Nucleic Acids Research, 2024, **52**, 940–952
https://doi.org/10.1093/nar/gkad1159
Advance access publication date: 12 December 2023
Structural Biology

Visualizing a two-state conformational ensemble in stem–loop 3 of the transcriptional regulator 7SK RNA

Momodou B. Camara[1], Bret Lange[1], Joseph D. Yesselman[1,2] and Catherine D. Eichhorn[1,2,*]

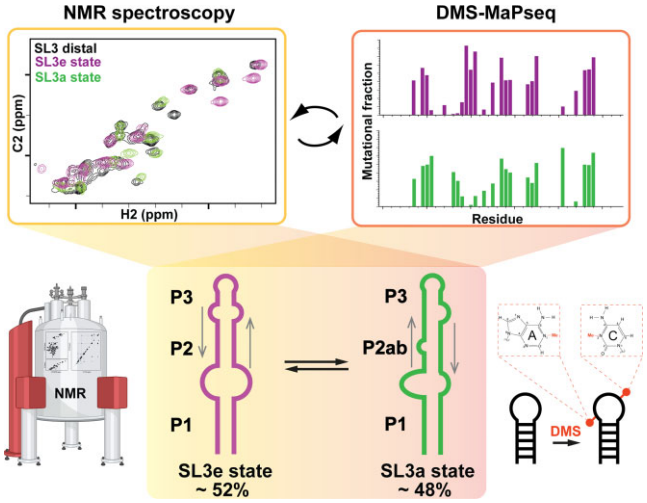
[1]Department of Chemistry, University of Nebraska, 639 North 12th St, Lincoln, NE 68588, USA
[2]Nebraska Center for Integrated Biomolecular Communication, Lincoln, NE, USA

*To whom correspondence should be addressed. Tel: +1 402 472 2426; Email: ceichhor@unl.edu

Abstract

Structural plasticity is integral to RNA function; however, there are currently few methods to quantitatively resolve RNAs that have multiple structural states. NMR spectroscopy is a powerful approach for resolving conformational ensembles but is size-limited. Chemical probing is well-suited for large RNAs but provides limited structural and kinetics information. Here, we integrate the two approaches to visualize a two-state conformational ensemble for the central stem–loop 3 (SL3) of 7SK RNA, a critical element for 7SK RNA function in transcription regulation. We find that the SL3 distal end exchanges between two equally populated yet structurally distinct states in both isolated SL3 constructs and full-length 7SK RNA. We rationally designed constructs that lock SL3 into a single state and demonstrate that both chemical probing and NMR data fit to a linear combination of the two states. Comparison of vertebrate 7SK RNA sequences shows either or both states are highly conserved. These results provide new insights into 7SK RNA structural dynamics and demonstrate the utility of integrating chemical probing with NMR spectroscopy to gain quantitative insights into RNA conformational ensembles.

Graphical abstract



Introduction

RNA secondary structure rearrangements are critical for cellular function (1–4); however, there are at present few methods to quantitatively resolve these dynamic ensembles. NMR spectroscopy is a powerful method for resolving conformational ensembles at atomic resolution (2,5–7) but can be challenging for large RNAs. In contrast chemical probing, particularly mutational profiling (MaP)-based chemical probing and next generation sequencing, is a powerful technique for identifying and deconvoluting conformational ensembles in large RNAs (8–12). Many MaP techniques have been developed (4); however, few of the identified RNA conformational ensembles have been validated by orthogonal approaches. At present, MaP techniques are limited in their ability to measure kinetics or resolve local differences in structure (4,10).

The noncoding 7SK RNA assembles with proteins to form the 7SK ribonucleoprotein (RNP) to negatively regulate eukaryotic transcription elongation through sequestration and inactivation of the positive transcription elongation factor b (P-TEFb) (13–17). The emerging mechanism of P-TEFb release involves large-scale structural remodeling of 7SK RNA, which is proposed to be driven through protein association (11,18,19). 7SK RNA folds into four primary stem–loop (SL) domains (Figure 1A). Several 7SK RNA secondary structure

Received: August 23, 2023. **Revised:** November 15, 2023. **Editorial Decision:** November 15, 2023. **Accepted:** November 17, 2023

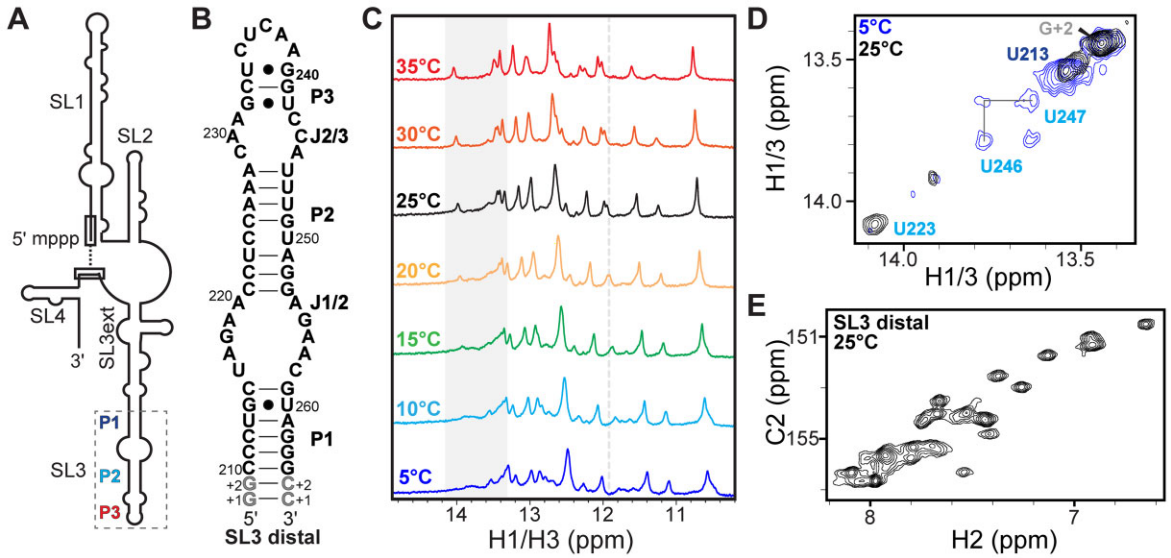


Figure 1. 7SK RNA SL3 construct design and NMR evidence of conformational exchange. (**A**) Cartoon schematic of 7SK RNA secondary structure; (**B**) SL3 distal domain (nts 210–264) constructs for NMR studies; (**C**) 1D ^{1}H imino proton NMR spectra of SL3 distal domain show new resonances that appear at elevated temperatures, colored with solid gray bars and dashed gray line; (**D**) 2D ^{1}H–^{1}H NOESY spectra at 5°C (blue) and 25°C (black) of imino region colored with a gray bar in panel C; (**E**) 2D ^{1}H–^{13}C HSQC spectrum of adenine C2H2 resonances shows over 20 resonances compared to the expected 16 resonances, indicative of slow exchange between multiple states.

models have been reported from chemical and enzymatic probing or bioinformatics techniques (11,13,15,18–23). In particular, multiple secondary structures have been proposed for the distal end of SL3, an element required to recruit accessory proteins to promote P-TEFb release on specific cellular cues (13,19,24–28). Given the functional significance of SL3, there is an urgent need to characterize SL3 structural features to resolve these discrepancies.

Here we use solution NMR spectroscopy, optical melting, and dimethyl sulfate mutational profiling with sequencing (DMS-MaPseq) chemical probing approaches to characterize the distal end of SL3. We find that SL3 slowly exchanges between two nearly equally populated secondary structural states. Using a domain clustering approach, we show that this two-state exchange is present in SL3 of an isolated domain construct as well as *in vitro* full-length 7SK RNA. Using rational design, we generate subdomain and point mutation constructs that lock SL3 into individual states. We demonstrate that the SL3 distal DMS-MaPseq data can be explained by a linear combination of the two states, providing strong evidence that these secondary structures are in fact populated. A survey of vertebrate 7SK RNA SL3 sequences shows broad conservation of either or both states, providing insights into the sequence determinants for two-state exchange equilibrium. Our results reconcile disparities in the literature by showing that rather than a single state, SL3 exists as an ensemble of two secondary structural states. Further, these studies demonstrate the utility in applying combined NMR and chemical probing approaches to quantitatively resolve RNA conformational ensembles.

Materials and methods

Sample preparation

Chemically synthesized DNA templates of SL3 domain constructs were purchased from Integrated DNA Technologies (IDT) with a 2′ O-methyl modification for the two nucleotides at the 5′ end (Supplementary Table S1). The DNA template for the wild-type full-length 7SK RNA sequence (NCBI Accession NR_001445.2), containing an upstream T7 RNAP promoter sequence, was cloned into a pUC19 plasmid (Supplementary Table S2). Site directed mutagenesis of this plasmid was performed to generate 7SK RNA constructs with point mutations. Linear DNA template was obtained from PCR amplification of the DNA template region and purified using the DNA clean and concentrator spin-column kit (Zymo). RNA samples were prepared by *in vitro* transcription (IVT) using T7 RNA polymerase (Addgene #124138 (29), prepared in-house). T7 RNA Polymerase, 500 nM DNA template, transcription buffer (40 mM Tris pH 8, 1 mM spermidine, 0.01% Triton-X, 40 mM $MgCl_2$, 2.5 mM DTT, 20% DMSO), and 2 mM each of either unlabeled rATP, rCTP, rUTP, rGTP (MP Biomedicals) or uniformly ^{13}C,^{15}N-labeled rATP, rCTP, rUTP, rGTP (Cambridge Isotope Laboratories) were incubated at 37°C for 6–8 h. Distal RNA constructs were purified by 15% denaturing polyacrylamide gel electrophoresis (PAGE) and full-length 7SK RNA constructs were purified by 5% denaturing PAGE. The RNA band was visualized by UV shadowing with a handheld UV lamp at 254 nm. After band excision, RNA was eluted from the gel using the 'crush and soak' method (30) by incubating gel pieces in crush and soak buffer (300 mM sodium acetate pH 5.2, 1 mM EDTA) for 24–48 h at room temperature. RNA was further purified to remove acrylamide contaminants by ion-exchange chromatography using a diethylaminoethanol (DEAE) column (GE Healthcare) and elution into buffer (10 mM sodium phosphate pH 7.6, 1 mM EDTA, 1.5 M KCl). RNA was diluted to <100 μM in ultrapure water and annealed by heating to 95 °C for 3 min, followed by snap cooling on ice for 1 h. RNA was then buffer exchanged into the appropriate buffer using a 3–10 kDa Amicon concentrator (Millipore Sigma).

SL3 distal constructs for chemical probing were prepared using established methods (31). Constructs were designed

with flanking 5′ GGAAGATCGAGTAGATCAAA 3′ and 5′ AAAGAAACAACAACAACAAC 3′ sequences for respective primer annealing during downstream PCR of dsDNA and reverse transcription of cDNA (Supplementary Table S3). Primers for PCR assembly of the DNA templates were designed using Primerize (32). PCR was performed using Q5 DNA Polymerase (NEB) followed by purification using DNA clean and concentrator spin-column purification (Zymo). RNA was transcribed using IVT as described above, purified using gel excision, and stored at –80°C in ultrapure water until use.

