

Transcriptome-wide RNA structure probing with temporal resolution

Danyang Wang¹, Yongkang Tang², Ang Li², Lei Sun², & Ralph E. Kleiner^{1*}

¹Department of Chemistry, Princeton University, Princeton, NJ 08544, USA

²Shandong Provincial Key Laboratory of Animal Cell and Developmental Biology, School of Life Sciences, Shandong University, Qingdao 266237, China.

*email: rkleiner@princeton.edu

ABSTRACT

Reactive small-molecule probes are widely used for RNA structure probing, however current approaches largely measure average RNA transcript dynamics and do not resolve structural differences that occur during folding or transcript maturation. Here, we present SNIPER-seq, an RNA structure probing method relying upon metabolic labeling with 2'-aminodeoxycytidine, structure-dependent 2'-amino reaction with an aromatic isothiocyanate, and high-throughput RNA sequencing. Our method maps cellular RNA structure transcriptome-wide with temporal resolution enabling determination of transcript age-dependent RNA structural dynamics. We benchmark our approach against known RNA structures and investigate the dynamics of human 5S rRNA during ribosome biogenesis, revealing specific structural changes in 5S rRNA loops that occur over the course of several hours. Taken together, our work sheds light on the maturation and coordinated conformational changes that take place during ribosome biogenesis and provides a general strategy for surveying evolving RNA structural dynamics across the transcriptome.

INTRODUCTION

The folding of RNA into complex three-dimensional structures is essential for its biological function. The development of chemical methods to probe RNA structure *in vitro* and in cells has provided key insights into RNA biology. Currently, a number of different chemistries are employed to analyze RNA structure including nucleobase alkylation by DMS¹, carbodiimides² or glyoxal³ and 2'-hydroxyl acylation by isatoic anhydrides⁴⁻⁶ or acylimidazoles⁷⁻⁹. When combined with high-throughput RNA sequencing workflows that can detect lesion-induced misincorporation or termination during reverse transcription, chemical structure probing methods can reveal transcriptome-wide RNA structure in living cells^{5, 10}. Innovations in chemistry, sequencing methods, library preparation, and bioinformatic analysis have contributed greatly to the widespread application of these methods to illuminate diverse phenomena in RNA biology¹¹.

RNA structure can be modulated through binding of proteins or small molecules, alterations in environmental conditions such as temperature, oxidative stress, or solute concentration, and post-transcriptional modifications¹². Coding and non-coding RNAs are often engaged in ribonucleoprotein (RNP) complexes, and the organization of these macromolecular assemblies, which is coordinated with RNA folding and post-transcriptional RNA processing/regulation can be highly dynamic¹³. Further, RNA folds co-transcriptionally, and therefore conditions that affect transcription can impact structural ensembles¹⁴. Current chemical probing methods measure the population average and obscure differences in behavior between mature RNA and newly synthesized RNA transcripts. Whereas RNA secondary and tertiary structure forms rapidly on the μ s to sec timescale¹⁵, and is better suited for *in vitro* analyses¹⁶, RNA/RNP processing, maturation, trafficking, and ultimately metabolism occur at slower timescales. For example, rough estimates of ribosome biogenesis in yeast and human cells derived from pulse labeling experiments range from minutes to hours¹⁷⁻¹⁹, making this important process amenable to characterization in cellular systems. Notably, ribosome assembly pathways and intermediates in eukaryotic systems have

been extensively studied using *in vitro* biochemical and structural approaches²⁰, but less is known about the structure and dynamics of ribosomal biogenesis intermediates in cells. Therefore, methods to increase the temporal resolution of structure probing analyses can provide new insights into molecular mechanisms underlying the dynamic behavior and regulation of RNA in its native context.

Metabolic labeling with modified nucleosides is a powerful method to introduce a temporal dimension into studies of RNA expression and dynamics²¹. A variety of ribonucleoside analogues have been introduced to enable analysis of nascent RNA through affinity isolation or direct detection, but few have been employed in RNA structure probing workflows. Isolation of nascent RNA can be coupled with reported chemical probing strategies²², but this approach has not been widely used. 4-thiouridine (4-SU) and 4-thiouracil (4-TU) labeling and nascent RNA isolation have been used to facilitate DMS or hydroxyl radical probing of pre-rRNA intermediates in *E. coli* and yeast^{23, 24}, but this strategy has not been applied transcriptome-wide or in human cells, and 4-SU/TU is known to perturb nucleolar morphology and rRNA processing²⁵. Alternatively, the metabolic label could serve as both an affinity isolation tag and as a unique chemical handle for reactivity-based structure probing, but to our knowledge this approach has not been realized.

Here, we present synchronized nucleoside incorporation and probing evaluated by reverse transcription and sequencing (SNIPER-seq), an RNA structure probing method based upon metabolic RNA labeling with 2'-aminodeoxynucleosides combined with selective 2'-amino functionalization chemistry, enrichment, and primer extension or sequencing analysis. We demonstrate that SNIPER-seq is compatible with *in vitro* and in-cell RNA structure probing and can be applied for transcriptome-wide experiments as well as targeted analysis of non-coding RNAs. Further, SNIPER-seq enables the interrogation of RNA structure at different stages in the lifetime of a transcript, and we leverage this ability to investigate the structure of 5S rRNA during

ribosomal assembly. Taken together, our strategy provides a new approach to measure transcriptome-wide RNA structure in living cells across the temporal dimension.

RESULTS

Quantitative analysis of 2'-amine acylation in synthetic oligonucleotides

To develop our method, we needed to identify a selective RNA functionalization reaction that reports on native structure and involves functional groups that can be introduced through metabolic labeling. Previously, Weeks and co-workers demonstrated that 2'-amino-modified RNA adopts native-like C3'-endo geometry and displays structure-dependent reactivity with NHS-esters analogous to labeling of native 2'-OH RNA with acylating agents²⁶. In addition, our lab and others have previously shown that 2'-azidodeoxynucleosides can be incorporated into RNA metabolically^{27, 28} and that 2'-azido substitutions do not dramatically perturb RNA structure²⁹, suggesting that small modifications at the 2'-position are metabolically tolerated and native-like. Therefore, we chose to pursue 2'-aminodeoxynucleosides as metabolically incorporated structure probes.

First, we synthesized stem-loop RNA oligonucleotides containing a single 2'-aminodeoxycytidine modification (2'-NH₂-C) in either the stem (oligo **2**) or the loop (oligo **3**) to evaluate structure-dependent reactivity with different amine-reactive probes (Fig. 1A and Supplementary Fig. 1). A control oligonucleotide (oligo **1**) that contained only unmodified RNA was also generated. We designed three probes containing an azide reporter group for enrichment and labeling via strain-promoted azide-alkyne cycloaddition (SPAAC)³⁰ and different amine-reactive warheads: 4-sulfotetrafluorophenyl ester (STP-N₃), N-hydroxysuccinimidyl ester (NHS-N₃) or aromatic isothiocyanate (isoS-N₃) (Fig. 1B). STP, NHS, and isothiocyanates are widely used for labeling primary amines in polypeptides³¹, and sulfo-NHS³² and isothiocyanates³³ have been shown to

react with 2'-amine-containing RNA *in vitro* with minimal 2'-OH reactivity. Oligonucleotides (1 μ M) were treated with an excess (5-20 mM) of the respective electrophilic probe for 15 min and then reacted with DBCO-biotin to generate biotinylated and mobility-shifted products (Fig. 1C). We quantified labeling efficiency of each electrophilic probe against oligos **1**, **2**, and **3** by densitometry. For all three electrophilic probes, we observed a similar trend in reactivity with unpaired 2'-NH₂-C (oligo **3**) consistently showing higher reactivity than base paired 2'-NH₂-C (oligo **2**), and the unmodified RNA (oligo **1**) showing lowest reactivity (Fig. 1C and 1D). While labeling of unmodified oligo **1** was low using NHS-N₃ and STP-N₃, with 8% or 4% conversion after 20 min, respectively, labeling with isoS-N₃ was undetectable, indicating that isothiocyanate has better specificity for the 2'-amino group (Fig. 1D). The isoS-N₃ probe also demonstrated the highest labeling efficiency (80% conversion of oligo **3**) and comparable selectivity for unpaired 2'-NH₂-C over base-paired 2'-NH₂-C (Fig. 1E). We performed a time-course labeling of the three oligos with isoS-N₃ and found that oligo **3** achieved 40% conversion in 2 min and reached a plateau after ~15 min (Fig. 1F, Supplementary Fig. 2). Taken together, we find that the isoS-N₃ probe displays superior reactivity and specificity for unpaired 2'-NH₂-C nucleotides of the three probes that we assayed, and therefore it was selected for further studies.

