In vivo structure of the Ty1 retrotransposon RNA genome

Angelika Andrzejewska [1], **Małgorzata Zawadzka** [1], **Julita Gumna** [1], **David J. Garfinkel** [2] **and Katarzyna Pachulska-Wieczorek** [1,*]

[1]Department of Structure and Function of Retrotransposons, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14, 61-704 Poznan, Poland and [2]Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA 30602, USA

ABSTRACT

Long terminal repeat (LTR)-retrotransposons constitute a significant part of eukaryotic genomes and influence their function and evolution. Like other RNA viruses, LTR-retrotransposons efficiently utilize their RNA genome to interact with host cell machinery during replication. Here, we provide the first genome-wide RNA secondary structure model for a LTR-retrotransposon in living cells. Using SHAPE probing, we explore the secondary structure of the yeast Ty1 retrotransposon RNA genome in its native *in vivo* state and under defined *in vitro* conditions. Comparative analyses reveal the strong impact of the cellular environment on folding of Ty1 RNA. *In vivo*, Ty1 genome RNA is significantly less structured and more dynamic but retains specific well-structured regions harboring functional *cis*-acting sequences. Ribosomes participate in the unfolding and remodeling of Ty1 RNA, and inhibition of translation initiation stabilizes Ty1 RNA structure. Together, our findings support the dual role of Ty1 genomic RNA as a template for protein synthesis and reverse transcription. This study also contributes to understanding how a complex multifunctional RNA genome folds *in vivo*, and strengthens the need for studying RNA structure in its natural cellular context.

INTRODUCTION

LTR-retrotransposons and RNA viruses possess a compact RNA genome that encodes information required for replication and encapsidation in nucleotide sequence and secondary and tertiary structures. Genome-wide *in vitro*, *in virio* and *ex virio* RNA structure models were presented for several infectious RNA viruses (1). However, since these RNA structure models may not reflect conditions *in vivo*,

elucidating the native structure of viral RNA genomes exposed to the cellular environment is necessary to reach a comprehensive view for how genome architecture influences viral replication. *In vivo* RNA structure models are available for Zika and Dengue viruses (*Flaviviridae*), and the coronavirus SARS-CoV-2 (2–5). There is also a transcriptome-wide *in vivo* secondary structure model of influenza A virus (IAV) mRNAs (*Orthomyxoviridae*) (6). These studies reveal important differences in the *in vivo* folding of viral RNAs, but it remains unclear what factors account for remodeling of the structure of viral RNAs in cells.

Reverse-transcribing single strand RNA viruses (ssRNA-RT viruses) are widespread in diverse eukaryotes. The most extensively studied are *Pseudoviridae* and *Metaviridae* families, better known as Ty1/Copia and Ty3/Gypsy long terminal repeat (LTR)-retrotransposons, respectively, and the *Retroviridae* family, which contains prominent human pathogens such as HIV-1 (7). These families share homologous proteins and show mechanistic similarities in genome replication, polyprotein processing and virus particle formation (8–11). As critical steps of their replication cycle, these retroelements encapsidate two copies of single-stranded RNA genome (gRNA), synthetize DNA copy of genome utilizing host tRNA as primers and integrate into host DNA (12,13).

Saccharomyces cerevisiae Ty1-5 retrotransposons continue to provide fundamental insights into the mechanism of retrotransposition and the impact of retroelements on eukaryotic cells (8,9,14). Compared to other eukaryotes, the compact yeast genome contains a relatively small fraction of transposon sequences (1.3–3.4%) (15). Except for Ty3, yeast Ty elements belong to Ty1/Copia family, and Ty1 is the most abundant mobile genetic element in many *S. cerevisiae* strains (16). Ty1 resembles simple retroviruses, but lacks an *ENV* (envelope) gene required for infectivity (17). Ty1 gRNA is 5.6-kb long and contains two partially overlapping *GAG* and *POL* ORFs, flanked by a 53-nt 5′UTR and 318-nt 3′UTR (18). Ty1 transcripts are not spliced and

about 15% of total Ty1 gRNA is polyadenylated. Unlike cellular mRNAs, however, Ty1 transcripts lacking a poly(A) tail do not undergo rapid degradation (19,20). Ty1 gRNA's unusual stability (21) contributes to its high abundance, comprising about 0.8% of total cellular RNA (22). Cytoplasmic Ty1 gRNA is utilized as mRNA for translation of Gag and Gag-Pol polyprotein or is directed to cytoplasmic foci, termed retrosomes (19), where full-length Ty1 gRNA and proteins assemble into virus like particles (VLPs) (23). Within the VLP, Ty1 gRNA is reverse transcribed to a linear dsDNA copy that integrates predominantly upstream of genes transcribed by RNA polymerase III (23–27). Numerous functional studies show that analogous to retroviruses, the 5′ and 3′ termini of Ty1 gRNA are rich in regulatory sequences directly involved in encapsidation, dimerization, cyclization and reverse transcription (25,28–32).

Secondary structure models of Ty1 gRNA derived by SHAPE (Selective 2′ Hydroxyl Acylation analyzed by Primer Extension) probing have been proposed for the first 1482 nts (26%) of *ex virio*, *in virio* and *in vitro* Ty1 gRNA (33), and for monomeric and dimeric states of the first 560 nts of the transcript *in vitro* (30–32). However, the *in vivo* structure of Ty1 gRNA or other LTR-retrotransposon transcripts has not been determined. Moreover, as compared to *Escherichia coli*, plant or mammalian cells, much less is known about the folding of mRNA in yeast (34). The first transcriptome-wide measurement of yeast mRNA structure was performed *in vitro* using an enzymatic PARS method (Parallel Analysis of RNA Structure), which suggested that mRNA transcripts with similar biological functions or cellular localization have common structural features (35). PARS analysis also suggests that yeast mRNAs contain more secondary structure in the coding sequences (CDSs) compared to the untranslated regions (UTRs), and unstructured sequences adjacent to the start codon is positively correlated with translation efficiency (TE). In contrast, transcriptome-wide DMS-seq mapping *in vivo* suggests that yeast mRNAs are largely unfolded, and secondary structures are similar in the CDS and the UTRs (36). Also, the correlation between the CDS structure and ribosome occupancy is not detected. Therefore, reconciling RNA structures determined *in vitro* with those determined *in vivo* is required to understand how structure relates to function.

Here, we provide the first genome-wide *in vivo* RNA secondary structure model for the endogenous yeast ssRNA-RT virus Ty1. Using SHAPE probing, we establish the secondary structure of the entire Ty1 RNA genome in its native *in vivo* state as well as under defined *in vitro* conditions. We find that *in vivo* Ty1 gRNA undergoes significant remodeling, is more dynamic and strongly destabilized, in large part by translating ribosomes. Nevertheless, Ty1 gRNA still contains well-structured regions harboring functional *cis*-sequences. The present work increases our understanding of viral RNA folding in living cells, and helps explain the interplay between translation and RNA structure. Furthermore, we show that commonly used SHAPE reagents, NMIA, NAI and 1M7 robustly modify RNA in yeast, as evidenced by strong correlations between position-dependent reactivities *in vivo* for all three reagents.

MATERIALS AND METHODS

Yeast strain, media and growth conditions

Strain DG3412 used in this study is Ty1-less *Saccharomyces paradoxus* strain (DG1768) (*MATα his3-Δ200hisG ura3 gal3 Spo⁻*) containing pBDG202 (pGTy1-H3CLA/URA3/2μ), a multicopy plasmid with Ty1 under the control of the *GAL1* promoter (37,38). Strain DG3412 was grown in SC-Ura 2% raffinose broth at 30°C with constant shaking at 250 rpm. Saturated cultures were diluted to $OD_{600 \, nm}$ of approximately 0.2 with SC-Ura 2% galactose broth to induce Ty1 expression from the *GAL1* promoter. Cultures were grown to a final $OD_{600 \, nm}$ of 1.0 at 22°C with constant shaking at 250 rpm. For antibiotic treatment, 10 mg/ml kasugamycin solution (Enzo Life Sciences) was added to final concentration of 1 mg/ml when the culture $OD_{600 \, nm}$ reached 1.0, and cells were incubated with shaking for 20 min. The reference strain BY4742 (*MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0*) was used for NMIA testing in *S. cerevisiae* and the growth conditions were as above except that SC medium was used instead of SC-Ura.

E. coli strain, media and growth conditions

Escherichia coli DH5α cells (Invitrogen) were grown in LB medium without antibiotics at 37°C overnight with constant shaking at 225 rpm.

In vivo RNA modification

A 100 ml culture of exponentially growing yeast cells or 1 ml overnight culture of *E. coli* cells were centrifuged, the cell pellet was washed with PBS and resuspended in 720 μl of PBS or PBS with 1 mg/ml kasugamycin. Each sample was divided into two equal amounts and treated with 40 μl SHAPE reagent in DMSO (10 mM NMIA, 10 mM 1M7 or 20 mM NAI, final concentration) or 40 μl DMSO alone. For each reagent, different concentrations were tested and the concentration yielding repeatable, high quality results, without over-modification signals, were chosen for final experiments. The modification reactions were carried at 37°C for 15 min (NMIA and NAI) or for 5 min (1M7). For NMIA, longer reaction time (45 min) was also tested and significant differences in the reactivity patterns were not observed. RNA modification reaction with NAI was quenched by adding 400 μl of 1 M DTT. Cells were collected at 8000 × *g* (10°C) for 7 min followed by total RNA isolation.

Isolation of total RNA from cells

Yeast pellets were resuspended in 1 ml of lysis buffer (10 mM Tris–HCl, pH 8.5, 5 mM EDTA, 2% SDS, 2% 2-mercaptoethanol) and incubated at 83°C for 20 min with constant shaking at 450 rpm. The tubes were centrifuged at 12 000 × *g* for 5 min. The supernatants were transferred to a fresh tubes and RNA was extracted twice with phenol (pH 8.0), followed by two extractions with phenol: chloroform (pH 4.5). RNA was recovered by LiCl precipitation overnight at –20°C. RNA pellets were washed twice in 70% ethanol and resuspended in 30 μl of water. Purified RNA

samples were stored at –20°C. RNA from *E. coli* cells was isolated using Direct-zol RNA MiniPrep Kit (Zymo Research) according to the manufacturer's protocol, and additionally purified by ethanol precipitation and resuspended in 10 μl of water.

In vitro transcription

pBDG433, which contains full-length Ty1 DNA under the control of the SP6 promoter, was digested with BamHI restriction enzyme (Thermo Scientific) to generate a linear sense-strand template for transcription *in vitro*. DNA was phenol–chloroform extracted, ethanol precipitated and resuspended in 20 μl of water. Ty1 RNA was synthetized using MEGAscript™ SP6 Transcription Kit (Invitrogen) and recovered by LiCl precipitation overnight at –20°C, and then washed twice in 70% ethanol. Purified RNA was resuspended in 30 μl of water. The integrity of the RNA was determined by high-resolution gel electrophoresis (1.5% agarose gel with formaldehyde). RNA was stored at −20°C.

In vitro NMIA modification

Ty1 *in vitro* transcript (3 pmol) was incubated in 100 μl of refolding buffer (10 mM Tris–HCl, pH 8.0, 100 mM KCl, 0.1 mM EDTA) at 95°C for 3 min, then cooled by placing on ice. RNA was incubated at 37°C for 30 min following the addition 50 μl of 3× folding buffer (120 mM Tris–HCl, pH 8.0, 600 mM KCl, 1.5 mM EDTA, 15 mM MgCl$_2$). The RNA sample was divided into two separate tubes and treated with 8 μl of NMIA in DMSO (1mM NMIA, final concentration) or DMSO alone. The reaction was performed as described above for modification of RNA *in vivo*. RNA was recovered by ethanol precipitation and resuspended in 10 μl of water.