NMR spectroscopy

Solution NMR spectroscopy experiments were performed at 278.15 and 298.15 K on a Bruker Neo 600 MHz NMR spectrometer equipped with a triple-resonance HCN cryoprobe. NMR samples were prepared in NMR buffer (20 mM sodium phosphate, 50 mM KCl, pH 6.0) with added 5% D_2O at 0.1–0.8 mM concentrations in 3 mm NMR tubes (Norell). Exchangeable (H1, H3, H41, H42) and nonexchangeable (H2, H5, H6, H8, H1') proton resonances were assigned using 2D ^{1}H–^{1}H NOESY spectra of unlabeled RNA samples with mixing times of 150 ms, 200 ms, and 250 ms using SD_noe11ezg (33) and noesyesggph pulse sequences. ^{1}H–^{15}N HSQC and ^{1}H–^{13}C HSQC spectra were collected using $^{13}C/^{15}N$ labeled RNA samples. Data were processed using NMRPipe (34) and analyzed using NMRFAM-Sparky 1.470 powered by Sparky 3.190 (35) in the NMRbox virtual machine (36). Weighted average chemical shift perturbations (CSP) were calculated using the equation $\sqrt{\Delta H^2 + 0.1\Delta N^2}$ (37). Imino resonances of SL3a and SL3e states from ^{1}H–^{15}N HSQC spectra were used to determine populations from peak integration following established protocols (38) at 25°C (residues U247, G252, G253, G259) and 37°C (residues G252, G253).

Circular dichroism (CD) spectroscopy

CD experiments were performed on a Jasco815 spectrometer equipped with a Peltier temperature control device. RNA samples were prepared at 25 μM in CD buffer (20 mM sodium phosphate, 50 mM KCl, 0.1 mM EDTA, pH 6.0). Thermal unfolding experiments were performed in triplicate with a temperature range of 2–100°C (forward) or 100–5°C (reverse) and a ramp rate of 1°C/min, 2 nm bandwidth, and molar ellipticity values measured at 258 nm every 1°C. Data was fitted and visualized using an in-house Python script adapted from the Delta-melt program developed by Al-Hashimi and coworkers (39,40). Data were fitted considering a two-state folding model (41) where the mean ellipticity is:

$$[\theta]_{obs} = \left((1 - f_{unfold}) * [\theta]_{fold}\right) + \left(f_{unfold} * [\theta]_{unfold}\right) \quad (1)$$

Assuming the mean ellipticity value as a function that changes linearly with temperature, this equation can be expanded into:

$$[\theta]_{obs} = (1 - f_{unfold}) * (a_{fold} + (b_{fold} * T)) + (a_{unfold} + (b_{unfold} * T) * f_{unfold}) \quad (2)$$

Where:

$$[\theta]_{fold} = (a_{fold} + (b_{fold} * T)) \text{and} [\theta]_{unfold} = ((a_{unfold} + (b_{unfold} * T) \quad (3)$$

a_{fold}, b_{fold}, a_{unfold} and b_{unfold} are linear coefficients that represent the slope and intercept of the RNA folded or unfolded states as a function of temperature, T is the temperature (K), and f_{unfold} is the fraction of the RNA folded/unfolded. f_{unfold} can be determined using the equation:

$$f_{unfold} = \frac{exp\left(\frac{1}{Tm} - \frac{1}{T}\right)\left(\frac{\Delta H^{\circ}}{R}\right)}{1 + exp\left(\frac{1}{Tm} - \frac{1}{T}\right)\left(\frac{\Delta H^{\circ}}{R}\right)} \quad (4)$$

where T_m is the melting temperature (K), ΔH° is the enthalpy change, and R is the gas constant.

DMS-MaPseq

DMS-MaPseq experiments were performed using established protocols (31). RNA samples were annealed by heating at 90°C in water for 2 min then 4°C for 5 min, followed by addition of 1× folding buffer (100 mM HEPES, pH 8.0) at 25°C. A 15% DMS (Sigma Aldrich) stock solution in ethanol was used for DMS-MaPseq experiments. 7.5 pmol RNA was incubated with 0.75–1.5% DMS solution at 25°C or 37°C for 6 min, then quenched with 25 μl of 100% β-mercaptoethanol (BME). As a control, reactions were performed using an equivalent volume of ethanol to estimate mutations induced by reverse transcriptase enzyme in the absence of DMS. DMS-treated samples were purified using RNA clean and concentrator-5 spin-column kits (Zymo). Complementary DNA (cDNA) was generated using either TGIRT-III (In-Gex) or Marathon (Kerafast) reverse transcriptases. RNA was reverse transcribed using 100 mM DTT, 10 mM dNTPs, RT buffer (250 mM Tris–HCl pH 8.0, 375 mM KCl, 15 mM $MgCl_2$), primers with custom barcodes for downstream demultiplexing (Supplementary Table S3) and TGIRT-III or Marathon RT enzyme. For TGIRT-III, the reaction was incubated at 57°C for 2 h. For Marathon, the reaction was incubated at 42°C for 3 h. The reaction was quenched by adding 5 μl of 0.4 M NaOH to a 12 μl RT reaction (final concentration of 167 mM NaOH), then heated at 90°C for 2 min and cooled at 4°C, followed by neutralization with the addition of 2.5 μl quench acid (5 M NaCl, 2 M HCl, 3 M NaAc). To optimize the neutralization procedure, the pH at each step was measured using pH paper for samples lacking nucleic acid. The cDNA was purified using Zymo oligo clean and concentrator-5 kit (Zymo). To prepare dsDNA for sequencing, purified cDNA was used as a template for PCR using Q5 DNA Polymerase (NEB) and forward and reverse sequencing primers (Supplementary Table S3). For SL3 distal constructs, PCR was performed using 16 cycles with a 15 s extension time. For full-length 7SK RNA constructs, PCR was performed using 20 cycles with a 20 s extension time. An annealing temperature of 62°C was used for both SL3 distal constructs and full-length 7SK RNA constructs. PCR products were visualized on a 4% agarose E-gel (Invitrogen) and gel purified using an Invitrogen gel purification kit (Invitrogen) following the manufacturer's protocol. The dsDNA samples were quantified using a qubit 1x dsDNA high-sensitivity assay kit (Invitrogen) and qubit flex fluorometer (Invitrogen). The initial library was created by preparing ∼1 nM solutions of individual dsDNA samples, then pooling dsDNA samples in equimolar amounts for a total of 395 pmol in the combined dsDNA library. For the final library, 39 pmol of the initial library was mixed with 36 pmol of PhiX (Illumina) for sequencing. SL3 distal, E-lock, and A-lock samples were sequenced on

the iSeq 100 (Illumina) using iSeq 100 i1 reagent v2 kit (300 cycle) 2 × 150 read length (Illumina). The total number of reads were 150 000–400 000 per sample, with five replicates (Supplementary Figure S3). Wild-type full-length 7SK RNA, 7SK E-lock and 7SK A-lock constructs were sequenced on the miSeq (Illumina) at the UNMC Sequencing core using miSeq reagent kit v3 (600 cycles) 2 × 300 read length (Illumina). The total number of reads were 1–2 million reads per sample.

Statistical analyses and visualization

Data analysis for DMS-MaPseq was performed using DREEM as previously described in Tomezsko *et al.* (42). To account for the stochastic nature of the DREEM algorithm, we performed three iterations of DREEM clustering for each experimental replicate to calculate the average population and standard deviation. A summary of the average populations and standard deviations are reported in Supplementary Tables S6 and S7. Data was visualized using in-house python scripts (https://github.com/eichhorn-lab) that utilize SciPy (43), Pandas (44), Matplotlib (45), NumPy (46), Seaborn (47) and Biopython (48) modules. Final secondary structure models (Supplementary Table S5) were prepared using Varna (49).

Linear combination analysis

We developed an in-house Python script (https://github.com/eichhorn-lab/Linear-combination) that reconstructs the best linear fit and computes populations of SL3e and SL3a from the wild-type SL3 distal and full-length 7SK RNA dataset. Linear combination analysis was performed using the DMS mutational fractions of the SL3 E-lock and A-lock for the SL3 distal construct, or 7SK E-lock and 7SK A-lock for the full-length 7SK RNA construct, as 'fingerprints' to generate a reference dataset. Briefly, the script iteratively loops through the wild-type sequence and selects the combination of E-lock and A-lock data that minimizes the sum of residual and mean absolute error with the wild-type data using Eq. (5):

$$d = p * d_{E-lock} + (1 - p) * d_{A-lock} \quad (5)$$

where p is percent, d_{E-lock} is the DMS mutational fraction for E-lock, d_{A-lock} is the DMS mutational fraction for A-lock construct and d is the best-fit data from the combination of d_{E-lock} *and* d_{A-lock}.

To account for end fraying and sequence differences at the mutation site of locked constructs, the terminal three residues at the 5′ and 3′ ends and mutation sites in the E-lock (C224) and A-lock (C221) were set to zero in the SL3 distal, full-length 7SK RNA (nts 210–264), E-lock, and A-lock constructs. To account for variations in DMS treatment, a weighting factor was applied using eq. 6 to scale the magnitude of the reference dataset to the SL3 distal or full-length 7SK RNA (nts 210–264) data.

$$d_{L\_weighed} = \left(d_{WT} \,/\, d_L\right) * d_L \quad (6)$$

where d_L is the DMS mutational fraction for d_{E-lock} or d_{A-lock}, d_{WT} is the DMS mutational fraction for the wild-type dataset and $d_{L\_weighed}$ is the weighted dataset for d_{E-lock} or d_{A-lock}. Statistical analysis of the linear combination fit was performed using R^2 correlation and determining the residual of the fit by taking a difference between the experimental data and linearly combined fit values (Supplementary Tables S6 and S7).

Phylogenetic analysis of 7SK RNA SL3

Multiple sequence alignment (MSA) for 7SK RNA SL3 vertebrate sequences from Gruber *et al* (50) was performed using LocARNA v1.9.1 (51,52) (Supplementary Table S4). The sequence conservation from the MSA was visualized on SL3e and SL3a secondary structures using R2R (53). The dot bracket secondary structure notation in the stockholm file was modified to either SL3e or SL3a secondary structures for R2R visualization. Due to large gaps in *Lampetra* and *Petromyzon*, the line '#=GF R2R keep allpairs' was added to the stockholm file to permit generation of the secondary structure image. Suboptimal 7SK RNA SL3 secondary structure prediction for each vertebrate organism was performed using RNAsubopt 2.6.2 (54). Free energies (ΔG) of the SL3a and SL3e states in the top five predicted secondary structures were computed using RNAeval (54) with parameters of 25°C and the RNA energy model developed by Andronescu *et al* (55).