2'-NH₂-C-based structure probing of tRNA^{Cys}

We further investigated whether reaction of 2'-NH₂-C with isoS-N₃ could be used to measure nucleotide flexibility in a biologically relevant RNA. Based on precedent³², we incorporated 2'-NH₂-C modifications into tRNA^{Cys} by *in vitro* transcription (IVT) with the corresponding modified NTP to yield, on average, a single 2'-NH₂-C per tRNA transcript, and then measured reactivity with isoS-N₃. We observed labeling of 2'-NH₂-C-modified tRNA, but not unmodified tRNA, under denaturing conditions and in tRNA folding buffer (Supplementary Fig. 3), indicating that 2'-NH₂-C is required for reactivity. Further, labeling with isoS-N₃ was more efficient under denaturing

conditions suggesting that RNA structure inhibits reaction kinetics. We next analyzed isoS-N₃ labeling of 2'-NH₂-C-modified tRNA^{Cys} by reverse transcription (RT)-primer extension (Fig. 1G). Consistent with known tRNA structure, we primarily observed RT termination at C residues in the anti-codon stem loop (ASL) (Fig. 1H), particularly at C35 in the anti-codon, indicating this position has the highest flexibility; in contrast, cytidine residues in the D-arm or T Ψ C-arm, which are involved in tertiary interactions, displayed lower reactivity.

Metabolic labeling with 2'-aminodeoxynucleosides and in vitro RNA structure probing

We next applied our approach to perform structure probing on native RNA transcripts. For this purpose, we tested metabolic incorporation of 2'-aminodeoxypyrimidine nucleosides. We treated HEK293T cells with 2'-aminodeoxycytidine (2'-NH₂-dCyd) or 2'-aminodeoxyuridine (2'-NH₂-dUrd) for 16 hr and analyzed incorporation into total RNA by nucleoside LC-QQQ-MS (Fig. 2A, 2B, 2C, Supplementary Table 2 and 3). Gratifyingly, 2'-NH₂-dCyd feeding resulted in 2'-NH₂-C accumulation to 6% of unmodified C in total RNA, which is comparable or better to widely used nucleoside analogues such as 5-EU³⁴ or 4-SU³⁵. In addition, we also detected incorporation of 2'-NH₂-U, likely resulting from deamination of 2'-NH₂-dCyd. Direct feeding of 2'-NH₂-dUrd resulted in ~1.5% incorporation of 2'-NH₂-U and some metabolic interconversion to 2'-NH₂-C (Fig. 2B and 2C). Notably, neither modified nucleoside showed overt cellular toxicity (Supplementary Fig. 5) or effects on levels of native RNA pyrimidine post-transcriptional modifications (Supplementary Fig. 6). We also characterized nucleolar integrity in 2'-NH₂-dCyd treated cells by immunostaining with NPM1 (Supplementary Fig. 7). In contrast to 4-SU, a widely used metabolic label that inhibits rRNA synthesis and induces nucleolar stress (visible by diffuse NPM1 staining), 2'-NH₂-dCyd-treated cells showed largely unperturbed nucleolar morphology at the concentrations and treatment times evaluated.

After isolation of 2'-NH₂-dCyd labeled total RNA, we evaluated *in vitro* labeling with isoS-N₃, STP-N₃, and NHS-N₃ under denaturing conditions (Supplementary Fig. 8). While reactivity of 2'-NH₂-modified RNA with all three probes was comparable, we observed higher background labeling of native RNA with either NHS-N₃ or STP-N₃, similar to our findings with synthetic oligos (Fig. 1C). We next performed a time-course analysis of the reactivity of refolded 2'-NH₂-C-labeled total RNA with isoS-N₃. As expected, we found that reactivity was less than under denaturing conditions but increased in a time-dependent manner from 2 min to 30 min (Fig. 2D). To verify the ability of our method to probe native RNA transcript structure, we investigated the 5S rRNA, which is highly abundant and can fold into a stable structure without the need for protein cofactors³⁶. After isoS-N₃ labeling of purified total RNA, we enriched reacted RNA through SPAAC with DBCO-disulfide-biotin and streptavidin pulldown, and measured labeling by RT-primer extension. We observed multiple RT stops corresponding to unpaired residues in 5S rRNA including bulges (C63) and loops (C26 in loop B, C34 and C39 in loop C) (Fig. 2E and 2F). Interestingly, we also found that cytidine residues at the joint of loop and stem structures showed high reactivity (i.e. C52, C57 and C67). This phenomenon was also reported when the SHAPE approach was applied to probe 5S rRNA structure⁷ and to probe structure within an RNA oligo library³⁷.

RNA structure probing in living cells using 2'-NH₂-dCyd and isoS-N₃

Encouraged by our *in vitro* structure probing results, we investigated whether isoS-N₃ could probe RNA structure in living cells labeled with 2'-NH₂-dCyd. We incubated HEK293T cells with 2'-NH₂-dCyd overnight and treated cells with 5 mM isoS-N₃. We then isolated total RNA and measured labeling efficiency using DBCO-Cy3. We were able to detect labeling of 2'-NH₂-containing RNA with isoS-N₃ as early as 2 min and reaching a plateau at ~15 min (Fig. 3A), indicating that isoS-N₃ is cell permeable and able to react with 2'-NH₂-modified RNA in living cells. The labeling kinetics for isoS-N₃ are comparable to widely used SHAPE reagents such as NAI or NAI-N₃, albeit

at 20-fold lower reaction concentration^{7,11}. We again used 5S rRNA to investigate isoS-N₃ reaction in living cells by RT-primer extension as described above and found major RT stops at several C residues including C63, C57, C52, C39 and C91 (Fig. 3B). Analysis of the x-ray crystal structure of the human 80S ribosome, which contains the 5S rRNA and its interacting proteins (PDB ID: 4UG0), showed that these cytidines are located at loop positions and highly exposed to solvent (Fig. 3C).

We next combined metabolic labeling and isoS-N₃ reaction with Illumina sequencing to study transcriptome-wide RNA structure (SNIPER-seq). Reaction of 2'-NH₂-modified RNA with isoS-N₃ was either performed *in vitro* or *in cellulo*, and unreacted native RNA was used as control. Illumina sequencing libraries were prepared following the icSHAPE method³⁸ with biotin-based enrichment of labeled RNA fragments, and isoS-N₃ reactivity was inferred by analyzing RT stop signatures. Two independent biological replicates were used for all conditions. Compared with untreated RNA samples, in which RT stops were equally distributed across all four nucleobases, we found that 2'-NH₂-dCyd-labeled and isoS-N₃ treated samples predominantly contained RT stops at C (70-80% of RT stops) (Fig. 3D, Supplementary Fig. 9A). We also found that RT stops in the 2'-NH₂-dCyd samples occurred more frequently at U (13-14%) than at A/G (6%), likely due to metabolic conversion of 2'-NH₂-dCyd to 2'-NH₂-dUrd, as shown previously by LC-MS analysis (Fig. 2C). Overall, >90% of RT stops occurred at C/U, demonstrating the selectivity of the isoS-N₃ labeling reaction. Further, we observed a high correlation of RT-stop levels at individual sites between the two *in cellulo* replicates, with Spearman correlation coefficient $r = 0.77$, indicating that our method is robust and reproducible (Fig. 3E).

To investigate the ability our method to recapitulate complex cellular RNA structure, we chose abundant RNA transcripts with known 3D structures for analysis. We compared the reactivity profile of SNIPER-seq on all C residues in the 18S rRNA (length of 1869 nt) and 28S rRNA (length of 5348 nt) with the known structure model collected from the RiboVision database³⁹, which

determines secondary structure from 3D structures. SNIPER-seq reactivity profiles fit the 18S or 28S rRNA models with high area under curve (AUC) of 0.719 and 0.683, respectively (Fig. 3F and 3G). These AUC values are comparable with those obtained using icSHAPE (AUC = 0.697 for 18S rRNA, AUC = 0.633 for 28S rRNA) (Fig. 3H)⁴⁰. In addition to rRNAs, a region of the long noncoding RNA MALAT1 was also used as a model to exam the accuracy of SNIPER-seq as its structure has been previously studied with SHAPE⁴¹ and it was highly abundant in our sequencing data; in this context, we also found that SNIPER-seq reactivity scores were similar to those previously determined using SHAPE and consistent with predicted structural motifs (Supplementary Fig. 9B). Taken together, our data demonstrates that SNIPER-seq can be used for *in cellulo* RNA structure probing at pyrimidine nucleosides across the transcriptome.

