Supplementary Information

Phosphorothioate-based site-specific labeling of large RNAs for structural and dynamic studies

Yanping Hu[1, ‡], Yan Wang[1, ‡], Jaideep Singh[2, ‡], Ruirui Sun[1, ‡], Lilei Xu[1], Xiaolin Niu[1], Keyun Huang[1], Jieqiong Chen[1], Guangcan Bai[3], Guoquan Liu[3], Xiaobing Zuo[4], Chunlai Chen[1], Peter Z Qin[2], Xianyang Fang[1, *]

[1]Beijing Advanced Innovation Center for Structural Biology, Beijing Frontier Research Center for Biological Structure, School of Life Sciences, Tsinghua University, Beijing 100084, China. [2]Department of Chemistry, University of Southern California, Los Angeles, California 90089, United States. [3]State Key Laboratory of Natural and Biomimetic Drugs, School of Pharmaceutical Sciences, Peking University, Beijing 100191, China. [4]X-ray Science Division, Argonne National Laboratory, Lemont IL 60439, USA.

[*]To whom correspondence should be addressed. Tel: +86 10 62771071; Email: fangxy@mail.tsinghua.edu.cn

[‡]The authors wish it to be known that, in their opinion, the first four authors should be regarded as Joint First Authors.

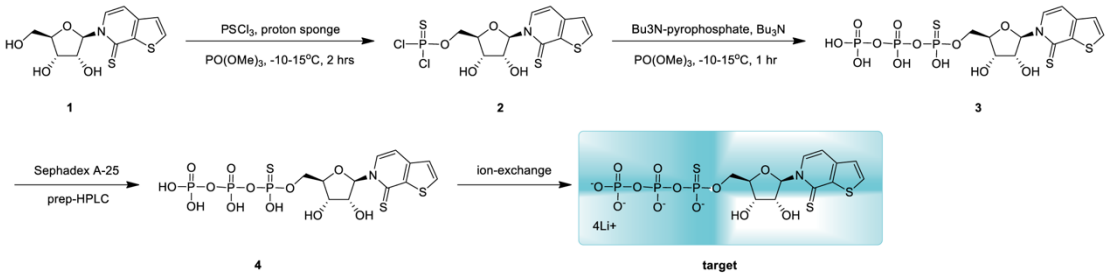
Table of Content

Synthetic procedures and characterizations of rTPT3$_{\alpha S}$-TP

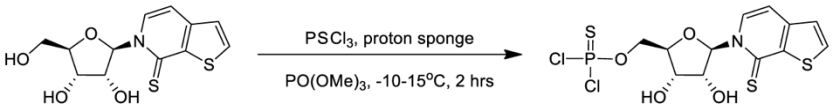
General

All solvents and reagents were purchased commercially and used without further purification. For synthetic procedures all reactions were carried out in oven-dried glassware under an inert atmosphere. Solvents were distilled and/or dried over 4 Å molecular sieves. NMR spectra were recorded on an AVANCE III 1 BAY 400 MHz Bruker NMR spectrometer and the chemical shifts were reported relative to the deuterated NMR solvent used [^{1}H-NMR: CDCl3 (7.26 ppm), DMSOd6 (2.50 ppm)]. Mass spectra were recorded on an Agilent 1200+G6110A.