Results

NMR reveals that the 7SK RNA SL3 distal domain adopts multiple conformations

We used solution NMR spectroscopy to characterize a domain construct of the SL3 distal sequence of human 7SK RNA (SL3 distal, nts 210–264) (Figure 1B). The 1D ^{1}H spectrum showed more imino proton resonances than expected, as well as temperature-dependent line broadening, indicating base-pair related chemical exchange (Figure 1C). Due to the observed chemical exchange, the secondary structure could not be unambiguously determined from ^{1}H–^{1}H NOESY experiments. Imino proton resonances for P1 and P3 stem residues could be assigned from ^{1}H–^{1}H NOESY experiments, although P3 stem imino proton resonances were only observed at reduced temperatures (Supplementary Figure S1). In contrast, few NOEs were observed for imino proton resonances attributed to the P2 stem (Figure 1D, Supplementary Figure S1). Consistent with 1D ^{1}H NMR spectra, the ^{1}H-^{1}H NOESY spectra of the imino proton region showed a prominent on-diagonal peak at 14.1 ppm (previously assigned as U223 (20)) at 25°C but not 5°C (Figure 1D). Located in the middle of the P2 stem, U223 is not expected to undergo chemical exchange. We prepared a ^{13}C/^{15}N isotopically labeled sample and observed a greater-than-expected number of resonances in ^{1}H–^{13}C and ^{1}H–^{15}N HSQC spectra (Supplementary Figure S1). For example, although the SL3 distal construct has 16 adenine residues, over 20 C2H2 (adenine) resonances were observed (Figure 1E). A possible explanation for this observation is that the RNA adopts more than one conformation that interconvert slowly on the NMR chemical shift timescale (typically millisecond or slower (56)). The imino proton resonance line broadening for P2 stem residues, and extra nucleobase resonances, indicates that SL3 has conformational exchange on slow exchange timescales.

DMS-MaPseq shows 7SK SL3 adopts two conformations with distinct secondary structures

As an orthogonal approach to investigate the secondary structure, we performed DMS-MaPseq (10) experiments on the isolated SL3 distal construct at 25°C. Due to the lack of commercial availability of TGIRT-III enzyme, both TGIRT-III and Marathon reverse transcriptases were used with excellent agreement, as seen in prior studies (57,58)

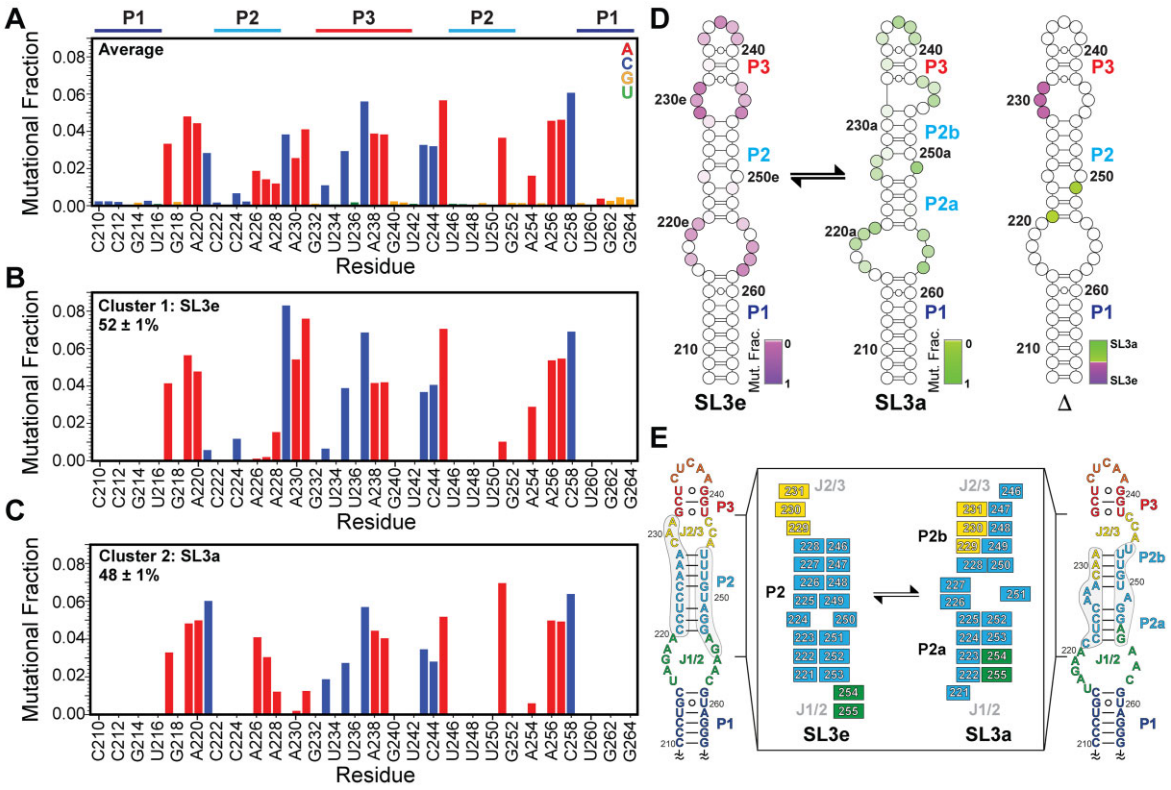


Figure 2. DMS-MaPseq and DREEM clustering supports the presence of a two-state ensemble. (**A**) Average mutational fraction of SL3 distal construct. Residues are colored by nucleotide (A, red; C, blue; G, yellow; U, green). (**B**, **C**) Mutational fraction after DREEM clustering. (**D**) Secondary structure models of SL3e (*left*), SL3a (*center*), and difference plot (*right*), colored according to DMS-MaPseq reactivity, shows key reporter residues in the P2 stem and J2/J3 loop. (**E**) SL3e and SL3a secondary structure models colored by motif (P1, dark blue; J1/2, green; P2, light blue; J2/3, yellow; P3, red). *Inset*: Schematic of differences between SL3e and SL3a secondary structures.

(Supplementary Figure S2). The secondary structure, assuming a single state, matched the secondary structure shown in Figure 1B. However, P2 stem residues C221 and A251 had intermediate mutational fraction values, suggesting that a single secondary structure may be insufficient to describe the data (Figure 2A). Next, we applied DREEM (Detection of RNA folding Ensembles using Expectation-Maximization) (42) to investigate whether the DMS-MaPseq data could be deconvoluted into two or more secondary structure states. From DREEM, two clusters were identified with near-equal populations (Figure 2B-D). Both clusters have identical P1 and P3 stems but differ significantly in the P2 stem and adjacent loops. A comparison of the mutational fraction values for the two clusters identified several key reporter residues in the center of the P2 stem and J1/2 and J2/3 loops that are present in a stem in one state and loop in the other state (Figure 2D).

The P2 stem of one cluster matched the secondary structure in Figure 1B, hereafter named SL3e for an extended P2 stem (Figure 2D, E). The P2 stem of SL3e has eight base-pairs with a central C•U mismatch. This secondary structure was observed in prior chemical probing and NMR studies (11,20). The P2 stem of the other cluster consists of two four base-pair stems, hereafter called SL3a for the intervening A-rich loop between P2a and P2b stems. A similar secondary structure was reported in early studies of 7SK RNA secondary structure (18,19,21,23) (Figure 2A). Consistent with the near-equal observed populations, free energy calculations using the RNAeval WebServer (54,59) showed similar computed free energy (ΔG) values for the two secondary structure states (Supplementary Table S4).

To evaluate the population equilibrium of the two SL3e and SL3a states at physiological temperature, we performed DMS-MaPseq at 37°C. While both SL3e and SL3a states are present, we observed an increase in the SL3a state population at 37°C relative to 25°C (Table 1, Supplementary Figure S4). This result is consistent with the NMR temperature dependence and NOESY experiments, where different imino resonances are observed at different temperatures (Figure 1C-D, Supplementary Figure S4). The two-state conformational ensemble from DMS-MaPseq data explains the chemical exchange observed in the NMR studies described above. Both SL3e and SL3a states have identical P1 and P3 stems, consistent with the strong NOE connectivities and unambiguous assignments for residues in these stems. The extra observed resonances and ambiguous assignments for resonances of P2 stem residues results from slow exchange of these residues between two distinct secondary structures. DMS-MaPseq shows equal populations of the two states at 25°C, consistent with similar resonance intensities observed in 2D HSQC spectra. Together, the combined NMR and DMS-MaPseq data show that the SL3 distal end exists as a two-state conformational ensemble.

SL3 two-state exchange is present in full-length 7SK RNA

To identify if the SL3 two-state conformational ensemble is present in the full-length 332 nt 7SK RNA, we performed DMS-MaPseq experiments using *in vitro* transcribed human 7SK RNA (Figure 3A). Strong agreement was observed for

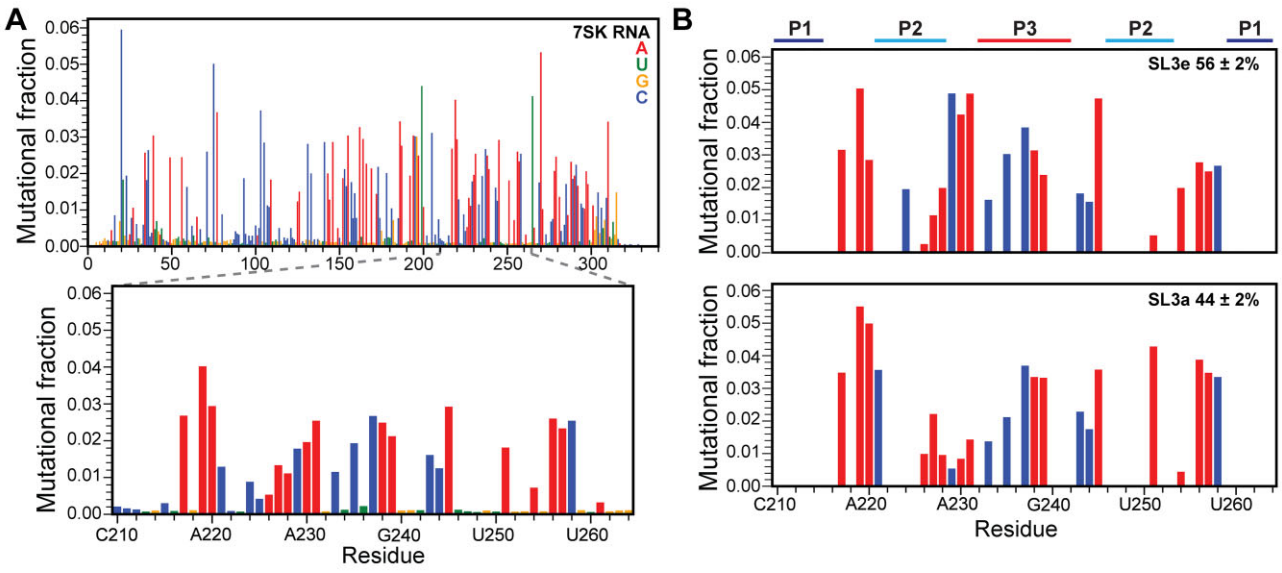


Figure 3. DMS-MaPseq shows two-state conformational ensemble of SL3 in *in vitro* full-length 7SK RNA. (**A**) *Top:* average mutational fraction for full-length 7SK RNA. *Bottom:* zoom-in panel of SL3 distal end region corresponding to the isolated domain construct (nts 210–264). (**B**) DREEM clustering shows SL3a and SL3e states with populations that are consistent with the SL3 domain construct.