Measuring transcript age-dependent structural changes with SNIPER-seq

We explored whether we could use SNIPER-seq to study time-dependent structural changes occurring in RNA in a pulse-chase experiment with 2'-NH₂-dCyd (Fig. 4A). The biogenesis of ribosomes involves the synthesis, processing, and assembly of multiple RNA transcripts including 28S, 18S, 5.8S, and 5S rRNA²⁰. 28S, 18S, and 5.8S rRNAs are processed from a single 45S rRNA transcript generated in the nucleolus by RNA Pol I, whereas the 5S rRNA is transcribed from a separate locus (outside of nucleoli) by RNA Pol III before it forms the 5S ribonucleoprotein particle (5S RNP) and is imported into the nucleolus and incorporated into the pre-60S particle. The 5S RNP then undergoes 180 degree rotation to form the central protuberance (CP) structure that is found in the mature 60S subunit. Whereas the order of these steps in yeast and human ribosome biogenesis have been established through *in vitro* structural studies^{42, 43}, their timing in cells and the native distribution of intermediates is poorly established. We fed cells 2'-NH₂-dCyd for 1 hr to label newly transcribed RNA, performed a chase for 3, 6, or 12 hr, and then reacted cells with isoS-N₃ at the end of the chase. We first analyzed isoS-N₃ labeling by performing

targeted RT-primer extension of 5S rRNA. Interestingly, we observed differences in reactivity at different chase time points (Fig. 4A). In particular, reactivity at positions C39 in loop C and C91 in loop D increased as a function of chase time, while labeling at C34 in loop C and C78 in helix V decreased in a time-dependent manner. In contrast, the reactivity at positions C52, C57, and C63 remained constant over the course of our experiment.

To further validate differences in nucleotide reactivity on the 5S rRNA, we measured adduct formation by mutational profiling immediately after 1 hr 2'-NH₂-dCyd pulse labeling or after 1 hr pulse labeling followed by a 6 hr chase. We reverse transcribed RNA in the presence of Mn²⁺, a condition known to induce read through with mutation at modified nucleotide adducts^{44, 45}, and performed high-throughput sequencing of the 5S rRNA amplicon. Total mutation rate at each nucleotide was calculated by combining the frequency of substitutions, deletions, and insertions and subtracting the mutation rate present in the control. In the 1 hr pulse sample, we found mutation rates above 5% at residues C29, C30, C39, C57, C91 and C95 (Fig. 4B and Supplementary Fig. 10). After a 6 hr chase (Fig. 4A), the most dramatic change in reactivity occurred at C39 in loop C, which showed a 4-fold increase in mutational frequency (this position also showed a time-dependent increase in reactivity in the gel-based RT stop analysis). We also found a time-dependent 1.7-fold increase in mutation rate at C36 in loop C and a 1.6-fold increase in mutation at C95 in loop D. In contrast, mutational frequencies at C30 in helix III and C57 in loop B remained largely unchanged during the chase, suggesting that they undergo only minimal changes in structure and reactivity. To understand the origin of reactivity differences at C39 in loop C, we examined structures of the 5S RNP from pre-60S subunit intermediates to the mature 80S ribosome in yeast and human^{42, 43}. Whereas the secondary and tertiary structure of the 5S rRNA is largely conserved throughout ribosome biogenesis, (Supplementary Fig. 11A), its interaction with ribosomal proteins is dynamic. During the initial association of 5S RNP with the pre-60S ribosome (State I1, PDB ID: 8FL2)⁴², loop C is situated within a binding pocket formed

by two ribosomal protein, L5 and L11, and nucleolar protein SDAD1 (Fig. 4C). After 180 degree rotation and in the mature 60S/80S (PDB ID: 4UG0)⁴⁶, loop C becomes more solvent exposed (Fig. 4D). We also found a hydrogen bond between the O2 atom of C39 and amino acid residue Thr73 in L11 within the pre-rotated state (Fig. 4C, inset). After 5S RNP rotation, C39 drifts apart from L11, losing the hydrogen bond between C39 and Thr73 (Fig. 4D, inset). We propose that the increased solvent accessibility and loss of hydrogen bond in the mature 5S RNP state are responsible for the increased reactivity in loop C observed in our SNIPER-seq experiment. Notably, we also found increased solvent accessibility in this region in structural intermediates from yeast ribosomal assembly (Supplementary Fig. 11B, 11C)^{43, 47}.

We were also curious how the folding and maturation of nascent rRNA is affected by stress conditions. Oxidative stress can affect rRNA processing⁴⁸ and heat stress has been shown to unfold mRNA⁴⁹ and tRNA⁵⁰ structure. We performed SNIPER-seq on newly transcribed 5S rRNA under oxidative stress (NaAsO₂) or heat stress conditions together with mutational profiling (Fig. 4E, 4F). Compared to the unstressed control, oxidative stress reduced isoS-N₃ reactivity at multiple sites, including C28 in helix III, C36 and C39 in loop C, and C91 in loop D (Fig. 4E, Fig. 2F). The reactivity of C26 and C57 in loop B remained at the same level. In contrast, only C28 showed decreased reactivity under heat stress, whereas C57 exhibited slightly higher reactivity. Upon nucleolar stress, such as heat shock, serum starvation and oxidative stress, integration of 5S RNP into the pre-60S is decreased⁵¹, and 5S RNP accumulates in the nucleoplasm, forming interactions with the E3 ubiquitin ligase HDM2⁵². These processes may collectively contribute to the observed differences in 5S rRNA structural probing results under stress conditions. Taken together, our study demonstrates how age-dependent changes in cellular RNA structure can be revealed by SNIPER-seq, providing a versatile method to study RNP maturation and condition-dependent structural changes.

DISCUSSION

In this study, we present a novel method for probing RNA structure in living cells with the ability to selectively study RNA transcripts as a function of their lifetime. Our method, SNIPER-seq, relies upon metabolic labeling of cellular RNA with 2'-NH₂-dCyd, which is incorporated at high levels through native metabolic pathways with minimal cytotoxicity. Metabolic labeling facilitates temporal resolution of specific RNA populations and the 2'-amino modification serves as a non-perturbing structure-dependent reactivity handle for labeling and enrichment with a cell-permeable isothiocyanate probe. We validate our method using RNAs of known structure and analyze the age-dependent structure of 5S rRNA over the course of ~1-12 hr after its transcription to reveal dramatic and region-specific structure and reactivity differences in this RNA species during its lifetime.

SNIPER-seq endows reactivity-based RNA structure probing methods with temporal resolution. In this work we demonstrate how our method can be used to study the dynamics of RNA maturation within the context of ribosome biogenesis. Moving forward, we envision broad application of this approach to study age-dependent RNA dynamics transcriptome-wide on coding and non-coding RNAs. Whereas our method may currently lack the ability to characterize co-transcriptional folding (which typically occurs on the μ s-sec time scale), structural dynamics that result from changes in protein association, complex assembly, or cellular localization during the lifetime of an RNA should be amenable to our approach, thereby establishing the absolute timing of these processes. Further, pulse-labeling allows the specific interrogation of perturbations to co-transcriptional folding or RNA processing pathways that result in longer-lived structural intermediates that can be probed. Our method therefore provides a complementary approach to structure probing methods that operate on the entire transcriptome and should provide new insights into RNA structure and interactomes in cells.

We applied SNIPER-seq to study the age-dependent dynamics in the 5S rRNA. We measure changes in RNA reactivity at specific loop residues that occur over the span of multiple hours, such as the dramatic increase in reactivity in loop C. RNA pulse labeling in human cells has established rough estimates for rRNA maturation^{17, 19}, however the absolute timing of specific events and intermediates in human ribosome biogenesis are still largely unknown. Our study suggests that conformational changes in the 5S rRNA occur later than estimates for maturation of the 18S and 28S rRNA. In contrast to other canonical rRNAs that are generated from the 45S rRNA, 5S is transcribed from a separate locus outside of the nucleolus by RNA Pol III and incorporated into the pre-60S subunit later as the 5S RNP, consistent with slower maturation kinetics. Alternatively, the observed structural dynamics for 5S may reflect conformational changes occurring in cytosolic ribosomes.