Primer extension and data processing

The optimal amounts of modified (+) and control (−) RNAs were mixed with 1 μl of 10 μM fluorescently labeled primer [Cy5 (+) or Cy5.5 (−)] and 1 μl 2 mM EDTA (for *in vivo* additionally with 3.84 μl of 5M betaine) to a final volume 12 μl. Primer–template solutions were incubated at 95°C for 3 min, 37°C for 10 min and 55°C for 2 min, then 8 μl of reverse transcriptase mix (SuperScript III, Invitrogen) was added to each tube and RNA was reverse transcribed at 50°C for 45 min (*in vitro*) or 90 min (*in vivo*) to provide optimal reaction efficiency. Sequencing ladders were prepared using primers labeled with WellRed D2 (ddC) or LicorIRD-800 (ddA) and a Thermo Sequenase Cycle Sequencing kit (Applied Biosystems) according to the manufacturer's protocol. Samples and sequencing ladders were purified using ZR DNA Sequencing Clean-up Kit (Zymo Research) and analyzed on a GenomeLab GeXP Analysis System (Beckman-Coulter). Electropherograms were processed using ShapeFinder software (39), normalized as described previously (40). Exemplary raw electropherograms are presented on Supplementary Figure S1. Reactivity information was obtained from eighteen overlapping reads of ∼350 nts each and at least two independent experiments were obtained for each read. Primers are listed in Supplementary Table S1.

Signal-to-background ratio and correlation calculation

Signal-to-background (S/B) ratio (Figure 1B) was calculated based on the output files from ShapeFinder. Medians for signal and background were estimated from the averaged values of signal or background peak areas for each nucleotide. S/B is the ratio of the obtained signal and background medians. Spearman's correlations and linear regressions were computed using the GraphPad Prism 8 software.

RNA secondary structure modeling and analysis

SuperFold software (41) was used for minimum free energy (MFE) secondary structure prediction, calculation of base pairing probabilities for all possible canonical base pairs and identification of regions with well-defined structures. SuperFold uses Partition and Fold functions implemented in RNAstructure (42) and the SHAPE reactivities were used as a pseudo-energy constraints. For all calculations slope and intercept folding parameters were set to 1.8 and –0.6 kcal mol^{-1}, respectively.

Prediction of RNA pseudoknots. ShapeKnots (implemented in RNAstructure) was used to identify pseudoknots in the *in vivo* and *in vitro* Ty1 RNAs. Folding was performed in 600-nt sliding windows with 100-nt increments. Four additional folds were computed at the 5′ and 3′ ends to increase sampling of terminal regions. Pseudoknots occurring in >50% of the windows were included in the MFE structure.

Partition function calculation. Partition was run in 1200-nt sliding windows, incremented by 100 nts, with four additional calculations performed on the 5′ and 3′ ends of the Ty1 RNA. Next, 300 nts were trimmed from the 5′ and 3′ ends of each window to eliminate end effects. The base-pairing distance was limited to 600 nts. Base-pair probabilities from the multiple partition function windows were combined and pseudoknot nucleotides were involved. Shannon entropies were calculated from individual base-pairing probabilities and combined into a single profile in 55-nt sliding windows.

Minimum free energy structure prediction. The MFE structure models were predicted using the Fold function with pseudo-energy constraints, and base pairs with >99% probability were used as hard pairing constraints. Fold was performed in 3000-nt sliding windows, incremented by 300 nts to predict individual MFE structures over each window, with four additional fold calculations on the 5′ and 3′ ends of the genome. Next, 300 nts were trimmed from the 5′ and 3′ ends of each window to eliminate end effects. These multiple folds were then combined into a final MFE structure by requiring that base pairs appeared in a majority of the windowed folds. The pseudoknotted helices were added to the final RNA structure. The maximum pairing distance was set to 600 nts. Additional prediction of Ty1 MFE structure was performed with tRNA$_i$^{Met} binding, dimerization and cyclization sequences forced to be single-stranded. The MFE structure model of the *in virio* Ty1 RNA + 1–1482 region was generated based on previously published SHAPE data (33), using the RNAstructure.
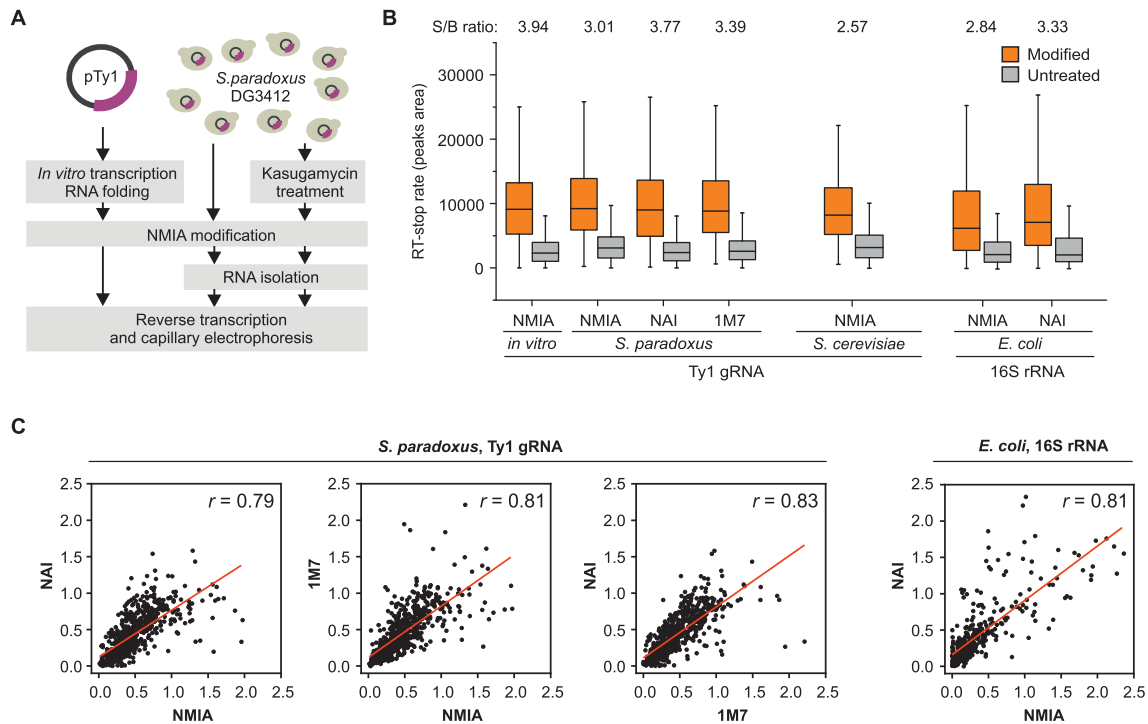
Figure 1. Experimental workflow and SHAPE reagent comparison. (**A**) Experimental workflow. (**B**) Box plot analysis of RT-stop rate (CE peaks area) measurement with medians for signal and background from capillary electrophoresis experiments for different SHAPE reagents and different cells. Plots present data for approximately 600 nts of Ty1 gRNA *in vitro* and in *S. paradoxus* and approximately 400 nts of Ty1 gRNA in *S. cerevisiae* and 400 nts of 16S r RNA in *E. coli*. (**C**) Correlation of position-dependent NMIA, NAI and 1M7 reactivities from *in vivo* probing of Ty1 gRNA in *S. paradoxus* and 16S rRNA in *E.coli*.

Identification of regions with well-defined structures. Well-defined regions (lowSS) were identified by selecting regions of at least 40 nts with both median SHAPE reactivities and median Shannon entropies below the global medians, as described previously (43). Some lowSS regions were expanded or combined to preserve the entirety of helices predicted in the MFE structure. The pseudoknot helix was ignored in lowSS regions identification, because its presence caused the lowSS region 1 extension to high entropy and low base-pair probability regions. RNA structures were visualized with VARNA (http://varna-gui.software.informer.com/) (44).

Sensitivity and positive prediction value (PPV) calculations. The sensitivity and PPV for the obtained MFE models were calculated using Scorer function (implemented in RNAstructure). If the region of interest contained only half the predicted helix, calculations were performed with the manual addition of nucleotides from disrupted helix. For the entire Ty1 RNA and identified lowSS regions, the *in vitro* structure was used as the reference state while in other cases the reference structure was indicated in the figures. Sensitivity is defined as the percentage of base pairs in the reference MFE structure that are also present in the compared MFE structure, and the PPV represents the percentage of base pairs in the MFE structure that are also present in the reference model (45).

RESULTS

High-resolution probing of Ty1 genome under *in vivo* and *in vitro* conditions

We used SHAPE followed by capillary electrophoresis to explore the structure of gRNA of an active Ty1 retrotransposon at single-nucleotide resolution. We probed Ty1 gRNA structure under three experimental conditions: in *S. paradoxus* cells expressing an active Ty1 element, *in vitro* in the absence of cellular components, and in cells partially inhibited for translation following treatment with kasugamycin (Figure 1A). The parental yeast strain DG1768 lacks full-length chromosomal Ty1 sequences [(37) and Garfinkel and Bergman, unpublished results], and expression of Ty1 occurs exclusively from a multicopy pGTy1 plasmid driven by the inducible *GAL1* promoter (23,46). Galactose induction produces a high level of correctly initiated Ty1 RNA that participates in all steps of the replication cycle, including translation, co-localization into retrosomes, encapsidation into VLPs, and reverse transcription (38). The resulting cDNA copy integrates into host DNA, and a high frequency of Ty1 retromobility can be measured. Based on the available data, we assume that the major fraction of Ty1 RNA is present in the cytoplasm or associated with retrosomes, and a much smaller amount is packaged in VLPs (38). As this approach allowed us to obtain SHAPE reactivity data from a homogenous Ty1 gRNA population, a more reliable RNA secondary structure model was gener-

ated. In contrast, the widely used *S. cerevisiae* S288c laboratory strain background contains 32 full length Ty1 elements comprised of three subfamilies: Ty1, Ty1/Ty2 hybrids, and the ancestral Ty1′ element (15,47,48). Ty1 elements are expressed at different levels that vary up to 50-fold, with 75% of total Ty1 expression coming from eleven highly expressed elements, and eight of these are Ty1/Ty2 hybrids (49). Additionally, some full length Ty1 copies contain deletions and mutations that may affect transposition, but they can be still expressed.

SHAPE exploits 2′-hydroxyl-selective reagents to interrogate local backbone flexibility in RNA by forming covalent 2′-O-adducts at flexible nucleotides (50). Generally, SHAPE reactivities are inversely proportional to local nucleotide conformations, such that flexible or unpaired nucleotides are reactive, while structurally constrained or base-paired nucleotides tend to be less reactive (51). Diverse cell types have different permeabilities to SHAPE reagents, and in general, those growing in a suspension culture are more permeable than adherent cell lines (52). Surprisingly little is known about yeast permeability to SHAPE reagents and only NAI (2-methylnicotinic acid imidazolide) has been used in SHAPE analyses in yeast (53–55). To the best of our knowledge, other well-validated SHAPE reagents, such as NMIA (N-methylisatoic anhydride) or closely related 1M7 (1-methyl-7-nitroisatoic anhydride) have not been used for RNA structure mapping in yeast, however, their ability to modify RNA occurs in other cell types (52,56–59). In contrast, one group reports that NMIA reactivity is not detected in an adherent mouse cell line (mESCs) (53), and a very low 1M7 signal is detected in mESCs cells and *E. coli* (60).