the SL3 distal end between average mutational fraction values of the isolated domain and full-length 7SK RNA constructs ($R^2 = 0.89$) (Supplementary Figure S5), indicating that the isolated SL3 domain construct has similar behavior to the full-length 7SK RNA construct. Overall, the DREEM-identified secondary structure of 7SK RNA was nearly identical to previous studies using SHAPE or DMS-based chemical probing experiments, particularly SL1 and SL4 (11,18,20,21) (Supplementary Figure S6). When a single secondary structure is assumed, SL3 has a SL3a-like state, identified by the J1/2 loop and P2a stem, with melted P2b and P3 stems. Similar to the isolated SL3 domain construct, elevated reactivities were observed for P2 stem residues C221, A227, A230, A231 and A251 (Figure 3A). We initially performed DREEM clustering for the DMS-MaPseq data of the entire 332 nt 7SK RNA sequence and found that the secondary structures in the two clusters showed significant differences in the SL1 basal end and SL2 regions (Supplementary Figure S6). The observed populations and secondary structure models are remarkably consistent with previous *in vitro* DANCE-MaP studies of 7SK RNA (11). The major populated state (71%) resembles the 'linear' secondary structure identified in previous studies (18,20,21). The minor populated state (29%) is similar to the 'circular' secondary structure identified in previous studies (11,22,60), which has an additional SL0 stem from long-range base-pairing between 5′ and 3′ ends.

Nevertheless, in both clusters, the SL3 distal end shows only the SL3e state. MaP techniques are limited in their ability to deconvolute local structural differences for large RNAs (4). To address this limitation, we performed DREEM clustering of DMS-MaPseq data from the SL3 distal end (nts 210–264) alone. With this approach, we observed both SL3e and SL3a states with near-equal populations of 56 ± 2% and 44 ± 2%, respectively (Figure 3B, Table 1, and Supplementary Figure S5C). These results demonstrate that the two-state conformational ensemble is present both in the isolated SL3 construct and full-length 7SK RNA.

Rational construct design locks SL3 into individual states to quench exchange

Due to chemical exchange, P2 stem resonances of the SL3 distal construct could not be unambiguously assigned. To further probe the two conformational states and enable resonance assignment, subdomain constructs were designed for the top and bottom halves of the SL3 distal end for each state (Supplementary Figure S8A). Leveraging 5′-3′ boundary differences between the two states, constructs were designed to stabilize either SL3a (SL3a-top, nts 222–255) or SL3e (SL3e-top, nts 221–253) states (Figure 4A). 1D ^{1}H NMR spectra of imino proton resonances showed the expected number of resonances and were able to be assigned by ^{1}H–^{1}H NOESY experiments (Figure 4B and Supplementary Figure S7) indicating a single folded state. 2D ^{1}H–^{13}C HSQC spectra of C2H2 (adenine) resonances showed the expected number of resonances for each construct, further supporting the existence of a single state (Supplementary Figure S8B). Resonances in 2D HSQC spectra of all constructs showed excellent agreement with SL3 distal construct resonances (Supplementary Figure S8B). From NMR spectra of subdomain constructs, resonances in the SL3 distal construct were assigned to either the SL3e state or SL3a state (Supplementary Figure S7-S8). For example, the characteristic U223 imino proton resonance at 14.1 ppm that was observed in the SL3 distal construct at elevated temperatures is also observed in SL3a-top, denoted as U223a (Figure 4B, C). In SL3e-top, U223 is shifted upfield to ∼13.6 ppm, denoted as U223e (Figure 4C). SL3e-top showed the characteristic crosspeak between resonances ∼13.8 ppm, assigned as U246e-U247e, observed in SL3 distal at 5°C in the ^{1}H-^{1}H NOESY spectrum. Comparison of 2D ^{1}H-^{15}N HSQC spectra showed that SL3a-top and SL3e-top imino resonances overlay with subsets of SL3 distal resonances, confirming that SL3 distal has slow exchange between two states (Figure 4C and Supplemental Fig. S8C). Together, the NMR spectra of the four subdomain constructs superimpose nearly completely onto the SL3 distal construct (Supplementary Figure S8),

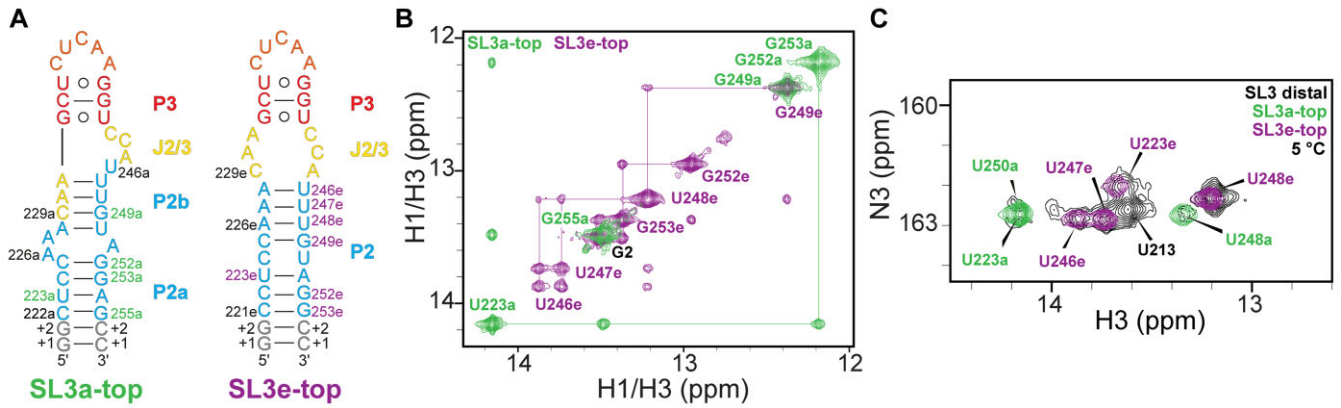


Figure 4. Subdomain constructs of the SL3 top stem lock SL3 into a single state. (**A**) Secondary structures of SL3a-top and SL3e-top constructs. Constructs are colored by secondary structure motifs in SL3e as seen in Figure 2E, with number labels colored green or red for corresponding SL3a- or SL3e-state resonance assignments. (**B**) ^{1}H–^{1}H NOESY spectrum of imino region of SL3a-top (green) and SL3e-top (purple) constructs. (**C**) ^{1}H–^{15}N 2D HSQC spectrum of uridine N3H3 resonances for SL3 distal (black), SL3a-top (green), and SL3e-top (purple) shows SL3 distal construct is a combination of SL3a and SL3e states.

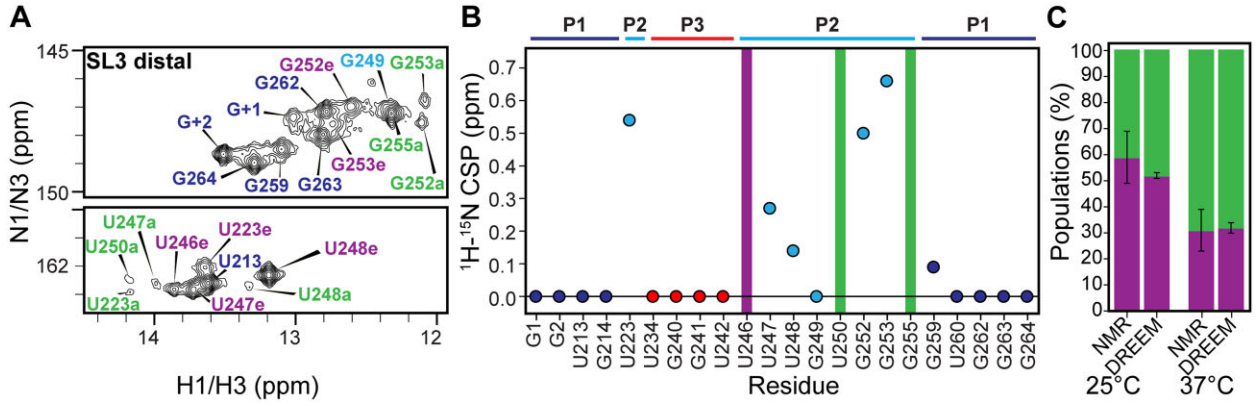


Figure 5. Comparison of SL3e and SL3a populations from NMR and DREEM analysis. (**A**) ^{1}H–^{15}N imino HSQC spectrum of SL3 distal at 5°C. Labels for P1 stem resonances are colored dark blue, SL3e state resonances are colored purple, SL3a state resonances are colored green, and P2 stem resonances that are not in exchange are colored light blue. (**B**) Plot of weighted average chemical shift perturbations of SL3e- and SL3a-state imino resonances of SL3 distal in panel A. Solid bars indicate imino resonances that are only observed in the SL3e (purple) or SL3a (green) states. (**C**) NMR and DREEM analysis of SL3 distal show excellent agreement in the determined populations of SL3e (purple) and SL3a (green) states at 25°C and 37°C.

indicating both SL3e and SL3a states are sufficient to explain the chemical exchange observed in the NMR data of the SL3 distal construct.

NMR assignments from the subdomain constructs were transferred onto the SL3 distal construct (Figure 5A). Three imino resonances were observed in one state but not the other: U246e (SL3e), U250a (SL3a), and G255a (SL3a), and chemical shift perturbations were only observed in the P2 stem (Figure 5B). We next performed peak integration of these resonances to determine populations of SL3e and SL3a states and found that the populations of SL3e and SL3a states were consistent with those determined from DREEM analysis at both 25°C and 37°C (Table 1 and Figure 5C).