Metabolic labeling with modified nucleosides has been widely used to install orthogonal chemical functionalities for RNA detection and affinity isolation. Here we use this approach to directly incorporate an artificial functional group that reports on RNA structure. Whereas the 2'-amino group is metabolically tolerated and compatible with native RNA structure, its reactivity is greatly enhanced relative to 2'-OH or any of the nucleobase exocyclic amines present on RNA, which enables selective functionalization of modified nucleotides. Here we present our strategy using 2'-NH₂-dCyd, but it is likely that other 2'-amino modified nucleosides will also be compatible with our approach, expanding its utility to probe all 4 nucleotides; 2'-thionucleosides have also been incorporated into RNA transcripts⁵³ and are likely compatible with cellular metabolism and our approach. Further, we envision that perturbation of nucleotide metabolism to increase flux through the nucleoside salvage pathways may offer a strategy to enhance nucleoside uptake and provide even higher temporal resolution. Such studies are underway in our group and will be reported in due course.

ACKNOWLEDGMENTS

We thank D. Dhinghani for assistance with bioinformatic analysis. R.E.K. acknowledges support from the NIH (R01 GM132189) and NSF (MCB-1942565). R.E.K., D.W., and A.L. acknowledge financial support from Princeton University.

COMPETING INTERESTS

The authors declare no competing financial interests.

REFERENCES

- (1) Peattie, D. A.; Gilbert, W. Chemical probes for higher-order structure in RNA. *Proc Natl Acad Sci U S A* **1980**, 77 (8), 4679-4682. DOI: 10.1073/pnas.77.8.4679 From NLM.
- (2) Ehresmann, C.; Baudin, F.; Mougél, M.; Romby, P.; Ebel, J. P.; Ehresmann, B. Probing the structure of RNAs in solution. *Nucleic Acids Res* **1987**, 15 (22), 9109-9128. DOI: 10.1093/nar/15.22.9109 From NLM.
- (3) Mitchell, D., 3rd; Ritchey, L. E.; Park, H.; Babitzke, P.; Assmann, S. M.; Bevilacqua, P. C. Glyoxals as in vivo RNA structural probes of guanine base-pairing. *Rna* **2018**, 24 (1), 114-124. DOI: 10.1261/rna.064014.117 From NLM.
- (4) Merino, E. J.; Wilkinson, K. A.; Coughlan, J. L.; Weeks, K. M. RNA structure analysis at single nucleotide resolution by selective 2'-hydroxyl acylation and primer extension (SHAPE). *J Am Chem Soc* **2005**, 127 (12), 4223-4231. DOI: 10.1021/ja043822v From NLM.
- (5) Ding, Y.; Tang, Y.; Kwok, C. K.; Zhang, Y.; Bevilacqua, P. C.; Assmann, S. M. In vivo genome-wide profiling of RNA secondary structure reveals novel regulatory features. *Nature* **2014**, 505 (7485), 696-700. DOI: 10.1038/nature12756 From NLM.
- (6) Mortimer, S. A.; Weeks, K. M. A fast-acting reagent for accurate analysis of RNA secondary and tertiary structure by SHAPE chemistry. *J Am Chem Soc* **2007**, 129 (14), 4144-4145. DOI: 10.1021/ja0704028 From NLM.
- (7) Spitale, R. C.; Crisalli, P.; Flynn, R. A.; Torre, E. A.; Kool, E. T.; Chang, H. Y. RNA SHAPE analysis in living cells. *Nat Chem Biol* **2013**, 9 (1), 18-20. DOI: 10.1038/nchembio.1131 From NLM.
- (8) Xiao, L.; Fang, L.; Chatterjee, S.; Kool, E. T. Diverse Reagent Scaffolds Provide Differential Selectivity of 2'-OH Acylation in RNA. *J Am Chem Soc* **2023**, 145 (1), 143-151. DOI: 10.1021/jacs.2c09040 From NLM.
- (9) Marinus, T.; Fessler, A. B.; Ogle, C. A.; Incarnato, D. A novel SHAPE reagent enables the analysis of RNA structure in living cells with unprecedented accuracy. *Nucleic Acids Res* **2021**, 49 (6), e34. DOI: 10.1093/nar/gkaa1255 From NLM.
- (10) Incarnato, D.; Neri, F.; Anselmi, F.; Oliviero, S. Genome-wide profiling of mouse RNA secondary structures reveals key features of the mammalian transcriptome. *Genome Biol* **2014**, 15 (10), 491. DOI: 10.1186/s13059-014-0491-2 From NLM.
- (11) 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.; Chang, H. Y. Structural imprints in vivo decode RNA regulatory mechanisms. *Nature* **2015**, 519 (7544), 486-490. DOI: 10.1038/nature14263 From NLM.

- (12) Strobel, E. J.; Yu, A. M.; Lucks, J. B. High-throughput determination of RNA structures. *Nat Rev Genet* **2018**, *19* (10), 615-634. DOI: 10.1038/s41576-018-0034-x From NLM.
- (13) Woodson, S. A. RNA folding and ribosome assembly. *Curr Opin Chem Biol* **2008**, *12* (6), 667-673. DOI: 10.1016/j.cbpa.2008.09.024 From NLM Medline.
- (14) Ganser, L. R.; Kelly, M. L.; Herschlag, D.; Al-Hashimi, H. M. The roles of structural dynamics in the cellular functions of RNAs. *Nat Rev Mol Cell Biol* **2019**, *20* (8), 474-489. DOI: 10.1038/s41580-019-0136-0 From NLM.
- (15) Leamy, K. A.; Assmann, S. M.; Mathews, D. H.; Bevilacqua, P. C. Bridging the gap between in vitro and in vivo RNA folding. *Q Rev Biophys* **2016**, *49*, e10. DOI: 10.1017/S003358351600007X From NLM Publisher.
- (16) Watters, K. E.; Strobel, E. J.; Yu, A. M.; Lis, J. T.; Lucks, J. B. Cotranscriptional folding of a riboswitch at nucleotide resolution. *Nat Struct Mol Biol* **2016**, *23* (12), 1124-1131. DOI: 10.1038/nsmb.3316 From NLM.
- (17) Riback, J. A.; Eeftens, J. M.; Lee, D. S. W.; Quinodoz, S. A.; Donlic, A.; Orlovsky, N.; Wiesner, L.; Beckers, L.; Becker, L. A.; Strom, A. R.; Rana, U.; Tolbert, M.; Purse, B. W.; Kleiner, R.; Kriwacki, R.; Brangwynne, C. P. Viscoelasticity and advective flow of RNA underlies nucleolar form and function. *Mol Cell* **2023**, *83* (17), 3095-3107 e3099. DOI: 10.1016/j.molcel.2023.08.006 From NLM Medline.
- (18) Warner, J. R. The assembly of ribosomes in yeast. *J Biol Chem* **1971**, *246* (2), 447-454. From NLM Medline.
- (19) Rovera, G.; Berman, S.; Baserga, R. Pulse labeling of RNA of mammalian cells. *Proc Natl Acad Sci U S A* **1970**, *65* (4), 876-883. DOI: 10.1073/pnas.65.4.876 From NLM Medline.
- (20) Klinge, S.; Woolford, J. L., Jr. Ribosome assembly coming into focus. *Nat Rev Mol Cell Biol* **2019**, *20* (2), 116-131. DOI: 10.1038/s41580-018-0078-y From NLM.
- (21) Kleiner, R. E. Interrogating the transcriptome with metabolically incorporated ribonucleosides. *Mol Omics* **2021**, *17* (6), 833-841. DOI: 10.1039/d1mo00334h From NLM Medline.
- (22) Kubota, M.; Tran, C.; Spitale, R. C. Progress and challenges for chemical probing of RNA structure inside living cells. *Nat Chem Biol* **2015**, *11* (12), 933-941. DOI: 10.1038/nchembio.1958 From NLM.
- (23) Hulscher, R. M.; Bohon, J.; Rappé, M. C.; Gupta, S.; D'Mello, R.; Sullivan, M.; Ralston, C. Y.; Chance, M. R.; Woodson, S. A. Probing the structure of ribosome assembly intermediates in vivo using DMS and hydroxyl radical footprinting. *Methods* **2016**, *103*, 49-56. DOI: 10.1016/j.ymeth.2016.03.012 From NLM.
- (24) Swiatkowska, A.; Wlotzka, W.; Tuck, A.; Barrass, J. D.; Beggs, J. D.; Tollervey, D. Kinetic analysis of pre-ribosome structure in vivo. *RNA* **2012**, *18* (12), 2187-2200. DOI: 10.1261/rna.034751.112 From NLM Medline.
- (25) Burger, K.; Mühl, B.; Kellner, M.; Rohmoser, M.; Gruber-Eber, A.; Windhager, L.; Friedel, C. C.; Dölken, L.; Eick, D. 4-thiouridine inhibits rRNA synthesis and causes a nucleolar stress response. *RNA Biol* **2013**, *10* (10), 1623-1630. DOI: 10.4161/rna.26214 From NLM.
- (26) Gherghe, C. M.; Krahn, J. M.; Weeks, K. M. Crystal structures, reactivity and inferred acylation transition states for 2'-amine substituted RNA. *J Am Chem Soc* **2005**, *127* (39), 13622-13628. DOI: 10.1021/ja053647y From NLM.
- (27) Wang, D.; Zhang, Y.; Kleiner, R. E. Cell- and Polymerase-Selective Metabolic Labeling of Cellular RNA with 2'-Azidocytidine. *J Am Chem Soc* **2020**, *142* (34), 14417-14421. DOI: 10.1021/jacs.0c04566 From NLM Medline.
- (28) Nainar, S.; Beasley, S.; Fazio, M.; Kubota, M.; Dai, N.; Correa, I. R., Jr.; Spitale, R. C. Metabolic Incorporation of Azide Functionality into Cellular RNA. *Chembiochem* **2016**, *17* (22), 2149-2152. DOI: 10.1002/cbic.201600300 From NLM Medline.
- (29) Fauster, K.; Hartl, M.; Santner, T.; Aigner, M.; Kreutz, C.; Bister, K.; Ennifar, E.; Micura, R. 2'-Azido RNA, a versatile tool for chemical biology: synthesis, X-ray structure, siRNA