SHAPE reagents have local-structure and nucleobase biases that can impact reactivity profiles (52,61,62). All prior studies of Ty1 RNA structure have been performed using NMIA as the probing reagent (30,32,33,63). We reasoned that the application of the same SHAPE reagent for Ty1 RNA structure mapping *in vivo* would allow for direct and unbiased comparison of reactivity patterns with those presented for Ty1 RNA 5′ terminal region *in virio* and *ex virio* (33). This approach also facilitates detection of NMIA reactivity changes *in vivo* that result from Gag-induced RNA dimerization (30), pseudoknot formation (32,33), and primer tRNA annealing (30,33). Therefore, we initially tested the ability of NMIA to modify the 5′ end of Ty1 gRNA in yeast, and compared this data with that obtained from NAI and 1M7 probing. We clearly detected high reactivity signals well above background for all three reagents using capillary electrophoresis for analysis of adduct-induced cDNA truncations (Figure 1B). Since the signal-to-background ratio is the raw measurement of structure-dependent RNA modification, these results show that NMIA, NAI and 1M7 robustly modify RNA in *S. paradoxus*. Additionally, *E. coli* and *S. cerevisiae* treated with NMIA also displayed high reactivity above the background (Figure 1B). Our data indicate that NAI is slightly more efficient than NMIA in both yeast and *E. coli*, but a recent study suggests that NAI has stronger nucleotide biases and less effectively differentiates paired and unpaired C and G residues (52). Despite that, we observed strong correlations between position-dependent NMIA, 1M7 and NAI

reactivities *in vivo*, but some local single-nucleotide reactivity differences were also detected (Figure 1C). Therefore, the reactivity patterns obtained with the same SHAPE reagents are more informative for detailed analysis of RNA structure alterations specific for diverse biological or experimental states. Using NMIA, we obtained high-quality SHAPE data for 98% of the nucleotides in the Ty1 gRNA *in vivo* and under defined conditions *in vitro*. SHAPE data derived from *in vivo* and *in vitro* experiments were used for structural modeling of Ty1 gRNA using SuperFold pipeline, which incorporates a windowing approach along with the *Partition* and *Fold* functions implemented in RNAstructure (41,42).

Impact of cell environment on SHAPE reactivities, base-pairing probabilities and Shannon entropies

To determine the impact of cell environment on Ty1 gRNA folding, we compared the *in vivo* and *in vitro* SHAPE reactivities and base-pairing probabilities combined with the Shannon entropy calculation. The difference in overall median SHAPE reactivities between *in vivo* and *in vitro* Ty1 gRNAs was smaller than expected (0.07) for such distinct environmental conditions (Figure 2A). However, the correlation between datasets from these two states was moderate ($r = 0.49$), which indicates a significant change in the pattern of SHAPE modifications *in vivo* (Figure 2B). In contrast to experiments under defined *in vitro* conditions, SHAPE reactivities *in vivo* are dependent not only on local nucleotide flexibility, but also on interactions with proteins or other RNA molecules (59,64). In general, increased *in vivo* SHAPE reactivities, relative to the *in vitro* state, usually reflect RNA conformational changes induced by the cellular environment, whereas decreased reactivity *in vivo* tends to reflect the binding of proteins and other biomolecules (65). The comparison of median SHAPE reactivity profiles across the Ty1 RNA genome showed that many regions were significantly more reactive *in vivo* than *in vitro,* some were less modified *in vivo*, and several had quite similar reactivity in both states (Figure 2C, D). Overall, the differences in SHAPE reactivity between both states (ΔSHAPE, Figure 2D) suggest there is a strong destabilization of Ty1 gRNA in the cellular environment, and the few specific regions protected from NMIA modification *in vivo* may serve as protein binding sites. However, all protected regions are probably not detected due to strong destabilization of RNA structure *in vivo* that reduces the impact of protein binding on the total SHAPE reactivity measurement.

The difference in SHAPE reactivity between *in vivo* and *in vitro* measurements was the most evident for adenosine residues and the smallest difference was calculated for cytidines (Figure 2E). Since Ty1 gRNA is particularly rich in A-residues (35.5%) while the C-content is low (21.9%), specific Ty1 gRNA nucleotide content likely contributes to differences between *in vivo* and *in vitro* SHAPE reactivity profiles.

Ty1 gRNA contained almost 50% less unreactive nucleotides *in vivo*, but was richer in nucleotides with intermediate SHAPE reactivity (0.4 > 0.85) (Figure 2F). The SHAPE reactivity at any given nucleotide position does not provide information about a single RNA structure but represents the average reactivity over the ensemble of
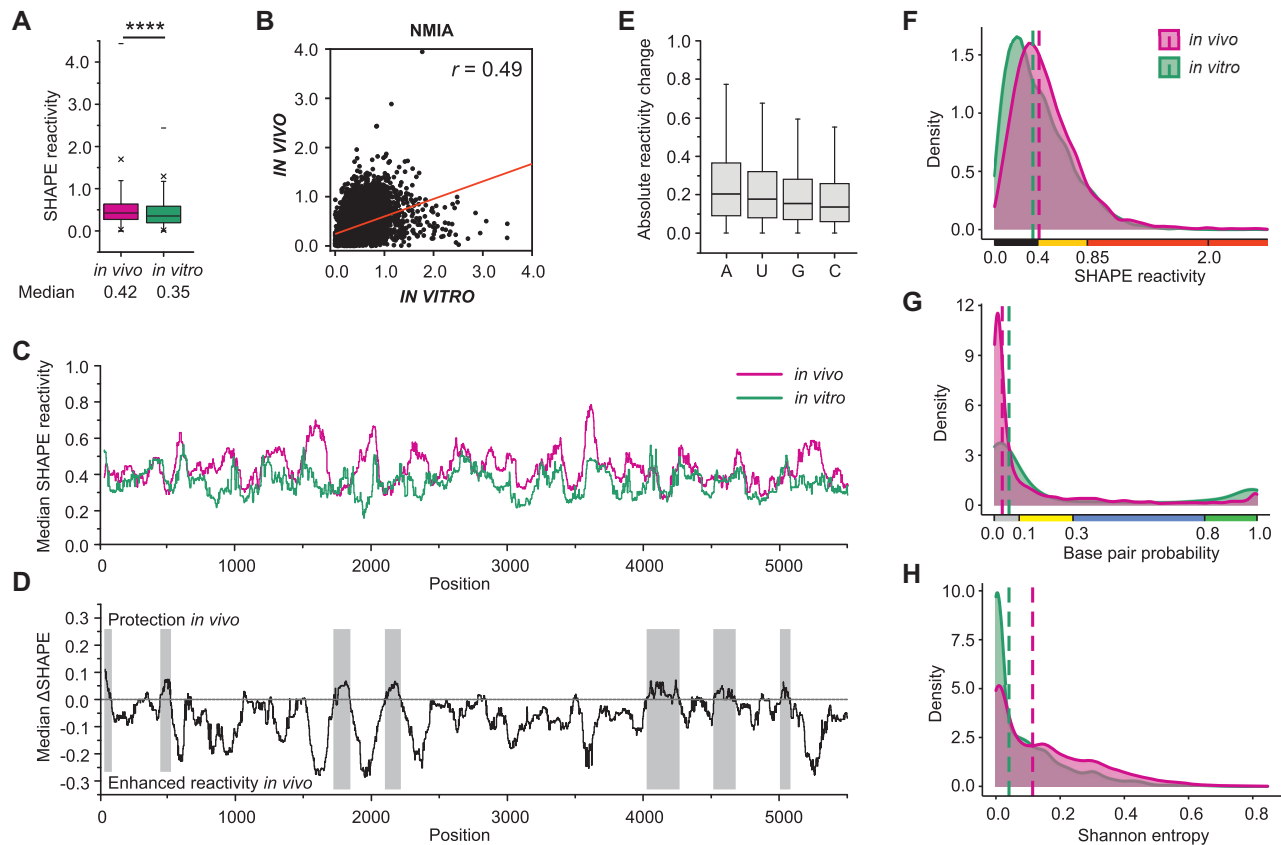
Figure 2. SHAPE reactivities, pairing probabilities and Shannon entropies for *in vivo* and *in vitro* Ty1 gRNAs. (**A**) Box plot analysis of SHAPE reactivity distributions with medians. Significance was computed by unpaired two-tailed Mann–Whitney test; ****$P < 0.0001$. (**B**) Correlation of position-dependent NMIA reactivities for *in vivo* and *in vitro* Ty1 gRNA states. (**C**) The median SHAPE reactivity profiles smoothed with a 75-nt window. (**D**) The median ΔSHAPE profile calculated by subtracting the *in vivo* reactivities from those of the *in vitro* RNA, smoothed with a 75-nt window. Gray shadings indicate regions that may serve as protein binding sites. (**E**) Distribution of the absolute difference in reactivity between *in vitro* and *in vivo* SHAPE measurements, calculated for each nucleotide type. (**F**) NMIA reactivity distributions with medians (dashed lines). (**G**) Base pair probability distributions with medians. (**H**) Shannon entropy distributions with medians.

all coexisting RNA structures (66,67). Nucleotides with low SHAPE reactivity (<0.4) are conformationally constrained and likely to be double-stranded in most conformers, whereas those with intermediate reactivity may be involved in diverse coexisting RNA conformations (68). To consider the entire structural ensemble, we calculated a probability for each base pair across all SHAPE-directed structures predicted for the Ty1 RNA *in vitro* and *in vivo* (45). A significant increase in the number of low-probability base pairs and a concurrent decrease of high-probability pairings was detected *in vivo* (Figures 2G and 3C, F). Consistent with these findings, Shannon (base-pairing) entropy distributions indicated that 2-fold fewer nucleotides form well-defined structures (with low entropy) *in vivo* and the median entropy was significantly higher than *in vitro* (0.11 and 0.04, respectively) (Figures 2H and 3B, E). Taken together, these results strongly suggest that Ty1 gRNA structure is destabilized in yeast cells and is more heterogeneous than under *in vitro* conditions.

We queried the RASP database (69) to determine whether structural information for Ty1 transcripts can be retrieved from published transcriptome-wide studies in yeast. The only available data are from DMS-MaPseq experiments *in vitro* and *in vivo* (70). However, in all Ty1 transcripts, only

about 30% of nucleotides are mapped and all transcripts contain many long 100–200 nts sequences without structural data. In addition, we observe a low correlation between SHAPE and DMS reactivity patterns for fragments with higher DMS data coverage. This may result from different probing reagents and methods of adduct detection.