To further characterize each of the SL3 conformational states, we designed point mutations to stabilize either the SL3e or SL3a state in the SL3 distal and full-length 7SK RNA constructs (Figure 6A, B). For the locked SL3e construct (E-lock), a C224A substitution replaces the C224•U250 mismatch with a canonical A224-U250 base-pair (Figure 6A and Supplementary Figure S9). This substitution disfavors the SL3a state by introducing an A•G pair in the SL3a P2a stem. DMS-MaPseq experiments showed 100% SL3e state with high mutational fraction values for reporter residues C229, A230, A231, and A254 (Figure 6C), confirming stabilization of a single state. DMS-MaPseq profiles had excellent agreement between the SL3e cluster in SL3 distal and E-lock constructs ($R^2 = 0.97$) (Figure 6D). A comparison of ^{1}H–^{15}N 2D HSQC spectra of the imino region for SL3 distal and E-lock showed good agreement to SL3e chemical shifts (Figure 6E) with only U223e and U248e, near the mutation site, showing chemical shift perturbations. In addition, a new U250 imino resonance was observed consistent with the new A224–U250 base-pair. Similarly, comparison of ^{1}H–^{13}C 2D HSQC spectra of adenine C2H2 resonances showed good agreement to SL3e chemical shifts, with a new A224e resonance and chemical shift perturbations to A251e and A226e, near the mutation site (Figure 6I). Importantly, resonances attributed to the SL3a state are not observed in the E-lock construct, consistent with DMS-MaPseq experiments showing a single SL3e state.

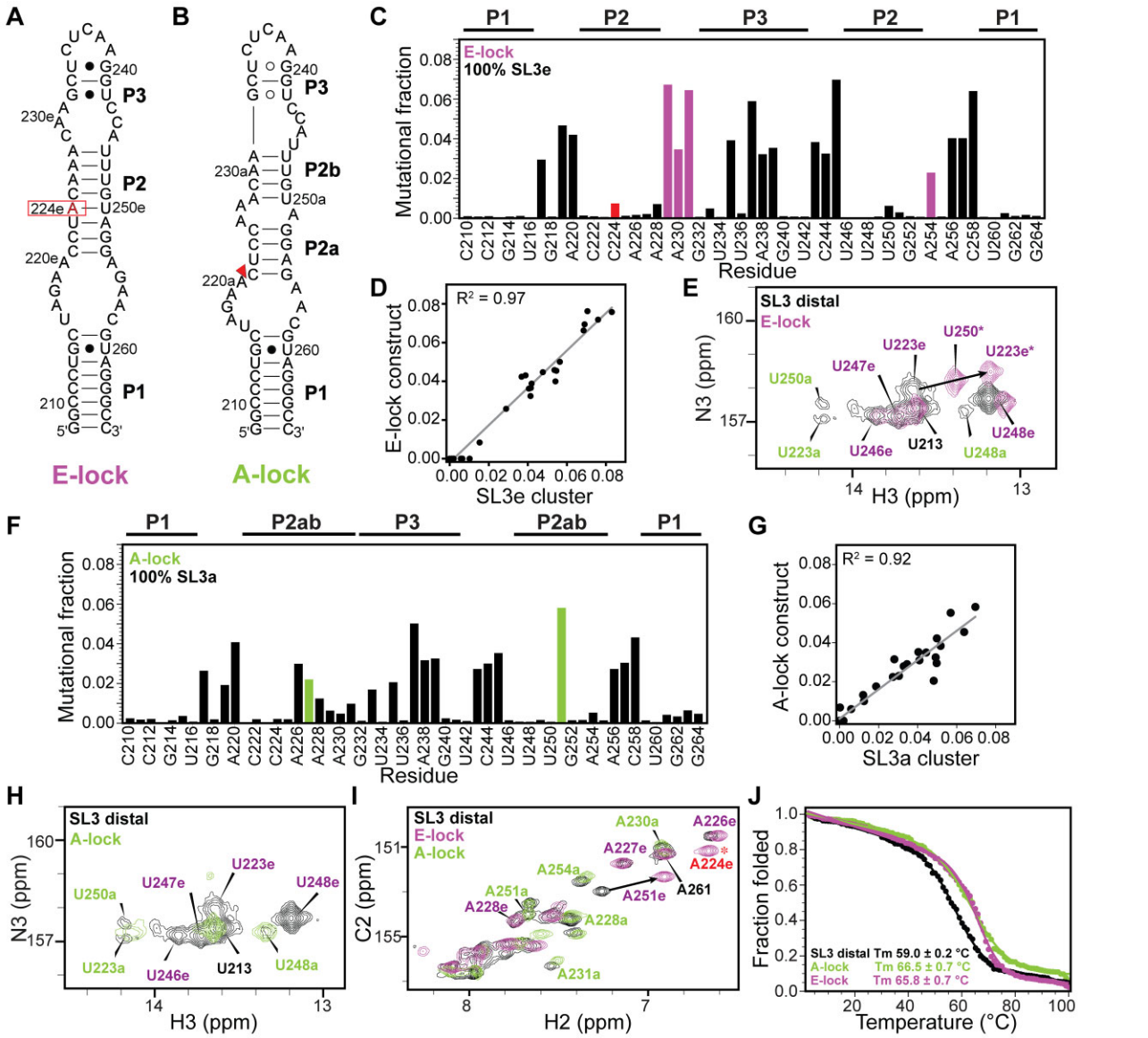


Figure 6. Point substitutions lock SL3 into single conformer. Secondary structure models of SL3 distal constructs. (**A**) E-lock and (**B**) A-lock with mutations indicated in red. (**C**) DMS-MaPseq profile of E-lock shown as bar plot. Red bar indicates point substitution, purple bars indicate reporter residues identified in Figure 2D. (**D**) Correlation plot of DMS-MaPseq data of E-lock and SL3 distal construct SL3e cluster show excellent agreement. (**E**) ^{1}H–^{15}N 2D HSQC of uridine imino resonances of SL3 distal (black) and E-lock (purple) show good agreement with superimposition of resonances corresponding to the SL3e state. SL3a resonance labels are colored green. Asterisks next to U223 and U250 resonances indicate perturbed and new chemical shift, respectively, due to C224A substitution. (**F**) DMS-MaPseq profile of A-lock shown as bar plot. Green bars indicate reporter residues identified in Figure 2D. (**G**) Correlation plot of DMS-MaPseq data of A-lock and SL3 distal construct SL3a cluster shows excellent agreement. (**H**) ^{1}H–^{15}N 2D HSQC of uridine imino resonances of SL3 distal (black) and A-lock (green) show good agreement with superimposition of resonances corresponding to the SL3a state. SL3e resonance labels are colored purple. (**I**) ^{1}H–^{13}C 2D HSQC of adenine C2H2 resonances of SL3 distal (black), E-lock (purple), and A-lock (green). New A224 C2H2 resonance is labeled in red with an asterisk. (**J**) Thermal unfolding profiles of SL3 distal (black), E-lock (purple) and A-lock (green) constructs.

To stabilize the SL3a state, we performed a deletion of C221 (A-lock) to the SL3 distal construct (Figure 6B and Supplementary Figure S10). C221 is located in the J1/2 loop in SL3a and is the closing base-pair in SL3e. We predicted that this loop deletion would minimally impact the SL3a state and disfavor the SL3e state by eliminating one base-pair. DMS-MaPseq experiments showed 100% SL3a state with high mutational fraction values for A227 and A251 (Figure 6F) and excellent agreement to the SL3a cluster in the SL3 distal construct (R^2 = 0.92) (Figure 6G). The ^{1}H–^{15}N 2D HSQC spectrum of the imino region showed excellent agreement to SL3a chemical shifts (Figure 6H), in support of formation of a single SL3a state. Similarly, inspection of ^{1}H–^{13}C 2D HSQC spectra of adenine C2H2 resonances shows excellent agreement to SL3a resonances (Figure 6I and Supplementary Figure S10) with minimal chemical shift perturbations. Resonances assigned to SL3e are not observed in the A-lock construct, consistent with DMS-MaPseq data showing a single SL3a state. E-lock and A-lock mutation constructs of full-length 7SK RNA were generated, and DMS-MaPseq experiments were performed to verify that the mutations can lock SL3 into SL3e or SL3a states within the

Table 1. SL3 populations determined from NMR, DREEM and Linear Combination analysis

Construct	Temperature	NMR		DREEM		Linear combination	
		SL3e (%)	SL3a (%)	SL3e (%)	SL3a (%)	SL3e (%)	SL3a (%)
SL3 distal	25°C	59 ± 10	41 ± 10	52 ± 1	48 ± 1	54 ± 1	46 ± 1
SL3 distal	37°C	31 ± 8	69 ± 8	32 ± 2	68 ± 2	23 ± 9	77 ± 9
A-lock	25°C	0 ± 0	100%	0 ± 0	100 ± 0	n/a	n/a
E-lock	25°C	100%	0 ± 0	100 ± 0	0 ± 0	n/a	n/a
7SK RNA	25°C	n/a	n/a	56 ± 2	44 ± 2	51 ± 0.5	50 ± 0.5

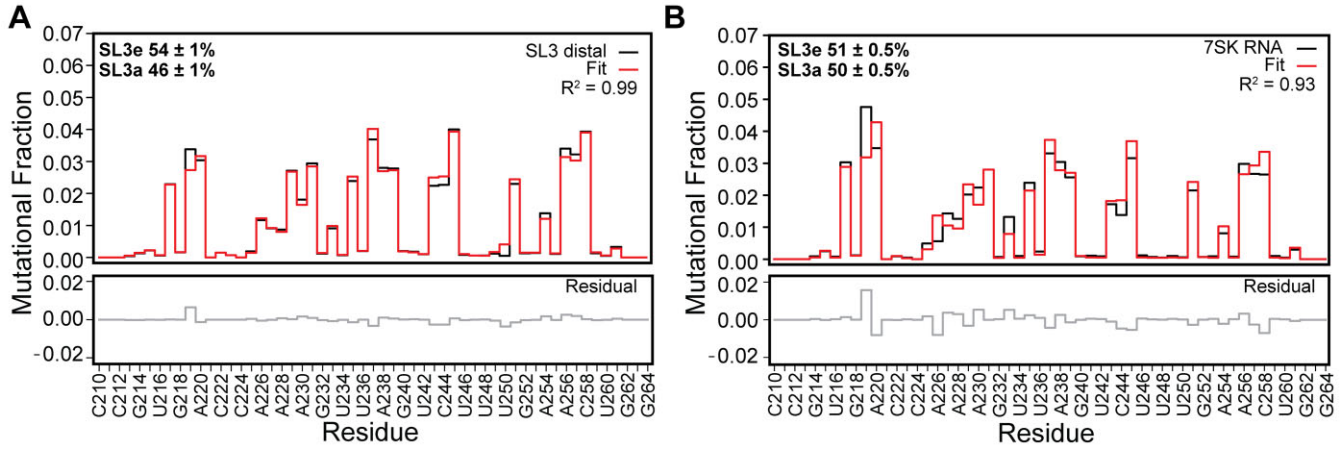


Figure 7. Linear combination analysis validates a two-state physical model. DMS-MaPseq profiles of (**A**) SL3 distal construct and (**B**) SL3 region of full-length 7SK RNA show excellent agreement to a linear combination of SL3e and SL3a states, with populations consistent with those determined using DREEM. DMS-MaPseq mutational fraction values colored black, fit line colored red and residual colored gray.

full-length 7SK RNA. An excellent correlation in the mutational fraction values was observed when comparing locked domain and full-length 7SK RNA constructs, with R^2 correlation values of 0.95 and 0.89 for E-lock and A-lock constructs, respectively (Supplementary Figures S11A, B and S12A, B). Consistent with domain constructs, DREEM clustering analysis of the 7SK RNA locked constructs shows that the 7SK E-lock construct has 100% SL3e state (Supplementary Figure S11C), while the 7SK A-lock construct has 100% SL3a state (Supplementary Figure S12C). Together, the subdomain and point mutant constructs in this study validate the presence of two distinct states in the SL3 P2 stem and demonstrate that 7SK RNA SL3 can be locked into a single conformer.