applications, click labeling. *ACS Chem Biol* **2012**, 7 (3), 581-589. DOI: 10.1021/cb200510k From NLM.

(30) Agard, N. J.; Prescher, J. A.; Bertozzi, C. R. A strain-promoted [3 + 2] azide-alkyne cycloaddition for covalent modification of biomolecules in living systems. *J Am Chem Soc* **2004**, 126 (46), 15046-15047. DOI: 10.1021/ja044996f From NLM Medline.

(31) Koniev, O.; Wagner, A. Developments and recent advancements in the field of endogenous amino acid selective bond forming reactions for bioconjugation. *Chem Soc Rev* **2015**, 44 (15), 5495-5551. DOI: 10.1039/c5cs00048c From NLM.

(32) Chamberlin, S. I.; Weeks, K. M. Mapping Local Nucleotide Flexibility by Selective Acylation of 2'-Amine Substituted RNA. *Journal of the American Chemical Society* **2000**, 122 (2), 216-224. DOI: 10.1021/ja9914137.

(33) Saha, S.; Jagtap, A. P.; Sigurdsson, S. T. Site-directed spin labeling of 2'-amino groups in RNA with isoindoline nitroxides that are resistant to reduction. *Chem Commun (Camb)* **2015**, 51 (66), 13142-13145. DOI: 10.1039/c5cc05014f From NLM.

(34) Jao, C. Y.; Salic, A. Exploring RNA transcription and turnover in vivo by using click chemistry. *Proc Natl Acad Sci U S A* **2008**, 105 (41), 15779-15784. DOI: 10.1073/pnas.0808480105.

(35) Herzog, V. A.; Reichholf, B.; Neumann, T.; Rescheneder, P.; Bhat, P.; Burkard, T. R.; Wlotzka, W.; von Haeseler, A.; Zuber, J.; Ameres, S. L. Thiol-linked alkylation of RNA to assess expression dynamics. *Nat Methods* **2017**, 14 (12), 1198-1204. DOI: 10.1038/nmeth.4435 From NLM Medline.

(36) Spitale, R. C.; Flynn, R. A.; Torre, E. A.; Kool, E. T.; Chang, H. Y. RNA structural analysis by evolving SHAPE chemistry. *Wiley Interdiscip Rev RNA* **2014**, 5 (6), 867-881. DOI: 10.1002/wrna.1253 From NLM.

(37) Xiao, L.; Fang, L.; Kool, E. T. Acylation probing of "generic" RNA libraries reveals critical influence of loop constraints on reactivity. *Cell Chem Biol* **2022**, 29 (8), 1341-1352.e1348. DOI: 10.1016/j.chembiol.2022.05.005 From NLM.

(38) Flynn, R. A.; Zhang, Q. C.; Spitale, R. C.; Lee, B.; Mumbach, M. R.; Chang, H. Y. Transcriptome-wide interrogation of RNA secondary structure in living cells with icSHAPE. *Nat Protoc* **2016**, 11 (2), 273-290. DOI: 10.1038/nprot.2016.011 From NLM.

(39) Bernier, C. R.; Petrov, A. S.; Waterbury, C. C.; Jett, J.; Li, F.; Freil, L. E.; Xiong, X.; Wang, L.; Migliozi, B. L.; Herschkovits, E.; Xue, Y.; Hsiao, C.; Bowman, J. C.; Harvey, S. C.; Grover, M. A.; Wartell, Z. J.; Williams, L. D. RiboVision suite for visualization and analysis of ribosomes. *Faraday Discuss* **2014**, 169, 195-207. DOI: 10.1039/c3fd00126a From NLM.

(40) Li, P.; Shi, R.; Zhang, Q. C. icSHAPE-pipe: A comprehensive toolkit for icSHAPE data analysis and evaluation. *Methods* **2020**, 178, 96-103. DOI: 10.1016/j.ymeth.2019.09.020 From NLM.

(41) Anais, M.-E.; Colin, T.; Chase, W.; Christina, B.; Alain, L. Structural conservation of MALAT1 long non-coding RNA in cells and in evolution. *bioRxiv* **2022**, 2022.2007.2029.502018. DOI: 10.1101/2022.07.29.502018.

(42) Vanden Broeck, A.; Klinge, S. Principles of human pre-60S biogenesis. *Science* **2023**, 381 (6653), eadh3892. DOI: 10.1126/science.adh3892 From NLM.

(43) Wu, S.; Tutuncuoglu, B.; Yan, K.; Brown, H.; Zhang, Y.; Tan, D.; Gamalinda, M.; Yuan, Y.; Li, Z.; Jakovljevic, J.; Ma, C.; Lei, J.; Dong, M. Q.; Woolford, J. L., Jr.; Gao, N. Diverse roles of assembly factors revealed by structures of late nuclear pre-60S ribosomes. *Nature* **2016**, 534 (7605), 133-137. DOI: 10.1038/nature17942 From NLM.

(44) Smola, M. J.; Weeks, K. M. In-cell RNA structure probing with SHAPE-MaP. *Nat Protoc* **2018**, 13 (6), 1181-1195. DOI: 10.1038/nprot.2018.010 From NLM.

(45) Zubradt, M.; Gupta, P.; Persad, S.; Lambowitz, A. M.; Weissman, J. S.; Rouskin, S. DMS-MaPseq for genome-wide or targeted RNA structure probing in vivo. *Nat Methods* **2017**, 14 (1), 75-82. DOI: 10.1038/nmeth.4057 From NLM.