Identification of Ty1 RNA regions with well-determined secondary structure

Long RNA molecules often adopt a mosaic structure, such that conformationally dynamic regions are interspaced with regions with well-determined secondary structure (43,66,71,72). Importantly, RNA regions with persistent stable structural motifs tend to be functionally important and contain *cis*-acting regulatory sequences (73–75). Based on low SHAPE and low Shannon entropy metrics (lowSS regions), we identified 11 lowSS regions within *in vivo* and *in vitro* Ty1 RNAs (Figure 3B,E; numbered dark violet strips). They were separated by conformationally dynamic regions that exhibited higher Shannon entropy even if SHAPE values remained below the global median reactivity. The 5′ and 3′ terminal lowSS regions (regions 1 and 11) detected *in vivo* correspond well with sequences required for
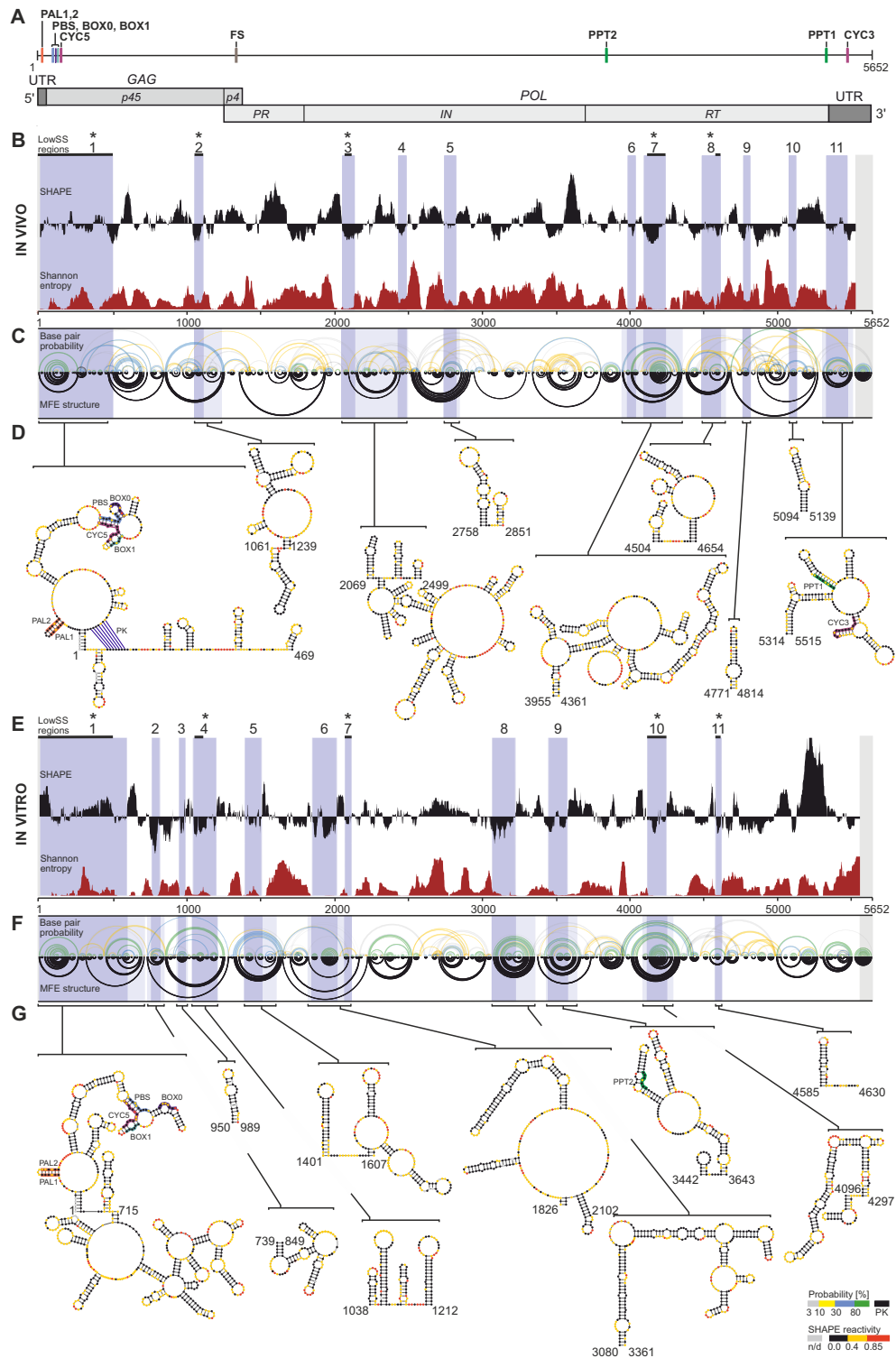
Figure 3. SHAPE-directed structure models and structurally well-defined regions for *in vivo* and *in vitro* Ty1 gRNAs. (**A**) Functional organization of Ty1 gRNA. The known *cis*-acting sequences are indicated: dimerization sequences (PAL1, PAL2), primer-binding sites (PBS, BOX0, BOX1), cyclization sites (CYC5, CYC3), frameshift site (FS) and polypurine tracts (PPT1, PPT2). (**B**) *In vivo* median SHAPE reactivity and Shannon entropy profiles smoothed with a 55-nt window. LowSS regions are marked by dark violet shadings with light violet shadings extended to encompass entire intersecting helices from MFE structures. Grey shadings indicate regions without experimental data. Stars mark regions that overlap in both Ty1 gRNAs and black horizontal lines indicate overlapping fragments. (**C**) Arc plots showing the base-pairing probabilities and the predicted MFE structure for *in vivo* Ty1 gRNA. At the top base pairs are colored with respect to their probabilities (see scale), black arcs indicate a predicted pseudoknot (PK). (**D**) Secondary structure models for *in vivo* lowSS regions. Nucleotides are colored by SHAPE reactivity (see scale). Panels (**E**), (**F**) and (**G**) show the analogous analysis performed for *in vitro* Ty1 gRNA. The MFE structure models were predicted with the maximum pairing distance of 600 nts, and they do not include CYC5/CYC3 long-range interaction. High-resolution structures of lowSS regions and entire Ty1 gRNAs are provided in Supplementary Figures S2 and S3.

for transposition of a minimal Ty1 element when Gag and Gag-Pol are provided *in trans* (28,76). The 5′ and 3′ ends of Ty1 RNA contain *cis*-acting sequences directly involved in retromobility, such as palindromes (PAL) required for Ty1 RNA dimerization, PBS, BOX0 and BOX1 that anneal with primer $tRNA_i^{Met}$, CYC5 and CYC3 required for genome cyclization, and the polypurine tract, which is required for plus-strand synthesis (23,25,28–30,32) (Figure 3A). These important RNA sequences were embedded in low SS regions identified *in vivo*. Thus, our data highlight the correlation between well-determined RNA conformation and biological function. An interesting but expected exception is the sequence mediating Ty1 translational frameshifting. Although the frameshift heptamer is not nested within a well-structured region, Ty1 frameshifting depends on the low abundance of a specific isoacceptor tRNA and does not require a specific RNA structure (77).

Only 5 lowSS regions detected *in vivo* overlapped with those identified *in vitro* (Figure 3B,E; strips marked with asterisks), while others were unique to the *in vivo* or *in vitro* state. Interestingly, the 5′ terminus containing the majority of known *cis*-acting sequences was well-structured both *in vivo* and *in vitro*. However, the 3′ terminal region with CYC3 and PPT1 was well-structured only in the *in vivo* state. Thus, our data suggest that detection of regions meeting low SHAPE and low entropy criteria *in vivo* might enable more precise identification of the positions of unknown regulatory RNA sequences. Here, we identified eight lowSS regions in the *POL* ORF of *in vivo* Ty1 RNA. Although their functions remain to be determined, four of them (regions 3, 6, 7 and 8) overlapped well with the regions protected from NMIA modification *in vivo*, perhaps resulting from protein binding (Figure 2D).

A prior study of the entire HIV-1 RNA genome structure *in vitro* shows decreases in SHAPE at RNA regions that encode protein domain junctions (78). These more structured RNA regions are proposed to induce ribosome pausing that facilitates proper protein folding. However, we do not detect decreases in SHAPE reactivity in regions of Ty1 gRNA that encode protein domain boundaries *in vivo* or *in vitro*. Our data indicate that such regions in Ty1 gRNA exhibit equal or higher reactivity values than the median reactivity in each state. A decrease in SHAPE reactivity at boundary RNA regions is also not detected for Zika and Dengue gRNA *in vivo*, suggesting different strategies of ribosome elongation regulation (3). Indeed, RNA structure is one of the mechanisms promoting ribosome pausing, but there are also other contributors to this process, such as mRNA sequence, charged tRNA and translation factor availability, and the nascent protein chain itself (79).

Comparison of minimum free energy structure models of the *in vivo* and *in vitro* Ty RNA genome

The consensus minimum free energy (MFE) secondary structure of Ty1 gRNA analyzed *in vivo* and *in vitro* was predicted using the SuperFold pipeline, which is dedicated to modeling structures of large RNAs (Figure 3C, F and Supplementary Figure S3) (41). SuperFold takes a windowing approach, combines base pairs predicted for overlapping windows, incorporates base pairs with the highest probabili-

ties (>99%), and the final structure is predicted by requiring base pairs that occur in more than one-half of the windows. Comparing the resulting MFE structures in terms of sensitivity and positive predictive value (PPV) showed a significant difference in the folding of Ty1 RNAs. Both parameters were in the range of 49–55%, which indicated that only about half of the base pairs were shared between *in vitro* and *in vivo* MFE structures, while the rest of the parings remained unique for each model (Figure 4A, B). Despite these differences, roughly 50% of the nucleotides were predicted to be base-paired in both MFE structures, suggesting similar levels of Ty1 gRNA structure *in vitro* and *in vivo* (Figure 4E). In contrast, analysis of SHAPE reactivities, base-pairing probabilities and Shannon entropies strongly suggest that Ty1 gRNA structure is destabilized in yeast cells. A single MFE model, even predicted as a consensus structure, represents one conformer from an ensemble of possible RNA conformers with similar free energies (ΔG) (42,45). Thus, focusing on a single MFE structure can lead to an incomplete or biased understanding of RNA folding, especially for structurally dynamic long RNAs.

We reasoned that alterations occurring in well-structured regions (lowSS) may better reflect changes in Ty1 gRNA folding under different experimental conditions. Therefore, we compared the MFE structures of Ty1 RNAs across all identified *in vivo* and *in vitro* lowSS regions by calculating a sensitivity value and PPV for these selected RNA segments. Most of the base pairs predicted in the *in vivo* lowSS regions were retained in the *in vitro* MFE structure, even if they were not in lowSS regions of the *in vitro* RNA (median sensitivity and PPV – 92% and 76%) (Figure 4C). This trend was not observed when we compared base pairs from *in vitro* lowSS regions with the *in vivo* MFE structure (median sensitivity and PPV – 30.5% and 38.5%) (Figure 4D). However, higher structural similarity was found for lowSS regions that partially overlapped in *in vitro* and *in vivo* Ty1 gRNA. Lower sensitivity and PPV metrics for overlapping region 1 can be caused by *cis*-acting sequences that mediate inter- and intramolecular RNA interactions that may occur *in vivo* but not *in vitro* (see below). Together, these analyses suggest that a large fraction of well-defined structural motifs detected *in vivo* are preserved *in vitro*, but not vice versa.

Analysis of base-pairing probabilities supports destabilization of Ty1 RNA *in vivo*

To better understand Ty1 gRNA folding, we repeated the comparative analysis of the *in vivo* and *in vitro* Ty1 gRNA structures considering only highly-probable base pairs (HP bps, pairing probability > 80%). HP bps are predicted with the highest confidence and are more likely to be present in the gRNA structure (45). Although a similar number of MFE base pairs was predicted for both Ty1 gRNA states, the contribution of high-probability base pairs in the *in vivo* MFE structure was much lower. Accordingly, 17% of Ty1 gRNA nucleotides participated in HP bps *in vivo*, representing 37.5% of the base pairs in the MFE structure, whereas 34.1% of nucleotides were involved in HP bps *in vitro*, and they represented 68% of the MFE base pairs (Figure 4E, F). We found that only 39% of *in vitro* HP bps were also predicted to occur in the *in vivo* MFE structure, while

Figure 4. Comparison of SHAPE-directed MFE structures predicted for *in vivo* and *in vitro* Ty1 gRNAs. (**A**) Calculation of sensitivity (sens) and PPV parameters for the MFE structures of *in vivo* and *in vitro* Ty1 RNAs. *In vitro* Ty1 gRNA MFE structure was used as a reference structure. Values are colored from low (red) to high (green). (**B**) Combined model of MFE structures of *in vivo* and *in vitro* Ty1 gRNAs. Common base pairs are indicated by black arcs, base pairs present only *in vivo* by pink arcs and those unique for *in vitro* gRNA by dark green arcs. Calculation of sensitivity and PPV parameters for lowSS regions identified *in vivo* (**C**) and *in vitro* (**D**). (**E**) Percentages of paired nucleotides in the predicted MFE structures, including high probability base pairs (HP bps, pairing probability > 80%). (**F**) Venn diagram showing overlap of highly probable base pairs between *in vivo* and *in vitro* MFE structures. (**G**) Comparison of high probability base pairs in the *in vivo* and *in vitro* MFE structure models of Ty1 gRNA.

almost 80% of *in vivo* HP bps were shared with *in vitro* MFE structure (Figure 4F). This disparity arises from the lower number of HP bps predicted for the *in vivo* model of Ty1 gRNA. We observed that among HP bps the short-range interactions representing locally stable secondary RNA structures were preferentially shared between both Ty1 gRNA states (Figure 4G). In contrast, a majority of highly probable longer-range base pairings *in vitro* were disrupted or became less probable *in vivo*. These analyses confirm that Ty1 gRNA possesses a significantly lower and more dynamic structure *in vivo*, and show that only part of stable (highly probable) base pairs *in vitro* are retained *in vivo*, consistent with the findings from comparison of the MFE structures across lowSS regions.