Locking SL3 distal into a single state increases thermal stability

To evaluate the thermodynamic stability of SL3 distal constructs, optical melting was performed using circular dichroism (CD) spectroscopy. A single cooperative melting transition was observed with a melting temperature (T_m) of 59.0 ± 0.2°C for the wild-type SL3 distal construct (Figure 6J and Supplementary Figure S13). Surprisingly, both E-lock and A-lock constructs showed increased melting temperatures with Tm values of 66.5 ± 0.7°C and 65.8 ± 0.7°C, respectively. The similar Tm values for the locked constructs are consistent with SL3a and SL3e states having similar free energies.

A two-state physical model is sufficient to explain NMR and chemical probing data

At present, MaP-based approaches are limited by a reliance on correlated mutations and thermodynamics-guided secondary structure prediction (4,61). We sought to analyze DMS-MaPseq data independently of DREEM to provide additional validation for the observed two-state conformational ensemble. Given the excellent fit in DMS-MaPseq profiles between SL3 distal DREEM clusters and locked constructs (Figure 6D, G), we reasoned that these mutants could be used as 'fingerprints' to represent either the SL3a or SL3e states. To evaluate whether the DMS-MaPseq data could be fit to a two-state model, we generated in-house software that computes populations of two states using linear combination. Here, the DMS-MaPseq profiles of the A-lock and E-lock constructs were used as templates for SL3a and SL3e conformers, respectively. The computed populations were consistent with those determined by NMR or DREEM analysis (Table 1, Supplementary Tables S6-S7). For the SL3 distal construct, populations fit to 54 ± 1% (SL3e) and 46 ± 1% (SL3a) at 25°C (Table 1, Figure 7A, Supplementary Table S14). For SL3 in full-length 7SK RNA, populations fit to 51 ± 0.5% (SL3e) and 50 ± 0.5% (SL3a) at 25°C (Table 1, Figure 7B, Supplementary Figure S15). The correlation between experimental and fit data was excellent, with $R^2 = 0.99$ for the isolated SL3 domain and $R^2 = 0.93$ for SL3 in full-length 7SK RNA. These results show that the DMS MaPseq data can be explained by a two-state physical model in which SL3 exists as two distinct states, and validates the populations determined using DREEM.

Both SL3e and SL3a states are conserved among vertebrates

SL3 is a conserved structural domain identified in vertebrates, hexapoda, and coleoptera (22,50,62). For the human 7SK RNA SL3 distal end, we found that the RNAeval webserver (54) computed a free energy difference (ΔΔG) between SL3e

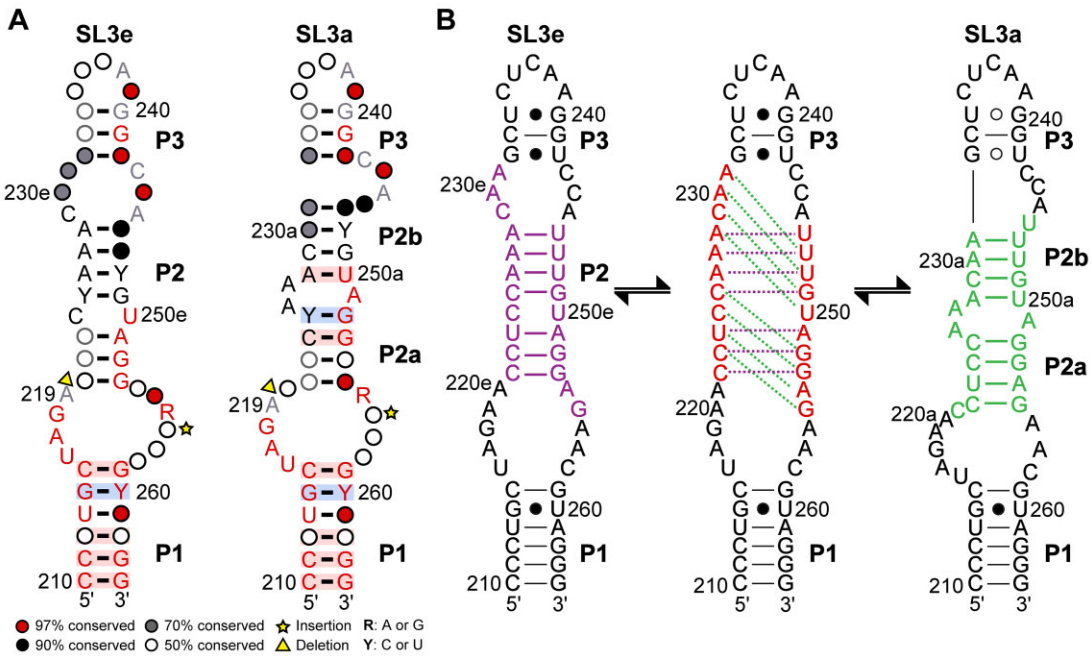


Figure 8. 7SK SL3 sequence conservation and two-state exchange model. (**A**) Sequence conservation plotted on SL3e and SL3a states using R2R (53) shows conservation of both potential states. (**B**) Proposed model of exchange between SL3e and SL3a states. Residues that are involved in base-pair rearrangements are shown in purple (SL3e), red (where SL3e base-pairs are indicated with purple dashed lines and SL3a base-pairs are indicated with green dashed lines), and green (SL3a).

and SL3a states to be approximately ∼0.4 kcal/mol, similar to the estimated ΔΔG assuming Boltzmann statistics, suggesting that the thermodynamic parameters used in the ViennaFold package perform reasonably well at predicting the energy differences between the two secondary structures.

To investigate whether the SL3 two-state conformational ensemble is specific to humans, or conserved more broadly across vertebrates, we used previously identified vertebrate 7SK RNA sequences from Gruber *et al.* (50) and RNAsubopt (54) to predict the five lowest-energy secondary structures for the SL3 distal end sequence. Consistent with prior studies, we found high sequence conservation among 7SK RNA sequences (22,50) that is compatible with both SL3a and SL3e states (Figure 8A, Supplementary Figure S16B). In particular, all mammals except for rodents and marsupials had identical sequences to human and, as expected, both SL3a and SL3e states were predicted with similar $\Delta\Delta G$ values (Supplementary Table S4, Supplementary Figure S17). In rodents, U246 is absent, resulting in the SL3a state having significantly lower energy compared to the SL3e state (Supplementary Figure S16B, S17). Consistent with this finding, the SL3a secondary structure was reported in prior chemical probing studies of 7SK RNA in rat Novikoff hepatoma cells (23) and mouse embryonic stem cells (19). In marsupials, an 8-nt deletion of P2 stem residues C222-C229 results in a unique secondary structure that is neither SL3a nor SL3e (Supplementary Figure S17). Across fishes, with the exception of *Danio rerio* (zebrafish), both states were predicted with equal energies (Supplementary Figure S16B, S17). For *Danio rerio*, which has insertions in J1/2 and J2/3 loops, only the SL3e state is predicted to form. Compared to human, *Mustelus* and *Lampetra* have substantial sequence differences and lack J1/2, J2/3 and apical loop sequences despite having respective SL3e-like and SL3a-like features in the P2 stem region. Taken together with our mutagenesis studies, the phylogenetic analysis shows remarkable sensitivity in the sequence dependence on the SL3 secondary structure.

Discussion

RNA conformational exchange and secondary structural rearrangements are essential to a wide range of cellular processes including protein recruitment and cellular function (1,63,64). Despite this significance, there are currently limited methods to measure these structural dynamics. The SL3 domain of 7SK RNA was previously identified as a conserved structural element of 7SK RNA (50,62), with prior studies reporting either SL3e (11,20) or SL3a (19,21,23) states. Using combined NMR, DMS-MaPseq, optical melting, and mutagenesis approaches, we show that both states are populated in human 7SK SL3 RNA as a two-state conformational ensemble. We demonstrate that this SL3 two-state ensemble is present in both domain and full-length 7SK RNA constructs. This local conformational exchange was masked when the entire 7SK RNA sequence was used for clustering, and was able to be detected when the SL3 distal sequence was used for clustering. This strategy may be applicable to other multidomain RNAs that have independently folded domains. This domain clustering strategy has potential to identify domains that do not fold independently or have long-range interactions, such as tertiary contacts by comparing MaP reactivity data between segments and full-length RNA constructs. Additionally, we developed a linear combination approach to validate the SL3e and SL3a states present in 7SK RNA SL3. This strategy could be applied toward other RNA multi-state ensembles; however, the rational design of locked constructs, with mutational fraction values representing each conformational state, is a prerequisite.

Prior *in vitro* and *in cellulo* chemical probing studies that identified the SL3e state showed elevated chemical reactivity in P2 stem residues (11,20), particularly residues C221 and A251 which we identify as reporter residues for the SL3a state. Our data shows that this reactivity is due to conformational exchange between SL3e and SL3a states. A previous study by Luo et al. combined DMS-MaPseq and NMR to study the SL3 domain and full-length 7SK RNA (20), where DMS-MaPseq reported the 7SK RNA secondary structure and identified conformational exchange at SL2 and the SL3 basal end (nts 153–200) and NMR reported the SL3e secondary structure and protein recognition of SL3 (nts 200–274). While the Luo et al. 7SK RNA secondary structure before clustering is consistent with our findings, the 7SK RNA secondary structure models identified in our study are more consistent with the previous DANCE-MaP study (11). For the SL3 region, overall there were very similar DMS mutational fraction patterns for the 7SK SL3 distal region between the Luo et al. study and our study, although exchange in the SL3 distal region was not observed. Our domain clustering approach, coupled with rational construct design, allowed us to detect conformational exchange in 7SK RNA SL3 in the full-length 7SK RNA. For the previous NMR study of the SL3 distal end, only the SL3e state was identified (20). However, imino resonances we identify as reporters of SL3a and SL3e states are both present in previously reported imino HSQC data (20) suggesting both states are present. In our study, unambiguous assignment was only possible through the rational design and assignment of subdomain constructs. The previous study and our study also had differences in NMR field strength and buffer conditions, which may also account for the reported differences. Our study reconciles disparities in the literature by showing that SL3 exists as a two-state conformational ensemble in human 7SK RNA, explaining why in some studies the SL3e state is identified and the SL3a state is identified in others. In addition, our phylogenetic analysis provides an explanation for mouse and rat studies that report the SL3a state.