- (46) Khatter, H.; Myasnikov, A. G.; Natchiar, S. K.; Klaholz, B. P. Structure of the human 80S ribosome. *Nature* **2015**, 520 (7549), 640-645. DOI: 10.1038/nature14427 From NLM.
- (47) Ben-Shem, A.; Jenner, L.; Yusupova, G.; Yusupov, M. Crystal structure of the eukaryotic ribosome. *Science* **2010**, 330 (6008), 1203-1209. DOI: 10.1126/science.1194294 From NLM.
- (48) Szaflarski, W.; Leśniczak-Staszak, M.; Sowiński, M.; Ojha, S.; Aulas, A.; Dave, D.; Malla, S.; Anderson, P.; Ivanov, P.; Lyons, S. M. Early rRNA processing is a stress-dependent regulatory event whose inhibition maintains nucleolar integrity. *Nucleic Acids Res* **2022**, 50 (2), 1033-1051. DOI: 10.1093/nar/gkab1231 From NLM.
- (49) Su, Z.; Tang, Y.; Ritchey, L. E.; Tack, D. C.; Zhu, M.; Bevilacqua, P. C.; Assmann, S. M. Genome-wide RNA structurome reprogramming by acute heat shock globally regulates mRNA abundance. *Proc Natl Acad Sci U S A* **2018**, 115 (48), 12170-12175. DOI: 10.1073/pnas.1807988115 From NLM.
- (50) Yamagami, R.; Sieg, J. P.; Assmann, S. M.; Bevilacqua, P. C. Genome-wide analysis of the in vivo tRNA structurome reveals RNA structural and modification dynamics under heat stress. *Proc Natl Acad Sci U S A* **2022**, 119 (25), e2201237119. DOI: 10.1073/pnas.2201237119 From NLM.
- (51) Witold, S.; Mateusz, S.; Marta, L.; Sandeep, O.; Anaïs, A.; Dhvani, D.; Sulochan, M.; Paul, A.; Pavel, I.; Shawn, M. L. A novel stress response pathway regulates rRNA biogenesis. *bioRxiv* **2020**, 2020.2008.2016.250183. DOI: 10.1101/2020.08.16.250183.
- (52) Bohnsack, K. E.; Bohnsack, M. T. Uncovering the assembly pathway of human ribosomes and its emerging links to disease. *The EMBO Journal* **2019**, 38 (13), e100278. DOI: <https://doi.org/10.15252/embj.2018100278> (accessed 2023/09/26).
- (53) Raines, K.; Gottlieb, P. A. Enzymatic incorporation of 2'-thio-CTP into the HDV ribozyme. *RNA* **1998**, 4 (3), 340-345. From NLM Medline.

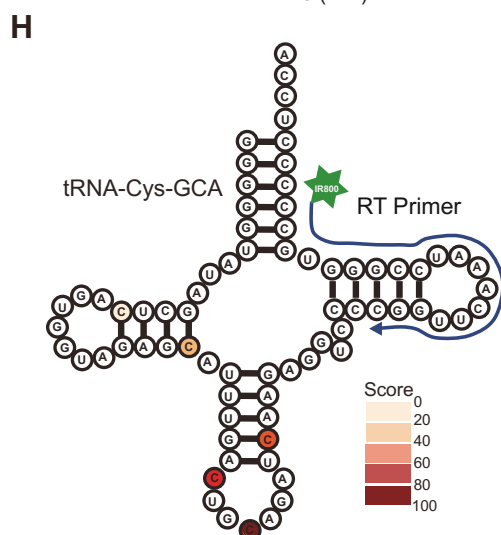
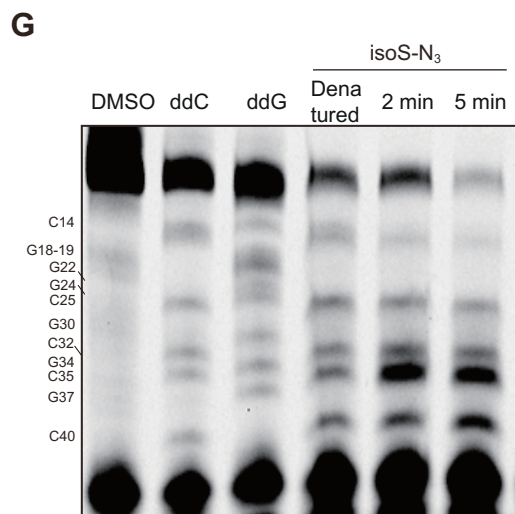
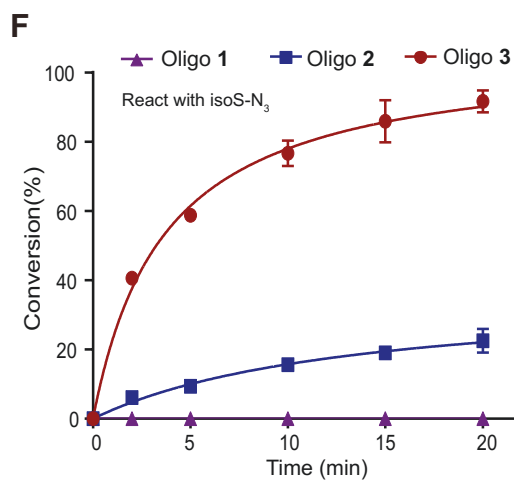
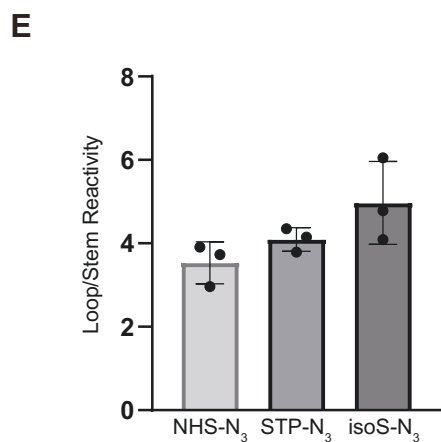
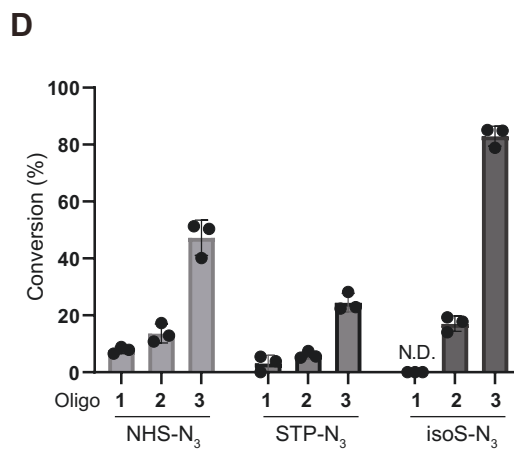
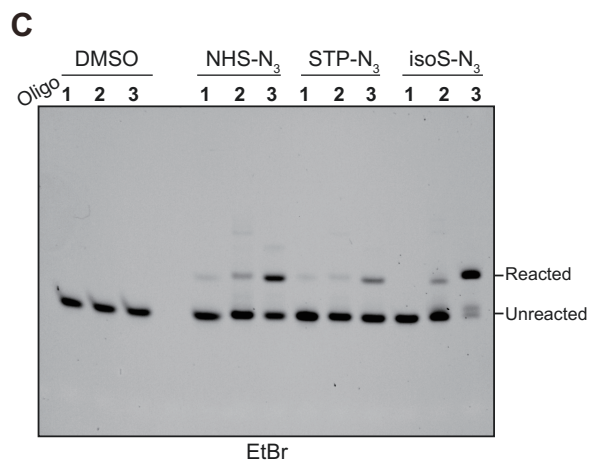
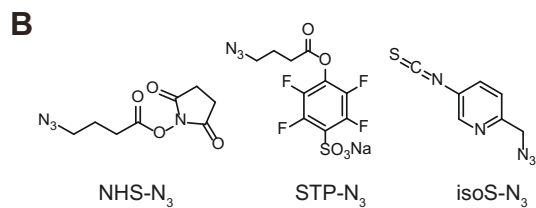
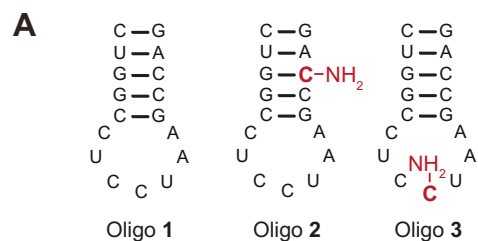


Figure 1. Structure-specific labeling of 2'-aminodeoxycytidine (2'-NH₂-dCyd) in RNA oligonucleotides. **(A)** Sequence and secondary structure of oligonucleotides used in this work. 2'-NH₂-dCyd modification is shown in red. **(B)** Chemical structure of azide-containing electrophiles used for 2'-NH₂-dCyd labeling. **(C)** Reaction of oligo 1, 2, or 3 with azide-containing electrophiles. 1 μ M oligo was reacted with 5 mM isoS-N₃, 20 mM NHS-N₃, or 20 mM STP-N₃ for 15 min and then reacted with an excess of DBCO-Biotin. Reactions were separated by denaturing polyacrylamide gel and stained with EtBr. **(D)** Gel densitometry-based quantification of labeling reactions in **(C)**. **(E)** Selectivity of each electrophilic probe for labeling structured (oligo **2**) as compared to unstructured (oligo **3**) 2'-NH₂-dCyd sites. **(F)** Time course oligo labeling with isoS-N₃. Reactions were performed as in **(C)**. Data represent mean \pm s.d. (n = 3). **(G)** RT-primer extension-based analysis of 2'-NH₂-dCyd-labeled IVT tRNA^{Cys} after isoS-N₃ reaction. Nucleotide ladders were generated using ddCTP and ddGTP in the primer extension reaction. **(H)** Gel densitometry-based quantification of reverse transcription (RT) "stops" in **(G)**. RT stop scores were normalized to the intensity of the band at C35, which exhibited the most prominent stop.