Ty1 RNA dimerization, cyclization and tRNA$_i^{Met}$ annealing *in vivo*

We analyzed SHAPE reactivity changes in the 5′-end of Ty1 gRNA in more detail to establish whether critical RNA - RNA interactions could occur *in vivo* prior to packaging

into VLPs. Ty1 gRNA dimerization, cyclization or primer tRNA annealing do not occur under the *in vitro* conditions employed here because Ty1 proteins such as Gag and tRNA are absent in the analysis. RNA sequences involved in tRNA$_i^{Met}$ binding (PBS, BOX0, and BOX1) and genome dimerization (PAL1 and PAL2) were engaged in intramolecular base-pairing *in vitro*, and we did not detect changes in reactivity patterns that could directly support these processes occurring *in vivo* (Figure 5A). This problem arises from a limitation of the SHAPE method that, like other classical RNA structure mapping techniques, cannot discriminate intra- from intermolecular base-pairing. However, changes in reactivity of nucleotides separating PALs were detected *in vivo*, analogous to those detected in VLPs and in Gag-induced Ty1 RNA dimerization *in vitro*, and supporting PALs involvement in the intermolecular PAL1 - PAL2 interactions (30,33). We also detected SHAPE reactivity increases for nucleotides between the PBS and BOX0 (105 – 107) that occur in Gag- or temperature-induced tRNA$_i^{Met}$/Ty1 RNA *in vitro* complexes and that are also observed in VLPs (30,33). Highly significant reactivity in-
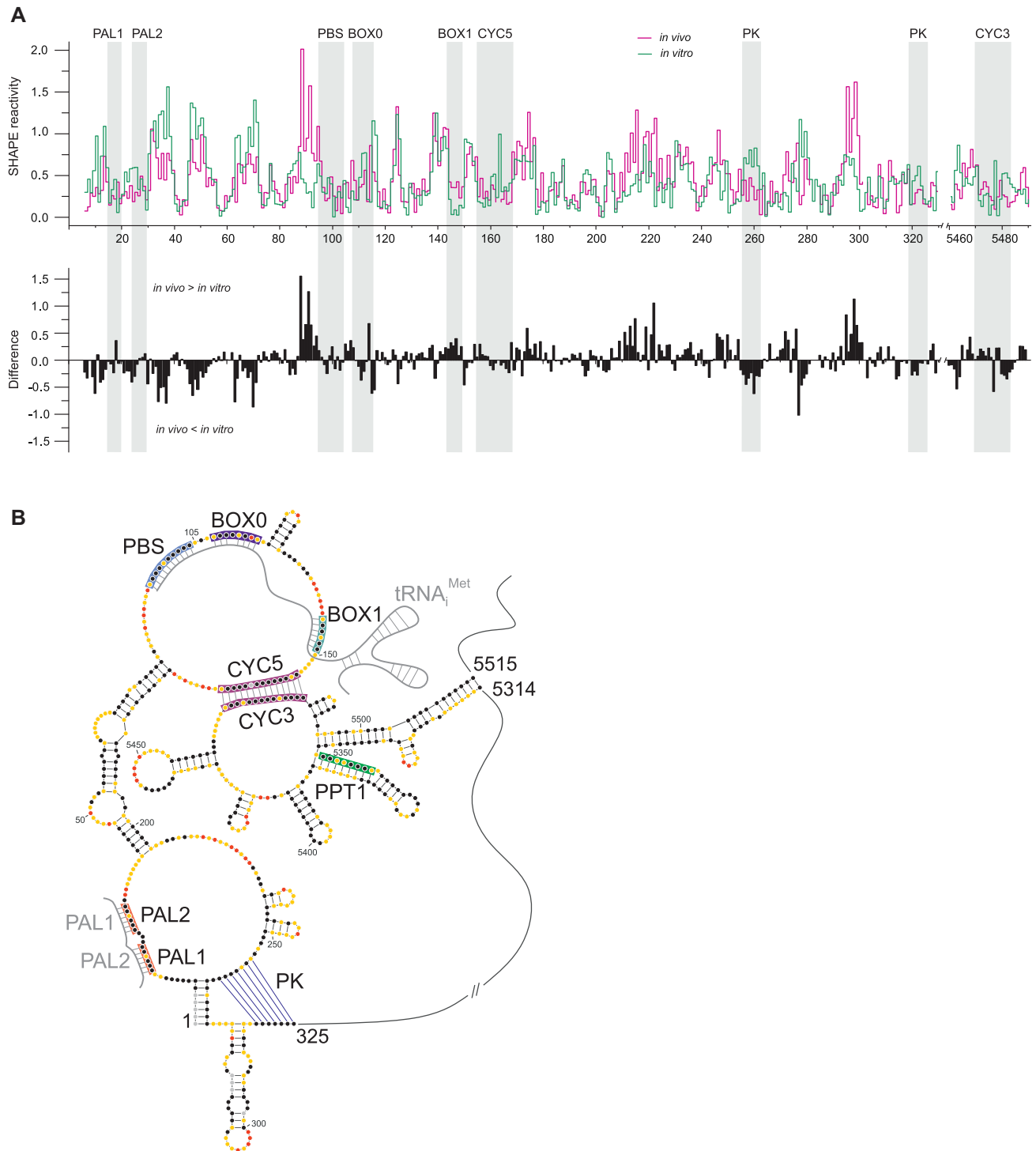
Figure 5. SHAPE-based analysis of Ty1 gRNA dimerization, cyclization and tRNA$_i^{Met}$ annealing. (**A**) The step plot (top) of NMIA reactivity for *in vivo* and *in vitro* gRNAs and the difference plot (bottom) calculated by subtracting the *in vitro* reactivities from those of the *in vivo* gRNA. Negative values indicate nucleotides that are less reactive *in vivo*. Locations of *cis*-acting sequences are indicated by gray stripes. (**B**) Ty1 gRNA in its dimeric and circular form with annealed tRNA$_i^{Met}$. *Cis*-acting sequences are marked with colored boxes, nucleotides are colored by SHAPE reactivity according to the scale presented in Figure 2. For simplicity, only 5′ and 3′ ends are shown. The present RNA structure was predicted with PAL, CYC and tRNA$_i^{Met}$ binding sequences forced to be single-stranded. The intermolecular base-pairs and long-range CYC5/CYC3 interaction were introduced manually based on differences in SHAPE reactivity patterns *in vitro* and *in vivo* and published data (30,32,33).

creases were also observed *in vivo* in the sequence (88–94) preceding the PBS (compare Figure 3D and G, leftmost structures). The majority of this sequence was base-paired and unreactive *in vitro*. *In vivo* SHAPE reactivity increases likely reflect disruption of the helical region preceding the PBS caused by tRNA$_i^{Met}$ hybridization (Figure 5A, B). This interpretation is further supported by reactivity increases in the opposite strand (173–176) of the helix. In addition, destabilization of this helix may be enhanced by CYC5 interaction with CYC3. In the absence of tRNA$_i^{Met}$, CYC5 base-paired with the PBS and BOX1 *in vitro*, thus the reactivity increases in CYC5 were expected as a consequence of tRNA$_i^{Met}$ hybridization *in vivo*. Since CYC5 remains unreactive *in vivo*, this region may be involved in long-range base-pairing with CYC3 at the 3′ end of Ty1 gRNA (Figure 5B), as suggested by genetic studies (29). Although SHAPE can detect RNA structure changes resulting from long-range interactions, the prediction of such interactions is challenging because the accuracy of the predicted RNA structure model decreases with distance between base-pairs (80).

Importantly, the SHAPE data supported the presence of the functionally important pseudoknot at the 5′ terminus of Ty1 gRNA *in vivo* (Figures 3D and 5A, B). The 5′ pseudoknot was identified *in virio*, and mutations disrupting its structure inhibit reverse transcription (32,33). Additionally, our current findings support previous work (33) indicating that the pseudoknot is not formed in the full-length Ty1 transcript *in vitro*.

Together, our data raise the possibility that functionally important RNA intramolecular interactions such as cyclization and pseudoknot formation occur in retrosomes or the cytoplasm prior to packaging of Ty1 gRNA into VLPs. The primer for reverse transcription may be annealed prior to packaging, which also occurs with primer tRNA^{Lys} annealing during HIV-1 propagation (81). Like retroviral RNA genomes, Ty1 gRNA may undergo dimerization prior to being packaged into VLPs. However, further confirmation of these observations is required.

Ty1 RNA structure changes during translation

Ty1 gRNA is both the template for reverse transcription and the mRNA for translation of Gag and the Gag-Pol polyprotein, thus translation may contribute to differences between Ty1 RNA structures *in vivo* and *in vitro*. The correlation between translation and *in vivo* mRNA structure is observed in transcriptome-wide structure probing experiments in *Escherichia coli* and zebrafish (73,82). However, interplay between ongoing translation and mRNA structure *in vivo* is not detected in global yeast transcriptome mapping, and the average *in vivo* structure of coding regions is not distinguishable from that of the UTRs (36). Interestingly, our results suggest Ty1 gRNA may not fit this pattern. The *in vivo* median SHAPE reactivities of *GAG* and *POL* ORFs were higher than that of UTRs, indicating that the Ty1 RNA coding region was less structured in cells than non-coding sequences, while this correlation was not detected *in vitro* (Figure 6A).

To analyze the relationship between translation and RNA structure, we probed Ty1 gRNA structure *in vivo* fol-lowing addition of the aminoglycoside kasugamycin, an antibiotic that specifically inhibits the initial step of protein synthesis by preventing formation of the translation initiation complex (83). This experimental approach was used recently to show the impact of translating ribosomes on mRNA structure in *E. coli* (73). Cells were treated with kasugamycin for 20 min to allow ongoing translation elongation cycles to finish (84) followed by the addition of NMIA. Comparison of median SHAPE reactivity profiles revealed that in cells treated with kasugamycin Ty1 gRNA was less reactive toward NMIA, and thus is likely more structured than under native *in vivo* conditions (Figure 6B). As expected for mRNA undergoing translation, statistically relevant decreases in SHAPE reactivity was observed for ORFs ($P < 0.0001$) but not for UTRs (Figure 6C). To better understand the effect of translating ribosomes on RNA conformation, MFE structures of a +1–2500 fragment of Ty1 RNA were compared in each probing condition. Ty1 gRNA from kasugamycin-treated cells contains more MFE and HP base pairs than RNA *in vivo*, but less than RNA *in vitro* (Figure 6D). The structural model of +1–2500 Ty1 gRNA from kasugamycin-treated cells shared roughly 60% of MFE base pairs with the *in vivo* or *in vitro* RNA model (Figure 6E). Nevertheless, the Ty1 gRNA in kasugamycin-treated cells appeared more similar to the *in vitro* state, since the gRNAs shared more highly probable base pairs with the *in vitro* than *in vivo* structure (70% and 50%, respectively) (Figure 6F). Together, our results show that ribosomes participate in the unfolding and remodeling of Ty1 gRNA *in vivo*, and support the idea that inhibition of translation initiation partially stabilizes Ty1 gRNA structure.

DISCUSSION

We took advantage of the well-developed Ty1 experimental system to define the first genome-wide RNA secondary structure model for a retrovirus-like retrotransposon *in vivo*. To reveal the effect of the cellular environment on folding of the Ty1 RNA genome, we directly compare the SHAPE-derived *in vivo* structure with that obtained under defined *in vitro* conditions. Furthermore, we analyze the role of active ribosomes in RNA structural remodeling by probing Ty1 gRNA structure in yeast treated with a translation initiation inhibitor. We find a strong impact of the cellular environment on folding Ty1 gRNA and show that ribosomes participate in the unfolding and remodeling of Ty1 gRNA.