While rearrangements of base-pairs within stems have been observed previously (65–67) there is often a substantial free energy difference resulting in a major populated state and minor alternative state. In sharp contrast, SL3 has similar free energies in the two states and near-equal populations indicating a bistable RNA at 25°C. While bistable RNAs have been rationally designed (68–71) and observed in riboswitches (72), base-pairing rearrangements typically take place between residues in an RNA hairpin and an adjacent single-stranded overhang, and in some cases strand exchange (73), not within a stem of a hairpin as seen in 7SK SL3. The transition between the two states in the SL3 distal end involves a rearrangement of eight base-pairs and is predominantly localized in the P2 stem (Figure 8B). The two states significantly differ in loop sequence and composition: while SL3e has symmetric J1/2 and J2/3 loops, SL3a has asymmetric J1/2 and J2/3 loops and a new A-rich loop in the P2ab stem (Figures 2E, 8B). These differences may be functionally relevant, as loops often serve as protein recognition sites. 7SK SL3 is required for P-TEFb release from 7SK RNP (26,27) and several proteins have been identified to bind SL3 (e.g. hnRNPA1, SRSF2, RBM7) in the P-TEFb release pathway (20,24,25,74).

In summary, this study highlights the power of combined application of NMR and chemical probing to provide quantitative insights into RNA conformational ensembles. We demonstrate that local base-pair rearrangements may be missed in current chemical probing approaches of large RNAs and anticipate that other functionally relevant RNAs may undergo similar structural transitions.

Data availability

Python scripts are deposited in GitHub (https://github.com/eichhorn-lab) and Figshare (https://doi.org/10.6084/m9.figshare.24573499). Raw demultiplexed chemical probing data have been deposited in the NCBI BioProject (PRJNA1004643). Processed chemical probing data is deposited in GitHub (https://github.com/eichhorn-lab) and Figshare (https://doi.org/10.6084/m9.figshare.24573580). Source code for the DREEM version used in this study, which performs clustering analysis, can be found in supplementary files. NMR resonance assignments have been deposited in the BMRB (ID 52092).

Supplementary data

Supplementary Data are available at NAR Online.

Acknowledgements

We gratefully acknowledge Addgene for the pQE30-His-T7RNAP plasmid used to prepare T7 RNA Polymerase enzyme for *in vitro* transcribed RNA samples, which was a gift from Sebastian Maerkl & Takuya Ueda (Addgene plasmid #124138). We thank Dr Robert Peterson and Dr Martha Morton for NMR assistance. We thank Dr Jennifer Bushing and the University of Nebraska Medical Center Genomics Core.

Author contributions: C.D.E. conceived and oversaw all aspects of the project. M.B.C., J.D.Y. and C.D.E. designed experiments and performed data analysis. M.B.C. prepared samples and performed experiments. B.L. performed library preparation for DMS-MaPseq experiments. J.D.Y. wrote the linear combination analysis software. M.B.C. and C.D.E. wrote the paper with input from all authors.

Funding

National Science Foundation [214363 to J.D.Y.]; National Science Foundation [2047328]; Nebraska Center for Integrated Biomolecular Communication [P20 GM113126]; UNL startup funds (to C.D.E.). Funding for open access charge: National Science Foundation.

Conflict of interest statement

None declared.

References

1. Ken,M.L., Roy,R., Geng,A., Ganser,L.R., Manghrani,A., Cullen,B.R., Schulze-Gahmen,U., Herschlag,D. and Al-Hashimi,H.M. (2023) RNA conformational propensities determine cellular activity. *Nature*, **617**, 835–841.
2. Liu,B., Shi,H. and Al-Hashimi,H.M. (2021) Developments in solution-state NMR yield broader and deeper views of the dynamic ensembles of nucleic acids. *Curr. Opin. Struct. Biol.*, **70**, 16–25.
3. Ganser,L.R., Kelly,M.L., Herschlag,D. and Al-Hashimi,H.M. (2019) The roles of structural dynamics in the cellular functions of RNAs. *Nat. Rev. Mol. Cell Biol.*, **20**, 474–489.

4. Spitale,R.C. and Incarnato,D. (2023) Probing the dynamic RNA structurome and its functions. *Nat. Rev. Genet.*, **24**, 178–196.
5. Marusic,M., Toplishek,M. and Plavec,J. (2023) NMR of RNA - structure and interactions. *Curr. Opin. Struct. Biol.*, **79**, 102532.
6. Marusic,M., Schlagnitweit,J. and Petzold,K. (2019) RNA dynamics by NMR spectroscopy. *ChemBioChem*, **20**, 2685–2710.
7. Rinnenthal,J., Buck,J., Ferner,J., Wacker,A., Furtig,B. and Schwalbe,H. (2011) Mapping the landscape of RNA dynamics with NMR spectroscopy. *Acc. Chem. Res.*, **44**, 1292–1301.
8. Morandi,E., Manfredonia,I., Simon,L.M., Anselmi,F., van Hemert,M.J., Oliviero,S. and Incarnato,D. (2021) Genome-scale deconvolution of RNA structure ensembles. *Nat. Methods*, **18**, 249–252.
9. Zubradt,M., Gupta,P., Persad,S., Lambowitz,A.M., Weissman,J.S. and Rouskin,S. (2017) DMS-MaPseq for genome-wide or targeted RNA structure probing in vivo. *Nat. Methods*, **14**, 75–82.
10. Tomezsko,P.J., Corbin,V.D.A., Gupta,P., Swaminathan,H., Glasgow,M., Persad,S., Edwards,M.D., McIntosh,L., Papenfuss,A.T., Emery,A., *et al.* (2020) Determination of RNA structural diversity and its role in HIV-1 RNA splicing. *Nature*, **582**, 438–442.
11. Olson,S.W., Turner,A.W., Arney,J.W., Saleem,I., Weidmann,C.A., Margolis,D.M., Weeks,K.M. and Mustoe,A.M. (2022) Discovery of a large-scale, cell-state-responsive allosteric switch in the 7SK RNA using DANCE-MaP. *Mol. Cell*, **82**, 1708–1723.
12. Smola,M.J., Rice,G.M., Busan,S., Siegfried,N.A. and Weeks,K.M. (2015) Selective 2'-hydroxyl acylation analyzed by primer extension and mutational profiling (SHAPE-MaP) for direct, versatile and accurate RNA structure analysis. *Nat. Protoc.*, **10**, 1643–1669.
13. Camara,M.B., Sobeh,A.M. and Eichhorn,C.D. (2023) Progress in 7SK ribonucleoprotein structural biology. *Front Mol. Biosci.*, **10**, 1154622.
14. AJ,C.Q., Bugai,A. and Barboric,M. (2016) Cracking the control of RNA polymerase II elongation by 7SK snRNP and P-TEFb. *Nucleic Acids Res.*, **44**, 7527–7539.
15. Peterlin,B.M., Brogie,J.E. and Price,D.H. (2012) 7SK snRNA: a noncoding RNA that plays a major role in regulating eukaryotic transcription. *Wiley Interdiscip. Rev. RNA*, **3**, 92–103.
16. Yang,Z., Zhu,Q., Luo,K. and Zhou,Q. (2001) The 7SK small nuclear RNA inhibits the CDK9/cyclin T1 kinase to control transcription. *Nature*, **414**, 317–322.
17. Nguyen,V.T., Kiss,T., Michels,A.A. and Bensaude,O. (2001) 7SK small nuclear RNA binds to and inhibits the activity of CDK9/cyclin T complexes. *Nature*, **414**, 322–325.
18. Brogie,J.E. and Price,D.H. (2017) Reconstitution of a functional 7SK snRNP. *Nucleic Acids Res.*, **45**, 6864–6880.
19. Flynn,R.A., Do,B.T., Rubin,A.J., Calo,E., Lee,B., Kuchelmeister,H., Rale,M., Chu,C., Kool,E.T., Wysocka,J., *et al.* (2016) 7SK-BAF axis controls pervasive transcription at enhancers. *Nat. Struct. Mol. Biol.*, **23**, 231–238.
20. Luo,L., Chiu,L.Y., Sugarman,A., Gupta,P., Rouskin,S. and Tolbert,B.S. (2021) HnRNP A1/A2 proteins assemble onto 7SK snRNA via context dependent interactions. *J. Mol. Biol.*, **433**, 166885.
21. Wassarman,D.A. and Steitz,J.A. (1991) Structural analyses of the 7SK ribonucleoprotein (RNP), the most abundant human small RNP of unknown function. *Mol. Cell. Biol.*, **11**, 3432–3445.
22. Marz,M., Donath,A., Verstraete,N., Nguyen,V.T., Stadler,P.F. and Bensaude,O. (2009) Evolution of 7SK RNA and its protein partners in metazoa. *Mol. Biol. Evol.*, **26**, 2821–2830.
23. Reddy,R., Henning,D., Subrahmanyam,C.S. and Busch,H. (1984) Primary and secondary structure of 7-3 (K) RNA of Novikoff hepatoma. *J. Biol. Chem.*, **259**, 12265–12270.
24. Ji,X., Zhou,Y., Pandit,S., Huang,J., Li,H., Lin,C.Y., Xiao,R., Burge,C.B. and Fu,X.D. (2013) SR proteins collaborate with 7SK and promoter-associated nascent RNA to release paused polymerase. *Cell*, **153**, 855–868.
25. Bugai,A., Quaresma,A.J.C., Friedel,C.C., Lenasi,T., Duster,R., Sibley,C.R., Fujinaga,K., Kukanja,P., Hennig,T., Blasius,M., *et al.* (2019) P-TEFb activation by RBM7 shapes a pro-survival transcriptional response to genotoxic stress. *Mol. Cell*, **74**, 254–267.
26. Van Herreweghe,E., Egloff,S., Goiffon,I., Jady,B.E., Froment,C., Monsarrat,B. and Kiss,T. (2007) Dynamic remodelling of human 7SK snRNP controls the nuclear level of active P-TEFb. *EMBO J.*, **26**, 3570–3580.
27. Briese,M., Saal-Bauernschubert,L., Ji,C., Moradi,M., Ghanawi,H., Uhl,M., Appenzeller,S., Backofen,R. and Sendtner,M. (2018) hnRNP R and its main interactor, the noncoding RNA 7SK, coregulate the axonal transcriptome of motoneurons. *Proc. Natl. Acad. Sci. U.S.A.*, **115**, E2859–E2868.
28. Barrandon,C., Bonnet,F., Nguyen,V.T., Labas,V. and Bensaude,O. (2007) The transcription-dependent dissociation of P-TEFb-HEXIM1-7SK RNA relies upon formation of hnRNP-7SK RNA complexes. *Mol. Cell. Biol.*, **27**, 6996–7006.
29. Shimizu,Y., Inoue,A., Tomari,Y., Suzuki,T., Yokogawa,T., Nishikawa,K. and Ueda,T. (2001) Cell-free translation reconstituted with purified components. *Nat. Biotechnol.*, **19**, 751–755.
30. Green,M.R. and Sambrook,J. (2019) Isolation of DNA fragments from polyacrylamide gels by the crush and soak method. *Cold Spring Harb. Protoc.*, **2019**, pdb.prot100479.
31. Cheng,C.Y., Kladwang,W., Yesselman,J.D. and Das,R. (2017) RNA structure inference through chemical mapping after accidental or intentional mutations. *Proc. Natl. Acad. Sci. U.S.A.*, **114**, 9876–9881.
32. Tian,S., Yesselman,J.D., Cordero,P. and Das,R. (2015) Primerize: automated primer assembly for transcribing non-coding RNA domains. *Nucleic Acids Res.*, **43**, W522–W526.
33. Macaya,R.F., Schultze,P., Smith,F.W., Roe,J.A. and Feigon,J. (1993) Thrombin-binding DNA aptamer forms a unimolecular quadruplex structure in solution. *Proc. Natl. Acad. Sci. U.S.A.*, **90**, 3745–3749.
34. Delaglio,F., Grzesiek,S., Vuister,G.W., Zhu,G., Pfeifer,J. and Bax,A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*, **6**, 277–293.
35. Lee,W., Tonelli,M. and Markley,J.L. (2015) NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics*, **31**, 1325–1327.
36. Maciejewski,M.W., Schuyler,A.D., Gryk,M.R., Moraru,I.I., Romero,P.R., Ulrich,E.L., Eghbalnia,H.R., Livny,M., Delaglio,F. and Hoch,J.C. (2017) NMRbox: a resource for biomolecular NMR computation. *Biophys. J.*, **112**, 1529–1534.
37. Cavanagh,J. (2007) In: *Protein NMR Spectroscopy : Principles and Practice*. 2nd edn., Academic Press, Amsterdam; Boston.
38. Furtig,B., Reining,A., Sochor,F., Oberhauser,E.M., Heckel,A. and Schwalbe,H. (2014) Characterization of conformational dynamics of bistable RNA by equilibrium and non-equilibrium NMR. *Curr. Protoc. Nucleic Acid Chem.*, **55**, 11.13.1–11.13.16.
39. Abou Assi,H., Rangadurai,A.K., Shi,H., Liu,B., Clay,M.C., Erharter,K., Kreutz,C., Holley,C.L. and Al-Hashimi,H.M. (2020) 2'-O-methylation can increase the abundance and lifetime of alternative RNA conformational states. *Nucleic Acids Res.*, **48**, 12365–12379.
40. Rangadurai,A., Shi,H., Xu,Y., Liu,B., Abou Assi,H., Boom,J.D., Zhou,H., Kimsey,I.J. and Al-Hashimi,H.M. (2022) Measuring thermodynamic preferences to form non-native conformations in nucleic acids using ultraviolet melting. *Proc. Natl. Acad. Sci. U.S.A.*, **119**, e2112496119.
41. Wahlberg,E. and Hard,T. (2006) Conformational stabilization of an engineered binding protein. *J. Am. Chem. Soc.*, **128**, 7651–7660.
42. Tomezsko,P., Swaminathan,H. and Rouskin,S. (2020) Viral RNA structure analysis using DMS-MaPseq. *Methods*, **183**, 68–75.
43. Virtanen,P., Gommers,R., Oliphant,T.E., Haberland,M., Reddy,T., Cournapeau,D., Burovski,E., Peterson,P., Weckesser,W., Bright,J.,