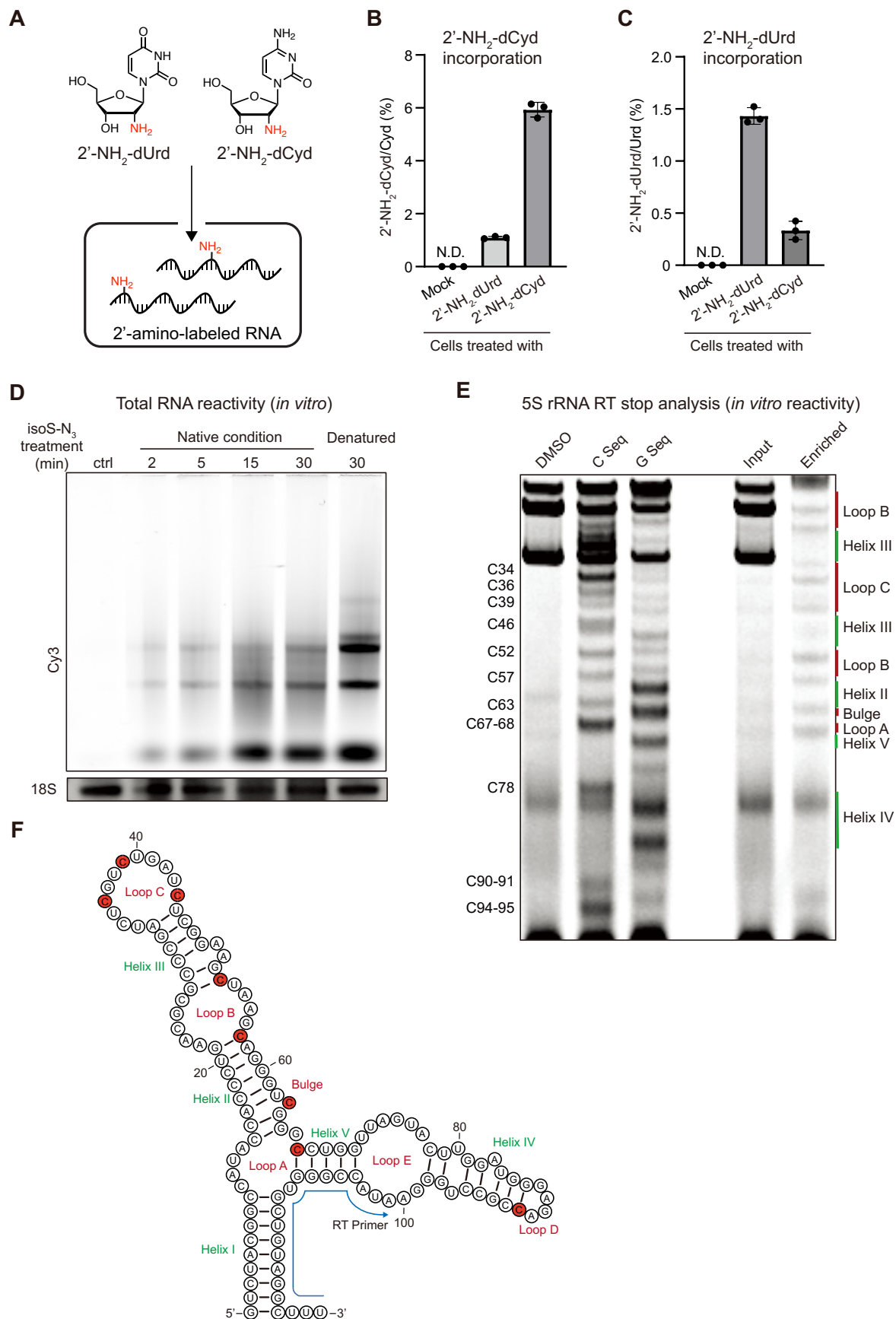


Figure 2. Metabolic RNA labeling with 2'-aminodeoxypyrimidines and isoS-N₃ structure probing. **(A)** Chemical structures of nucleosides used in this work. **(B)** and **(C)** LC-QQQ-MS analysis of total RNA isolated from HEK293T treated with 1 mM 2'-aminodeoxypyrimidine nucleosides. Data represent mean \pm s.d. (n = 3); n.d. = not detected. **(D)** Labeling of total RNA extracted from 2'-NH₂-dCyd-treated HEK293T cells. Isolated total RNA was reacted with isoS-N₃ and then treated with an excess of Cy3-DBCO. Cy3-labeled RNA was detected by in-gel fluorescence. **(E)** *In vitro* structure probing of 5S rRNA. Total 2'-NH₂-dCyd labeled cellular RNA was purified from HEK293T cells and reacted with 5 mM isoS-N₃ in folding buffer at 37 °C for 15 min, and then an excess of DBCO-Disulfide-Biotin for enrichment. Enriched isoS-N₃-labeled RNA was used as template for 5S rRNA primer extension. Nucleotide ladders were generated by adding ddCTP and ddGTP to the primer extension reactions. **(F)** Secondary structure of human 5S rRNA. Residues that react with isoS-N₃ in **(E)** are shown in red.