In support of previous observations made for viral RNAs (2,6) and cellular transcripts (36,53,82), we detect a moderate correlation between the *in vivo* and *in vitro* structure of Ty1 gRNA. About half of the MFE and <40% of HP base pairs are shared between *in vivo* and *in vitro* structure models, indicating significant remodeling of Ty1 gRNA occurs *in vivo*. Base-pairing probabilities and Shannon entropy calculations suggest that Ty1 gRNA adopts a significantly lower and more dynamic structure *in vivo* when compared to the structure *in vitro*. In the *in vivo* state, Ty1 gRNA contains 50% less HP base pairs, and is more likely to form alternative less probable pairings. The observed increase in structural heterogeneity *in vivo* may result from the parallel occurrence of various Ty1 gRNA conformers or from structural transitions through different stages of the replication
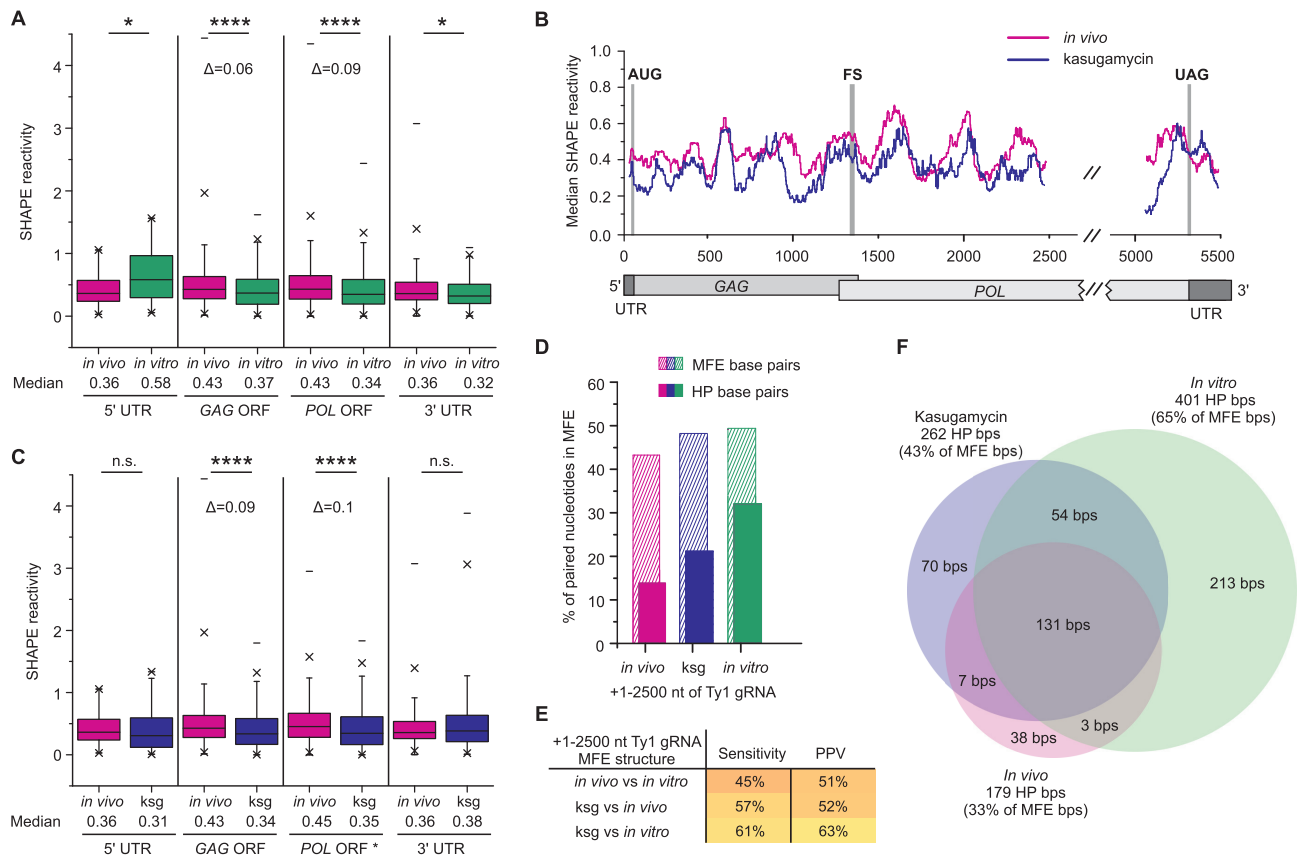
Figure 6. Correlating Ty1 RNA structure and translation. (**A**) Box plot analysis of SHAPE reactivity distributions with medians for coding and untranslated regions of *in vivo* and *in vitro* Ty1 gRNAs. (**B**) Median SHAPE reactivity profiles (smoothed with a 75-nt window) for Ty1 gRNA from cells treated with kasugamycin or not. The start codon (AUG), frameshift site (FS) and stop codon (UAG) were marked by grey strips. (**C**) Box plot analysis of SHAPE reactivity distributions with medians for coding and untranslated regions of Ty1 gRNA from cells treated with kasugamycin-treated (ksg) or not. The *POL* ORF* calculation was performed for fragments presented in panel (**B**). Significance was computed by unpaired two-tailed Mann–Whitney test; *$P < 0.05$, ****$P < 0.0001$, n.s., not significant. (**D**) Percentages of paired nucleotides in the predicted MFE structures of +1–2500 region of Ty1 gRNAs, including high probability base pairs (HP bps). (**E**) Sensitivity and PPV parameters calculated for the MFE structure models of +1–2500 Ty1 gRNA from cells with and without kasugamycin treatment or synthesized *in vitro* (colored as in Figure 3). (**F**) Venn diagram showing overlap of highly probable base pairs between MFE structures of +1–2500 region of Ty1 gRNAs.

cycle. Our results are in agreement with recent studies investigating *in vivo* and *in vitro* structures of viral RNAs, which suggest a more open structure for ZIKV gRNA (2) and a large unfolding of IAV mRNAs in infected cells (6). ZIKV and DENV gRNAs are also less structured in infected cells than *in virio* (3). Consistent with these findings, comparison of *in vivo* SHAPE reactivity data with those obtained previously *in virio* and *ex virio* for the first 1482 nts of Ty1 RNA (33) show that Ty1 gRNA is also less structured in cells than in VLPs (Supplementary Figure S4A). Additionally, the overall similarity between *in virio, ex virio* and *in vitro* MFE structures of Ty1 gRNA is quite high despite the fact that *in virio* RNA is probed in a protein-bound state (Supplementary Figure S4B). Whereas the *in vivo* structure differs significantly from those observed in other experimental states (Supplementary Figure S4B).

Not only are viral RNAs less structured in infected cells, transcriptome-wide structure probing analyses indicate a global decrease in the *in vivo* structure of eukaryotic and prokaryotic mRNAs (36,73,82,85). However, the reasons for RNA structural alteration *in vivo* are poorly understood. Whereas some studies show that ribosomes account

for RNA remodeling *in vivo* (73,82), others suggest a more significant role for RNA binding proteins, ATP-dependent helicases and covalent RNA modifications (36,85,86). To date, the mechanisms underlying viral RNA destabilization *in vivo* have not been thoroughly studied. Our observations indicate that ribosomes play an important role in Ty1 gRNA structure remodeling in yeast. First, the coding region of *in vivo* Ty1 gRNA is less structured than the UTRs. Secondly, kasugamycin treatment leads to stabilization of Ty1 gRNA structure. An analogous effect of translation initiation inhibitors is also observed for *E.coli* and zebrafish mRNAs (73,82). In yeast, a marked increase in mRNA structure is detected under ATP-depleted conditions, and the strong contribution of energy-dependent processes other than translation to mRNA remodeling *in vivo* is proposed (36). However translation is one of the major energy-consuming cellular processes, and upon ATP depletion yeast shut global protein synthesis (87,88). Thus, we conclude that translation is one of several energy-dependent processes contributing to mRNA structure destabilization in yeast. In addition, other cellular and viral factors may be involved in Ty1 gRNA structure remodeling *in vivo*.

Ty1 replication requires direct interactions between the RNA genome and Ty1 Gag (19,38), which displays nucleic acid chaperone activity and likely promotes structural rearrangements of genomic RNA and interactions between RNA partners, analogous to retroviral Gag polyproteins (40,89).

Although Ty1 gRNA is generally less structured *in vivo*, it retains well-structured 5′ and 3′ terminal regions that contain functional *cis*-acting sequences directly involved in dimerization, cyclization and packaging of the Ty1 genome, primer tRNA annealing, and reverse transcription (8,11). However, our results suggest that Ty1 gRNA may contain additional functional *cis*-acting sequences in *POL* where we identify eight well-structured regions. Recent studies show that low-SHAPE and low-entropy analyses can identify previously unknown functional motifs in viral RNA and cellular transcripts (43,74,75,90). Perhaps the stable RNA motifs in *POL* serve as specific binding sites for cellular proteins, as multiple cellular factors influence Ty1 replication (8). Alternatively, these well-structured motifs might contribute to the stability of Ty1 gRNA *in vivo* since *cis*-elements within the coding region help stabilize certain yeast transcripts in addition to the poly(A) tail (91). Although a small fraction of Ty1 transcripts are polyadenylated, Ty1 gRNA half-life is longer than the average half-life of many cellular transcripts (19–21), suggesting the presence of additional stabilizing elements. The correlation between the CDS structure *in vivo* and yeast RNA half-life remains unclear, and codon optimality is proposed as a major determinant of mRNA stability in yeast (92). A significant increase of yeast mRNA half-life is detected when non-optimal codons are converted to optimal codons in a manner minimizing changes in the GC content and the predicted RNA secondary structure (92). Interestingly codon-optimized Ty1 gRNA (CO-Ty1 gRNA) is about as stable as Ty1 gRNA (93). The secondary structure of CO-Ty1 RNA has not been analyzed, but the lack of an increase in stability suggests that the positive effect of codon optimization may be masked by recoding-induced changes in other Ty1 transcript features, such as RNA secondary structure in the CDS. Furthermore, Ty1 gRNA stability is enhanced by interactions with Ty1 Gag (38). Thus, the 1209 base changes in CO-Ty1 gRNA may also induce sequence or structure changes that alter Gag or cellular factor binding. Although extensive sequence recoding in the CO-Ty1 element does not alter the level of retrotransposition or protein expression, there are examples of mutations that strongly inhibit Ty1 mobility located in coding sequences and the UTRs (30,31,76). Indeed, a Rap1 binding site fortuitously introduced into CO-Ty1 during recoding dramatically alters Ty1 gRNA expression (93). Therefore, it stands to reason that RNA structural elements mimicking functions of natural Ty1 gRNA could also have been introduced in CO-Ty1.

Ty1 gRNA also shares structural properties with yeast mRNAs that contribute to stability. There is growing evidence that the low structure (high ΔG) of 5′UTRs and the high structure (low ΔG) of 3′UTRs are associated with longer half-lives and a higher abundance of yeast transcripts (94,95). The 5′UTR of Ty1 RNA has a low GC-content (30%) and is very short (53 nts), consequently, its folding free energy is relatively high. Our data also sup-port the formation of a long stem-loop structure within the 3′UTR of Ty1 RNA *in vivo* and *vitro* (Figure 3C, F and Supplementary Figure S3). Previous work indicates that stable stem-loop motifs in the 3′UTR but outside the poly(A) tail increase the stability of yeast mRNAs (96). Additional analyses will be required to address the function of well-structured regions in *POL* and the 3′ stem-loop motif in the Ty1 life cycle. Future work on RNA folding in the powerful yeast model will also be enhanced by our discovery that diverse SHAPE reagents can be used successfully in growing yeast cultures.