et al. (2020) SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods*, **17**, 261–272.
44. McKinney,W. (2010) Data structures for statistical computing in python. *Proceedings of the 9th Python in Science Conference*. Vol. **445**, pp. 56–61.
45. Hunter,J.D. (2007) Matplotlib: a 2D graphics environment. *Comput. Sci. Eng.*, **9**, 90–95.
46. Harris,C.R., Millman,K.J., van der Walt,S.J., Gommers,R., Virtanen,P., Cournapeau,D., Wieser,E., Taylor,J., Berg,S., Smith,N.J., *et al.* (2020) Array programming with NumPy. *Nature*, **585**, 357–362.
47. Waskom,M., Botvinnik,O., O'Kane,D., Hobson,P., Lukauskas,S., Gemperline,D., Augspurger,T., Halchenko,Y., Cole,J., Warmenhoven,J., *et al.* (2017) mwaskom/seaborn: v0.8.1 (September 2017). Zenodo, https://zenodo.org/records/883859.
48. Cock,P.J., Antao,T., Chang,J.T., Chapman,B.A., Cox,C.J., Dalke,A., Friedberg,I., Hamelryck,T., Kauff,F., Wilczynski,B., *et al.* (2009) Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics*, **25**, 1422–1423.
49. Darty,K., Denise,A. and Ponty,Y. (2009) VARNA: interactive drawing and editing of the RNA secondary structure. *Bioinformatics*, **25**, 1974–1975.
50. Gruber,A.R., Koper-Emde,D., Marz,M., Tafer,H., Bernhart,S., Obernosterer,G., Mosig,A., Hofacker,I.L., Stadler,P.F. and Benecke,B.J. (2008) Invertebrate 7SK snRNAs. *J. Mol. Evol.*, **66**, 107–115.
51. Will,S., Reiche,K., Hofacker,I.L., Stadler,P.F. and Backofen,R. (2007) Inferring noncoding RNA families and classes by means of genome-scale structure-based clustering. *PLoS Comput. Biol.*, **3**, e65.
52. Will,S., Joshi,T., Hofacker,I.L., Stadler,P.F. and Backofen,R. (2012) LocARNA-P: accurate boundary prediction and improved detection of structural RNAs. *RNA*, **18**, 900–914.
53. Weinberg,Z. and Breaker,R.R. (2011) R2R–software to speed the depiction of aesthetic consensus RNA secondary structures. *BMC Bioinf.*, **12**, 3.
54. Lorenz,R., Bernhart,S.H., Honer Zu Siederdissen,C., Tafer,H., Flamm,C., Stadler,P.F. and Hofacker,I.L. (2011) ViennaRNA package 2.0. *Algorithms Mol. Biol.*, **6**, 26.
55. Andronescu,M., Condon,A., Hoos,H.H., Mathews,D.H. and Murphy,K.P. (2007) Efficient parameter estimation for RNA secondary structure prediction. *Bioinformatics*, **23**, i19–i28.
56. Bothe,J.R., Nikolova,E.N., Eichhorn,C.D., Chugh,J., Hansen,A.L. and Al-Hashimi,H.M. (2011) Characterizing RNA dynamics at atomic resolution using solution-state NMR spectroscopy. *Nat. Methods*, **8**, 919–931.
57. Mitchell,D., Cotter,J., Saleem,I. and Mustoe,A.M. (2023) Mutation signature filtering enables high-fidelity RNA structure probing at all four nucleobases with DMS. *Nucleic Acids Res.*, **51**, 8744–8757.
58. Guo,L.T., Adams,R.L., Wan,H., Huston,N.C., Potapova,O., Olson,S., Gallardo,C.M., Graveley,B.R., Torbett,B.E. and Pyle,A.M. (2020) Sequencing and structure probing of long RNAs using MarathonRT: a next-generation reverse transcriptase. *J. Mol. Biol.*, **432**, 3338–3352.
59. Schuster,P., Fontana,W., Stadler,P.F. and Hofacker,I.L. (1994) From sequences to shapes and back: a case study in RNA secondary structures. *Proc. Biol. Sci.*, **255**, 279–284.
60. Yang,Y., Liu,S., Egloff,S., Eichhorn,C.D., Hadjian,T., Zhen,J., Kiss,T., Zhou,Z.H. and Feigon,J. (2022) Structural basis of RNA conformational switching in the transcriptional regulator 7SK RNP. *Mol. Cell*, **82**, 1724–1736.
61. Aviran,S. and Incarnato,D. (2022) Computational approaches for RNA structure ensemble deconvolution from structure probing data. *J. Mol. Biol.*, **434**, 167635.
62. Yazbeck,A.M., Tout,K.R. and Stadler,P.F. (2018) Detailed secondary structure models of invertebrate 7SK RNAs. *RNA Biology*, **15**, 158–164.
63. Solomatin,S.V., Greenfeld,M., Chu,S. and Herschlag,D. (2010) Multiple native states reveal persistent ruggedness of an RNA folding landscape. *Nature*, **463**, 681–684.
64. Winkler,W., Nahvi,A. and Breaker,R.R. (2002) Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature*, **419**, 952–956.
65. Han,G. and Xue,Y. (2022) Rational design of hairpin RNA excited states reveals multi-step transitions. *Nat. Commun.*, **13**, 1523.
66. Dethoff,E.A., Petzold,K., Chugh,J., Casiano-Negroni,A. and Al-Hashimi,H.M. (2012) Visualizing transient low-populated structures of RNA. *Nature*, **491**, 724–728.
67. Xue,Y., Kellogg,D., Kimsey,I.J., Sathyamoorthy,B., Stein,Z.W., McBrairty,M. and Al-Hashimi,H.M. (2015) Characterizing RNA excited states using NMR relaxation dispersion. *Methods Enzymol.*, **558**, 39–73.
68. Furtig,B., Wenter,P., Reymond,L., Richter,C., Pitsch,S. and Schwalbe,H. (2007) Conformational dynamics of bistable RNAs studied by time-resolved NMR spectroscopy. *J. Am. Chem. Soc.*, **129**, 16222–16229.
69. Andreasson,J.O.L., Gotrik,M.R., Wu,M.J., Wayment-Steele,H.K., Kladwang,W., Portela,F., Wellington-Oguri,R., Eterna,P., Das,R. and Greenleaf,W.J. (2022) Crowdsourced RNA design discovers diverse, reversible, efficient, self-contained molecular switches. *Proc. Natl. Acad. Sci. U.S.A.*, **119**, e2112979119.
70. Xu,X. and Chen,S.J. (2012) Kinetic mechanism of conformational switch between bistable RNA hairpins. *J. Am. Chem. Soc.*, **134**, 12499–12507.
71. Hobartner,C. and Micura,R. (2003) Bistable secondary structures of small RNAs and their structural probing by comparative imino proton NMR spectroscopy. *J. Mol. Biol.*, **325**, 421–431.
72. Garst,A.D., Edwards,A.L. and Batey,R.T. (2011) Riboswitches: structures and mechanisms. *Cold Spring Harb. Perspect. Biol.*, **3**, a003533.
73. Cheng,L., White,E.N., Brandt,N.L., Yu,A.M., Chen,A.A. and Lucks,J.B. (2022) Cotranscriptional RNA strand exchange underlies the gene regulation mechanism in a purine-sensing transcriptional riboswitch. *Nucleic Acids Res.*, **50**, 12001–12018.
74. Sobeh,A.M. and Eichhorn,C.D. (2023) C-terminal determinants for RNA binding motif 7 protein stability and RNA recognition. *Biophys. Chem.*, **292**, 106928.

Received: August 23, 2023. **Revised:** November 15, 2023. **Editorial Decision:** November 15, 2023. **Accepted:** November 17, 2023