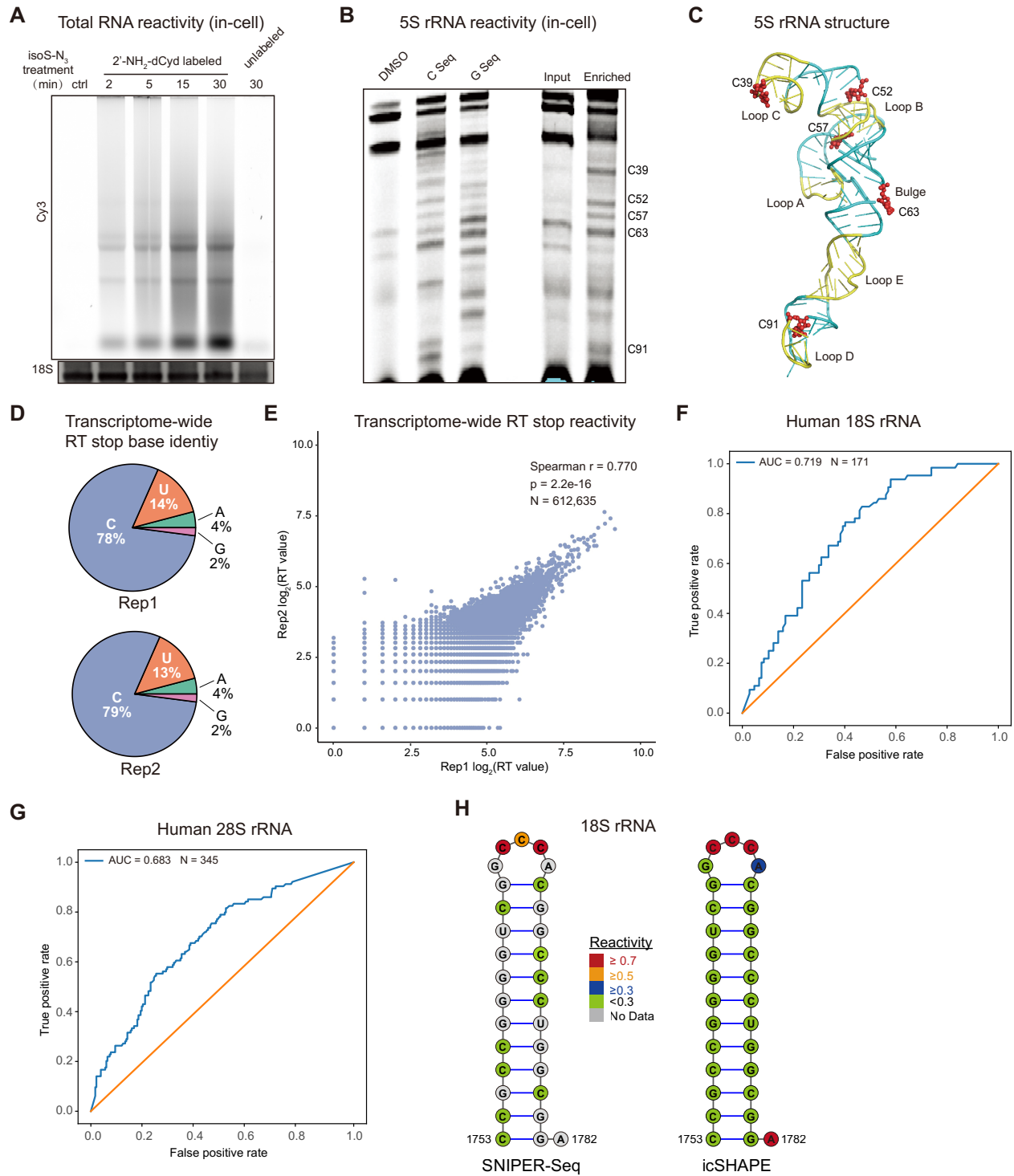


Figure 3. In-cell RNA structure probing with isoS-N₃ and 2'-NH₂-dCyd. **(A)** *In cellulo* RNA labeling with isoS-N₃. 2'-NH₂-dCyd treated HEK293T cells were labeled with 5 mM isoS-N₃ in PBS at 37 °C for the indicated time. Total cellular RNA was extracted and reacted with Cy3-DBCO for detection.

(B) *In cellulo* structure probing of 5S rRNA with isoS-N₃. 2'-NH₂-dCyd-labeled cells were treated with isoS-N₃ at 37 °C for 15 min prior to RNA isolation, enrichment, and RT-primer extension analysis with a 5S rRNA-specific primer. **(C)** Structure of human 5S rRNA (PDB ID: 4UG0). Loop regions are shown in yellow and reacted cytidine residues in **(B)** are highlighted in red. **(D)** Distribution of RT stop events in two independent SNIPER-seq samples. **(E)** Scatter plot of normalized RT stop counts across the transcriptome for two independent SNIPER-seq samples; ρ is Spearman correlation coefficient; total number of bases = 612,635. **(F)** and **(G)** Receiver Operating Characteristic (ROC) curve plots for SNIPER-seq reactivity scores as compared against reference 18S rRNA structure **(F)** and 28S rRNA structure **(G)** from the Comparative RNA Web (CRW) site. AUC scores for all ROC curves are included. The analysis includes C residues with solvent accessibility greater than 3. **(H)** Human 18S rRNA (nucleotides 1753–1782) secondary structure with both SNIPER-seq and icSHAPE reactivity data shown in color.

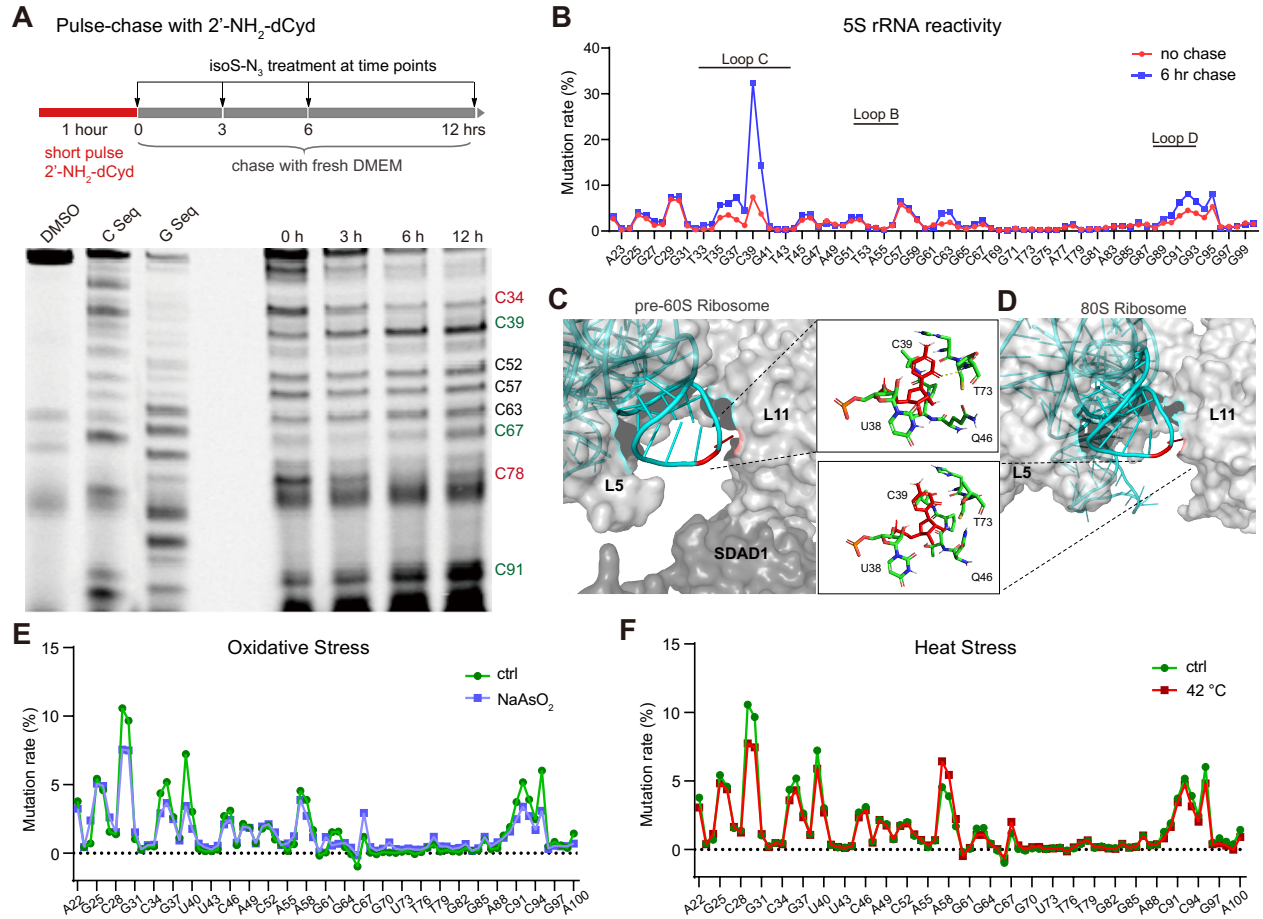


Figure 4. In-cell nascent RNA structure probing with SNIPER-seq. **(A)** Probing time-dependent 5S rRNA structural dynamics using pulse-chase feeding with 2'-NH₂-dCyd and isoS-N₃ labeling. Enriched isoS-N₃ labeled RNA was used for 5S rRNA primer extension reaction. **(B)** Mutational profiling to measure 5S rRNA structural dynamics. Cells were pulsed for 1 hr with 2'-NH₂-dCyd and processed immediately or chased for 6 hr. **(C)** and **(D)** Structure of 5S rRNA loop C (cyan) in complex with ribosomal proteins (gray) from human pre-60S ribosome (PDB ID: 3JCT) **(C)** or mature 80S ribosome (PDB ID:4V7R) **(D)**. C39 residue is shown in red and interactions with proximal protein residues are shown in the inset. **(E)** and **(F)** Structural dynamics of nascent 5S rRNA during oxidative stress **(E)** or heat stress **(F)**. For oxidative stress, cells were cotreated with 0.5 mM NaAsO₂ and 1 mM 2'-NH₂-dCyd for 1 hr at 37 °C. For heat stress, cells were incubated with 1 mM 2'-NH₂-dCyd at 42 °C for 1 hr.