DATA AVAILABILITY

The SHAPE reactivity values are provided in Supplemental Dataset. Other data that support the findings of this study are available from the corresponding author upon reasonable request.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

REFERENCES

1. Boerneke,M.A., Ehrhardt,J.E. and Weeks,K.M. (2019) Physical and functional analysis of viral RNA genomes by SHAPE. *Annu Rev Virol*, **6**, 93–117.
2. Li,P., Wei,Y., Mei,M., Tang,L., Sun,L., Huang,W., Zhou,J., Zou,C., Zhang,S., Qin,C.F. *et al.* (2018) Integrative analysis of Zika virus genome RNA structure reveals critical determinants of viral infectivity. *Cell Host Microbe*, **24**, 875–886.
3. Huber,R.G., Lim,X.N., Ng,W.C., Sim,A.Y.L., Poh,H.X., Shen,Y., Lim,S.Y., Sundstrom,K.B., Sun,X., Aw,J.G. *et al.* (2019) Structure mapping of dengue and Zika viruses reveals functional long-range interactions. *Nat. Commun.*, **10**, 1408.
4. Manfredonia,I., Nithin,C., Ponce-Salvatierra,A., Ghosh,P., Wirecki,T.K., Marinus,T., Ogando,N.S., Snijder,E.J., van Hemert,M.J., Bujnicki,J.M. *et al.* (2020) Genome-wide mapping of SARS-CoV-2 RNA structures identifies therapeutically-relevant elements. *Nucleic Acids Res.*, **48**, 12436–12452.

5. Huston,N.C., Wan,H., de Cesaris Araujo Tavares,R., Wilen,C. and Pyle,A.M. (2020) Comprehensive in-vivo secondary structure of the SARS-CoV-2 genome reveals novel regulatory motifs and mechanisms. doi:10.1101/2020.07.10.197079.

6. Simon,L.M., Morandi,E., Luganini,A., Gribaudo,G., Martinez-Sobrido,L., Turner,D.H., Oliviero,S. and Incarnato,D. (2019) In vivo analysis of influenza A mRNA secondary structures identifies critical regulatory motifs. *Nucleic Acids Res.*, **47**, 7003–7017.

7. Krupovic,M., Blomberg,J., Coffin,J.M., Dasgupta,I., Fan,H., Geering,A.D., Gifford,R., Harrach,B., Hull,R., Johnson,W. *et al.* (2018) Ortervirales: new virus order unifying five families of reverse-transcribing viruses. *J. Virol.*, **92**, e00515-18.

8. Curcio,M.J., Lutz,S. and Lesage,P. (2015) The Ty1 LTR-retrotransposon of budding yeast, *Saccharomyces cerevisiae*. *Microbiol. Spectr.*, **3**, MDNA3-0053-2014.

9. Sandmeyer,S., Patterson,K. and Bilanchone,V. (2015) Ty3, a position-specific retrotransposon in budding yeast. *Microbiol. Spectr.*, **3**, MDNA3-0057-2014.

10. Dodonova,S.O., Prinz,S., Bilanchone,V., Sandmeyer,S. and Briggs,J.A.G. (2019) Structure of the Ty3/Gypsy retrotransposon capsid and the evolution of retroviruses. *Proc. Natl. Acad. Sci. U.S.A.*, **116**, 10048–10057.

11. Pachulska-Wieczorek,K., Le Grice,S.F. and Purzycka,K.J. (2016) Determinants of genomic RNA encapsidation in the Saccharomyces cerevisiae long terminal repeat retrotransposons Ty1 and Ty3. *Viruses*, **8**, 193.

12. Dubois,N., Marquet,R., Paillart,J.C. and Bernacchi,S. (2018) Retroviral RNA Dimerization: from structure to functions. *Front. Microbiol.*, **9**, 527.

13. Menendez-Arias,L., Sebastian-Martin,A. and Alvarez,M. (2017) Viral reverse transcriptases. *Virus Res.*, **234**, 153–176.

14. Sultana,T., Zamborlini,A., Cristofari,G. and Lesage,P. (2017) Integration site selection by retroviruses and transposable elements in eukaryotes. *Nat. Rev. Genet.*, **18**, 292–308.

15. Carr,M., Bensasson,D. and Bergman,C.M. (2012) Evolutionary genomics of transposable elements in *Saccharomyces cerevisiae*. *PLoS One*, **7**, e50978.

16. Curcio,M.J. and Garfinkel,D.J. (1994) Heterogeneous functional Ty1 elements are abundant in the Saccharomyces cerevisiae genome. *Genetics*, **136**, 1245–1259.

17. Malik,H.S., Henikoff,S. and Eickbush,T.H. (2000) Poised for contagion: evolutionary origins of the infectious abilities of invertebrate retroviruses. *Genome Res.*, **10**, 1307–1318.

18. Elder,R.T., Loh,E.Y. and Davis,R.W. (1983) RNA from the yeast transposable element Ty1 has both ends in the direct repeats, a structure similar to retrovirus RNA. *Proc. Natl. Acad. Sci. U.S.A.*, **80**, 2432–2436.

19. Malagon,F. and Jensen,T.H. (2008) The T body, a new cytoplasmic RNA granule in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **28**, 6022–6032.

20. Voytas,D., Boeke,J., Craig,N., Craigie,R., Gellert,M. and Lambowitz,A. (2002) Mobile DNA II.

21. Nonet,M., Scafe,C., Sexton,J. and Young,R. (1987) Eucaryotic RNA polymerase conditional mutant that rapidly ceases mRNA synthesis. *Mol. Cell. Biol.*, **7**, 1602–1611.

22. Curcio,M.J., Hedge,A.M., Boeke,J.D. and Garfinkel,D.J. (1990) Ty RNA levels determine the spectrum of retrotransposition events that activate gene expression in Saccharomyces cerevisiae. *Mol. Gen. Genet.*, **220**, 213–221.

23. Boeke,J.D., Garfinkel,D.J., Styles,C.A. and Fink,G.R. (1985) Ty elements transpose through an RNA intermediate. *Cell*, **40**, 491–500.

24. Devine,S.E. and Boeke,J.D. (1996) Integration of the yeast retrotransposon Ty1 is targeted to regions upstream of genes transcribed by RNA polymerase III. *Genes Dev.*, **10**, 620–633.

25. Friant,S., Heyman,T., Bystrom,A.S., Wilhelm,M. and Wilhelm,F.X. (1998) Interactions between Ty1 retrotransposon RNA and the T and D regions of the tRNA(iMet) primer are required for initiation of reverse transcription in vivo. *Mol. Cell. Biol.*, **18**, 799–806.

26. Garfinkel,D.J., Boeke,J.D. and Fink,G.R. (1985) Ty element transposition: reverse transcriptase and virus-like particles. *Cell*, **42**, 507–517.

27. Asif-Laidin,A., Conesa,C., Bonnet,A., Grison,C., Adhya,I., Menouni,R., Fayol,H., Palmic,N., Acker,J. and Lesage,P. (2020) A small targeting domain in Ty1 integrase is sufficient to direct retrotransposon integration upstream of tRNA genes. *EMBO J.*, **39**, e104337.

28. Xu,H. and Boeke,J.D. (1990) Localization of sequences required in cis for yeast Ty1 element transposition near the long terminal repeats: analysis of mini-Ty1 elements. *Mol. Cell. Biol.*, **10**, 2695–2702.

29. Cristofari,G., Bampi,C., Wilhelm,M., Wilhelm,F.X. and Darlix,J.L. (2002) A 5′-3′ long-range interaction in Ty1 RNA controls its reverse transcription and retrotransposition. *EMBO J.*, **21**, 4368–4379.

30. Gumna,J., Purzycka,K.J., Ahn,H.W., Garfinkel,D.J. and Pachulska-Wieczorek,K. (2019) Retroviral-like determinants and functions required for dimerization of Ty1 retrotransposon RNA. *RNA Biol.*, **16**, 1749–1763.

31. Gamache,E.R., Doh,J.H., Ritz,J., Laederach,A., Bellaousov,S., Mathews,D.H. and Curcio,M.J. (2017) Structure-function model for kissing loop interactions that initiate dimerization of Ty1 RNA. *Viruses*, **9**, 93.

32. Huang,Q., Purzycka,K.J., Lusvarghi,S., Li,D., Legrice,S.F. and Boeke,J.D. (2013) Retrotransposon Ty1 RNA contains a 5′-terminal long-range pseudoknot required for efficient reverse transcription. *RNA*, **19**, 320–332.

33. Purzycka,K.J., Legiewicz,M., Matsuda,E., Eizentstat,L.D., Lusvarghi,S., Saha,A., Le Grice,S.F. and Garfinkel,D.J. (2013) Exploring Ty1 retrotransposon RNA structure within virus-like particles. *Nucleic Acids Res.*, **41**, 463–473.

34. Andrzejewska,A., Zawadzka,M. and Pachulska-Wieczorek,K. (2020) On the way to understanding the interplay between the RNA structure and functions in cells: A genome-wide perspective. *Int. J. Mol. Sci.*, **21**, 6770.

35. Kertesz,M., Wan,Y., Mazor,E., Rinn,J.L., Nutter,R.C., Chang,H.Y. and Segal,E. (2010) Genome-wide measurement of RNA secondary structure in yeast. *Nature*, **467**, 103–107.

36. Rouskin,S., Zubradt,M., Washietl,S., Kellis,M. and Weissman,J.S. (2014) Genome-wide probing of RNA structure reveals active unfolding of mRNA structures in vivo. *Nature*, **505**, 701–705.

37. Moore,S.P., Liti,G., Stefanisko,K.M., Nyswaner,K.M., Chang,C., Louis,E.J. and Garfinkel,D.J. (2004) Analysis of a Ty1-less variant of *Saccharomyces paradoxus*: the gain and loss of Ty1 elements. *Yeast*, **21**, 649–660.

38. Checkley,M.A., Mitchell,J.A., Eizenstat,L.D., Lockett,S.J. and Garfinkel,D.J. (2013) Ty1 gag enhances the stability and nuclear export of Ty1 mRNA. *Traffic*, **14**, 57–69.

39. Vasa,S.M., Guex,N., Wilkinson,K.A., Weeks,K.M. and Giddings,M.C. (2008) ShapeFinder: a software system for high-throughput quantitative analysis of nucleic acid reactivity information resolved by capillary electrophoresis. *RNA*, **14**, 1979–1990.

40. Pachulska-Wieczorek,K., Blaszczyk,L., Biesiada,M., Adamiak,R.W. and Purzycka,K.J. (2016) The matrix domain contributes to the nucleic acid chaperone activity of HIV-2 Gag. *Retrovirology*, **13**, 18.

41. 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.

42. Reuter,J.S. and Mathews,D.H. (2010) RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinformatics*, **11**, 129.

43. Siegfried,N.A., Busan,S., Rice,G.M., Nelson,J.A. and Weeks,K.M. (2014) RNA motif discovery by SHAPE and mutational profiling (SHAPE-MaP). *Nat. Methods*, **11**, 959–965.

44. Darty,K., Denise,A. and Ponty,Y. (2009) VARNA: interactive drawing and editing of the RNA secondary structure. *Bioinformatics*, **25**, 1974–1975.

45. Mathews,D.H. (2004) Using an RNA secondary structure partition function to determine confidence in base pairs predicted by free energy minimization. *RNA*, **10**, 1178–1190.

46. Garfinkel,D.J., Mastrangelo,M.F., Sanders,N.J., Shafer,B.K. and Strathern,J.N. (1988) Transposon tagging using Ty elements in yeast. *Genetics*, **120**, 95–108.

47. Kim,J.M., Vanguri,S., Boeke,J.D., Gabriel,A. and Voytas,D.F. (1998) Transposable elements and genome organization: a comprehensive survey of retrotransposons revealed by the complete Saccharomyces cerevisiae genome sequence. *Genome Res.*, **8**, 464–478.

48. Czaja,W., Bensasson,D., Ahn,H.W., Garfinkel,D.J. and Bergman,C.M. (2020) Evolution of Ty1 copy number control in yeast by horizontal transfer and recombination. *PLos Genet.*, **16**, e1008632.

