spin-X column, using a 1 mL pipet tip with the end trimmed using a sterile razorblade. Spin Costar column 3 min at 20,000 x g. ● Transfer eluate into fresh 1.5 ml Eppendorf tube. ● Add 800 µl 100% IsOH and 2 µl GlycoBlue. ● Precipitate nucleic acids by placing tube at -30°C for 30 min or more. ● Spin 45 min, 4°C at max speed \(20,000 x g). ● Remove supernatant. Wash pellet gently with 500 µl 70-80% EtOH and allow pellet to air dry 5 min. ● Resuspend pellet in 5 µl 10 mM Tris pH 8.0. **6. Fragment and add adaptors for sequencing** ● Use 1 ng of dsDNA from PCR for NexteraXT tagmentation, following the manufacturer protocol. ● Barcode and amplify tagmentation products using 12 cycles of PCR, using Nextera indices for multiplexing. ● Barcoded libraries were cleaned using 1.5x \(v/v) PCRClean beads, producing a library of 200-300 nucleotides in size. ● Libraries were quantified by Fragment Analyzer bioanalysis, pooled for multiplexing if needed, and sequenced on a HiSeq 4000 or other Illumina Nextera-compatible sequencer, using 50 bp single-end reads. **7. Variation of protocol to include a unique molecular index \(UMI)** It may be desirable to use a UMI when it is necessary to use many \(>35) cycles of PCR to amplify a low abundant RNA species. The UMI, combined with unique mismatch profiles from DMS modification itself, can be used to collapse to unique reads and remove any biases generated during rounds of PCR. ● Follow the protocol as above through reverse transcription, with one adjustment. The gene-specific RT primer should be designed with 10 random nucleotides on the 5’ end, preceded by a primer-binding overhang that will be used in the first round of PCR \(see the oMZ282 primer sequence). The Tm of the gene-specific binding region should still be > 60.5°C. ● After RT, use a forward gene-specific PCR primer and oMZ283 for the reverse primer, which complementary to the RT primer overhang. Amplicon size should be chosen based on the sequencing read length used downstream. For example, for a MiSeq 2x150 bp run, amplify PCR regions ~160bp in length to read through entire region in both

directions that will remove sequencing errors. Follow the PCR protocol using Advantage HF 2 polymerase as above. ● Amplicons were purified by TBE polyacrylamide gel and extracted with a Costar column as described above. ● Using 1 µl of dsDNA from the first PCR, complete a second round of PCR using Phusion in an 18 µl reaction \(3.4 µl 5x HF buffer, 0.4 µl 10 mM dNTPs, 0.2 µl Phusion, 11.9 µl RNase-free H2O, 0.2 µl 100 M oMZ284, and 1 µl 10 M indexing oligo) to add indices and adaptors for Illumina sequencing. Libraries were constructed so the UMI was sequenced first using custom Read1 sequencing primer oNTI202. PCR Program Initial denaturation: 98°C, 30 sec 20-25 cycles with the following settings: Denaturation: 94°C, 15 sec Annealing: 55°C, 5 sec Extension: 65°C, 10 sec ● Run out PCR product on 8% TBE gel and gel extract with Costar column as described previously. Resuspend pellet in 5 µl 10 mM Tris pH 8.0 and quantify by bioanalysis before sequencing.

Timing

3 days.

Troubleshooting

Certain PCR products may be more difficult to amplify than others. Decreasing the DMS modification level, either by altering time or concentration by up to two-fold, may be sufficient to generate a PCR product while still retaining excellent DMS signal. Given low levels of fragmentation that may occur on DMS-treated RNA, decreasing the length of the PCR amplicon could prove helpful, particularly for low abundance RNA targets. For G-rich regions that have the potential to form G-quadruplexes in vitro that might disrupt reverse transcription, we recommend using a higher salt buffer comprised of 200 mM NaCl instead of KCl as described in the protocol. Finally, if needed, one may switch to the Advantage GC 2 PCR kit \(639119, Clontech) instead of the Advantage HF 2 PCR kit. If using a low fidelity polymerase, it will also be important to amplify a non-DMS treated sample to assess any background signal produced during PCR. More generally, we do not use an untreated or denatured RNA structure sample to correct for our DMS-MaPseq signal, but it may prove useful in the case of a low fidelity polymerase, the identification of SNPs or endogenous modifications, or the empirical identification of PCR amplicon boundaries after sequencing.

Anticipated Results

This protocol will produce gene-specific libraries ready for sequencing on an Illumina platform.

References

1. Adiconis, X. et al. Comparative analysis of RNA sequencing methods for degraded or low-input samples. Nat. Methods 10, 623–629 \(2013).

Figures

name	purpose	sequence (5' to 3')
oMZ282	Reverse transcription and reverse primer (1st PCR) for targeted amplification with a unique molecular index	GCAGCGACAGGTTCAGAGTTCTACAGTCCGACGATC – $(N)_{10}$ – gene-specific primer
oMZ283	Forward primer (1st PCR) for targeted amplification with a unique molecular index	CTGAACCGCTCTTCCGATCT–gene-specific primer
oMZ284	Forward primer (2nd PCR) for targeted amplification with a unique molecular index	CAAGCAGAAGACGGCATACGAGACGGTCTCGGCAT TCCTGCTGAACCGCTCTTCCGATCT
Indexing primer with 6bp TruSeq index		aatgatacggcgaccaccgagatctacacgatcggaagagcacacgtctgaactcca gtcacNNNNNNcgacaggttcagagttc
oNTI202	Read1 sequencing primer	CGACAGGTTCAGAGTTCTACAGTCCGACGATC

Figure 1

Table 1 Primers for UMI Protocol Variation