49. Morillon,A., Benard,L., Springer,M. and Lesage,P. (2002) Differential effects of chromatin and Gcn4 on the 50-fold range of expression among individual yeast Ty1 retrotransposons. *Mol. Cell. Biol.*, **22**, 2078–2088.

50. Merino,E.J., Wilkinson,K.A., Coughlan,J.L. and Weeks,K.M. (2005) RNA structure analysis at single nucleotide resolution by selective 2′-hydroxyl acylation and primer extension (SHAPE). *J. Am. Chem. Soc.*, **127**, 4223–4231.

51. Low,J.T. and Weeks,K.M. (2010) SHAPE-directed RNA secondary structure prediction. *Methods*, **52**, 150–158.

52. Busan,S., Weidmann,C.A., Sengupta,A. and Weeks,K.M. (2019) Guidelines for SHAPE reagent choice and detection strategy for RNA structure probing studies. *Biochemistry*, **58**, 2655–2664.

53. Spitale,R.C., Crisalli,P., Flynn,R.A., Torre,E.A., Kool,E.T. and Chang,H.Y. (2013) RNA SHAPE analysis in living cells. *Nat. Chem. Biol.*, **9**, 18–20.

54. Guo,J.U. and Bartel,D.P. (2016) RNA G-quadruplexes are globally unfolded in eukaryotic cells and depleted in bacteria. *Science*, **353**, aaf5371.

55. Moqtaderi,Z., Geisberg,J.V. and Struhl,K. (2018) Extensive structural differences of closely related 3′ mRNA isoforms: links to Pab1 binding and mRNA stability. *Mol. Cell*, **72**, 849–861.

56. Watters,K.E., Abbott,T.R. and Lucks,J.B. (2016) Simultaneous characterization of cellular RNA structure and function with in-cell SHAPE-Seq. *Nucleic Acids Res.*, **44**, e12.

57. Sztuba-Solinska,J., Rausch,J.W., Smith,R., Miller,J.T., Whitby,D. and Le Grice,S.F.J. (2017) Kaposi's sarcoma-associated herpesvirus polyadenylated nuclear RNA: a structural scaffold for nuclear, cytoplasmic and viral proteins. *Nucleic Acids Res.*, **45**, 6805–6821.

58. Sherpa,C., Rausch,J.W. and Le Grice,S.F. (2018) Structural characterization of maternally expressed gene 3 RNA reveals conserved motifs and potential sites of interaction with polycomb repressive complex 2. *Nucleic Acids Res.*, **46**, 10432–10447.

59. Smola,M.J. and Weeks,K.M. (2018) In-cell RNA structure probing with SHAPE-MaP. *Nat. Protoc.*, **13**, 1181–1195.

60. Lee,B., Flynn,R.A., Kadina,A., Guo,J.K., Kool,E.T. and Chang,H.Y. (2017) Comparison of SHAPE reagents for mapping RNA structures inside living cells. *RNA*, **23**, 169–174.

61. Rice,G.M., Leonard,C.W. and Weeks,K.M. (2014) RNA secondary structure modeling at consistent high accuracy using differential SHAPE. *RNA*, **20**, 846–854.

62. Steen,K.A., Rice,G.M. and Weeks,K.M. (2012) Fingerprinting noncanonical and tertiary RNA structures by differential SHAPE reactivity. *J. Am. Chem. Soc.*, **134**, 13160–13163.

63. Blaszczyk,L., Biesiada,M., Saha,A., Garfinkel,D.J. and Purzycka,K.J. (2017) Structure of Ty1 internally initiated RNA influences restriction factor expression. *Viruses*, **9**, 74.

64. Wan,Y., Kertesz,M., Spitale,R.C., Segal,E. and Chang,H.Y. (2011) Understanding the transcriptome through RNA structure. *Nature reviews. Genetics*, **12**, 641–655.

65. Smola,M.J., Calabrese,J.M. and Weeks,K.M. (2015) Detection of RNA-protein interactions in living cells with SHAPE. *Biochemistry*, **54**, 6867–6875.

66. Kutchko,K.M. and Laederach,A. (2017) Transcending the prediction paradigm: novel applications of SHAPE to RNA function and evolution. *Wiley Interdiscipl. Rev. RNA*, **8**, e1374.

67. Deigan,K.E., Li,T.W., Mathews,D.H. and Weeks,K.M. (2009) Accurate SHAPE-directed RNA structure determination. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 97–102.

68. Weeks,K.M. and Mauger,D.M. (2011) Exploring RNA structural codes with SHAPE chemistry. *Acc. Chem. Res.*, **44**, 1280–1291.

69. Li,P., Zhou,X., Xu,K. and Zhang,Q.C. (2021) RASP: an atlas of transcriptome-wide RNA secondary structure probing data. *Nucleic Acids Res.*, **49**, D183–D191.

70. 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.

71. Giannetti,C.A., Busan,S., Weidmann,C.A. and Weeks,K.M. (2019) SHAPE probing reveals human rRNAs are largely unfolded in solution. *Biochemistry*, **58**, 3377–3385.

72. Smola,M.J., Christy,T.W., Inoue,K., Nicholson,C.O., Friedersdorf,M., Keene,J.D., Lee,D.M., Calabrese,J.M. and Weeks,K.M. (2016) SHAPE reveals transcript-wide interactions, complex structural domains, and protein interactions across the Xist lncRNA in living cells. *Proc. Natl. Acad. Sci. U.S.A.*, **113**, 10322–10327.

73. Mustoe,A.M., Busan,S., Rice,G.M., Hajdin,C.E., Peterson,B.K., Ruda,V.M., Kubica,N., Nutiu,R., Baryza,J.L. and Weeks,K.M. (2018) Pervasive regulatory functions of mRNA structure revealed by high-resolution SHAPE probing. *Cell*, **173**, 181–195.

74. Dethoff,E.A., Boerneke,M.A., Gokhale,N.S., Muhire,B.M., Martin,D.P., Sacco,M.T., McFadden,M.J., Weinstein,J.B., Messer,W.B., Horner,S.M. *et al.* (2018) Pervasive tertiary structure in the dengue virus RNA genome. *Proc. Natl. Acad. Sci. U.S.A.*, **115**, 11513–11518.

75. Mauger,D.M., Golden,M., Yamane,D., Williford,S., Lemon,S.M., Martin,D.P. and Weeks,K.M. (2015) Functionally conserved architecture of hepatitis C virus RNA genomes. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, 3692–3697.

76. Bolton,E.C., Coombes,C., Eby,Y., Cardell,M. and Boeke,J.D. (2005) Identification and characterization of critical cis-acting sequences within the yeast Ty1 retrotransposon. *RNA*, **11**, 308–322.

77. Belcourt,M.F. and Farabaugh,P.J. (1990) Ribosomal frameshifting in the yeast retrotransposon Ty: tRNAs induce slippage on a 7 nucleotide minimal site. *Cell*, **62**, 339–352.

78. Watts,J.M., Dang,K.K., Gorelick,R.J., Leonard,C.W., Bess,J.W. Jr, Swanstrom,R., Burch,C.L. and Weeks,K.M. (2009) Architecture and secondary structure of an entire HIV-1 RNA genome. *Nature*, **460**, 711–716.

79. Collart,M.A. and Weiss,B. (2020) Ribosome pausing, a dangerous necessity for co-translational events. *Nucleic Acids Res.*, **48**, 1043–1055.

80. Lange,S.J., Maticzka,D., Mohl,M., Gagnon,J.N., Brown,C.M. and Backofen,R. (2012) Global or local? Predicting secondary structure and accessibility in mRNAs. *Nucleic Acids Res.*, **40**, 5215–5226.

81. Kleiman,L., Jones,C.P. and Musier-Forsyth,K. (2010) Formation of the tRNALys packaging complex in HIV-1. *FEBS Lett.*, **584**, 359–365.

82. Beaudoin,J.D., Novoa,E.M., Vejnar,C.E., Yartseva,V., Takacs,C.M., Kellis,M. and Giraldez,A.J. (2018) Analyses of mRNA structure dynamics identify embryonic gene regulatory programs. *Nat. Struct. Mol. Biol.*, **25**, 677–686.

83. Schuwirth,B.S., Day,J.M., Hau,C.W., Janssen,G.R., Dahlberg,A.E., Cate,J.H. and Vila-Sanjurjo,A. (2006) Structural analysis of kasugamycin inhibition of translation. *Nat. Struct. Mol. Biol.*, **13**, 879–886.

84. Riba,A., Di Nanni,N., Mittal,N., Arhne,E., Schmidt,A. and Zavolan,M. (2019) Protein synthesis rates and ribosome occupancies reveal determinants of translation elongation rates. *Proc. Natl. Acad. Sci. U.S.A.*, **116**, 15023–15032.

85. Spitale,R.C., Flynn,R.A., Zhang,Q.C., Crisalli,P., Lee,B., Jung,J.W., Kuchelmeister,H.Y., Batista,P.J., Torre,E.A., Kool,E.T. *et al.* (2015) Structural imprints in vivo decode RNA regulatory mechanisms. *Nature*, **519**, 486–490.

86. Sun,L., Fazal,F.M., Li,P., Broughton,J.P., Lee,B., Tang,L., Huang,W., Kool,E.T., Chang,H.Y. and Zhang,Q.C. (2019) RNA structure maps across mammalian cellular compartments. *Nat. Struct. Mol. Biol.*, **26**, 322–330.

87. Stouthamer,A.H. (1973) A theoretical study on the amount of ATP required for synthesis of microbial cell material. *Antonie Van Leeuwenhoek*, **39**, 545–565.

88. Ashe,M.P., De Long,S.K. and Sachs,A.B. (2000) Glucose depletion rapidly inhibits translation initiation in yeast. *Mol. Biol. Cell*, **11**, 833–848.

89. Rein,A. (2010) Nucleic acid chaperone activity of retroviral Gag proteins. *RNA Biol*, **7**, 700–705.

90. Dethoff,E.A. and Weeks,K.M. (2019) Effects of refolding on large-scale RNA structure. *Biochemistry*, **58**, 3069–3077.

91. Vemula,M., Kandasamy,P., Oh,C.S., Chellappa,R., Gonzalez,C.I. and Martin,C.E. (2003) Maintenance and regulation of mRNA stability of the *Saccharomyces cerevisiae* OLE1 gene requires multiple elements within the transcript that act through translation-independent mechanisms. *J. Biol. Chem.*, **278**, 45269–45279.

92. Presnyak,V., Alhusaini,N., Chen,Y.H., Martin,S., Morris,N., Kline,N., Olson,S., Weinberg,D., Baker,K.E., Graveley,B.R. *et al.* (2015) Codon optimality is a major determinant of mRNA stability. *Cell*, **160**, 1111–1124.

93. Yarrington,R.M., Richardson,S.M., Lisa Huang,C.R. and Boeke,J.D. (2012) Novel transcript truncating function of Rap1p revealed by synthetic codon-optimized Ty1 retrotransposon. *Genetics*, **190**, 523–535.

94. Ringner,M. and Krogh,M. (2005) Folding free energies of 5′-UTRs impact post-transcriptional regulation on a genomic scale in yeast. *PLoS Comput. Biol.*, **1**, e72.

95. Aw,J.G., Shen,Y., Wilm,A., Sun,M., Lim,X.N., Boon,K.L., Tapsin,S., Chan,Y.S., Tan,C.P., Sim,A.Y. *et al.* (2016) In vivo mapping of Eukaryotic RNA interactomes reveals principles of higher-order organization and regulation. *Mol. Cell*, **62**, 603–617.

96. Geisberg,J.V., Moqtaderi,Z., Fan,X., Ozsolak,F. and Struhl,K. (2014) Global analysis of mRNA isoform half-lives reveals stabilizing and destabilizing elements in yeast. *Cell*, **156**, 812–824.