Synthetic schemes and procedures

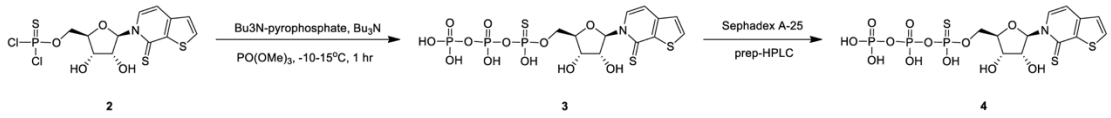


General procedure for preparation of compound 2



Compound 1 (3.00 g, 10.0 mmol, 1.00 *eq*) and proton sponge (2.15 g, 10.0 mmol, 1.00 *eq*) was dissolved in dry ACN. The mixture was concentrated under reduced pressure to remove water three times. To a solution of compound 1 (3.00 g, 10.0 mmol, 1.00 *eq*) and proton sponge (2.15 g, 10.0 mmol, 1.00 *eq*) in $PO(OMe)_3$ (18.0 mL) was added $PSCl_3$ (2.60 g, 15.4 mmol, 1.53 *eq*) at -10°C. The mixture was stirred at 15°C for 2hrs. LCMS (RT=1.26 min) showed compound 1 was consumed completely and one main peak with desired mass was detected. The reaction solution was used into the next step without further work up and purification. Compound 2 (8.66 g, crude) in $PO(OMe)_3$ (18.0 mL) was obtained as a yellow liquid.

General procedure for preparation of compound 4

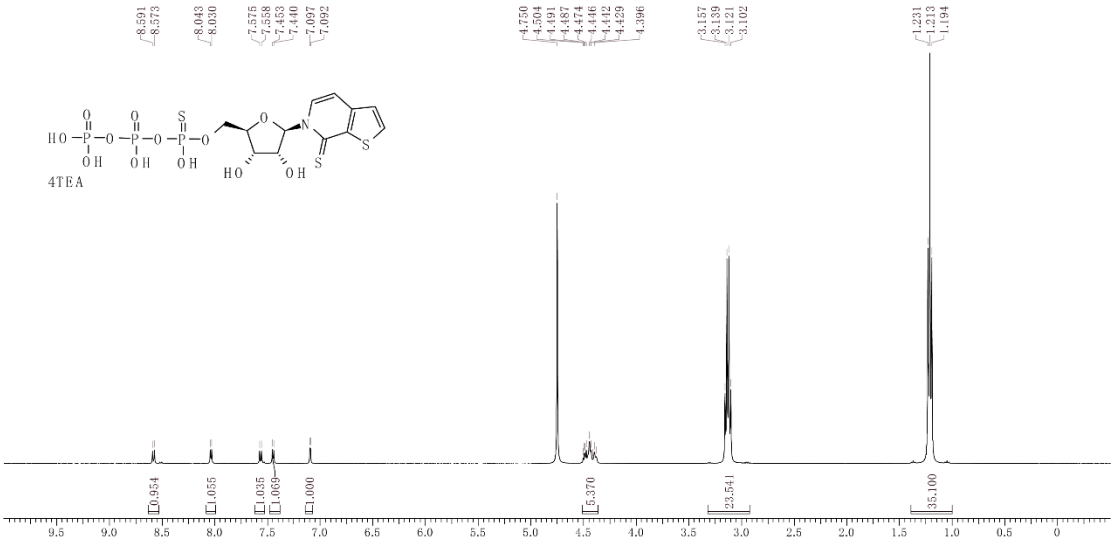


2 3 4

To a solution of compound 2 (4.33 g, 10.0 mmol, 1.00 eq) in $PO(OMe)_3$ (18.0 mL) was added Bu3N (11.2 g, 60.6 mmol, 14.4 mL, 6.05 eq) and Bu_3N-pyrophosphate (0.60 M, 83.5 mL, 5.00 eq) at -10°C. The mixture was stirred at 15°C for 1 hr. LCMS (ET42571-7-P1A1, RT = 1.08 min) showed compound 2 was consumed completely and one peak with desired mass was detected. The reaction solution was added TEAB to pH = 7. The reaction solution was diluted with H_2O (200 mL) and extracted with MTBE (400 mL x 3). The residue was purified by column chromatography (DEAE Sephadex A-25, TEAB=0.80 N~1.00 N). LCMS (RT=0.592 min) one peak with desired mass was detected. The residue was purified by prep-HPLC (neutral condition) column: Agela DuraShell C18 250 x 70mm x 10um; mobile phase: [TEAB (10mM)-ACN]; B%: 1%-12%, 20min. The mixture was lyophilized. Compound 4 (0.90 g, 1.62 mmol, 8.18% yield) was obtained as a yellow solid.

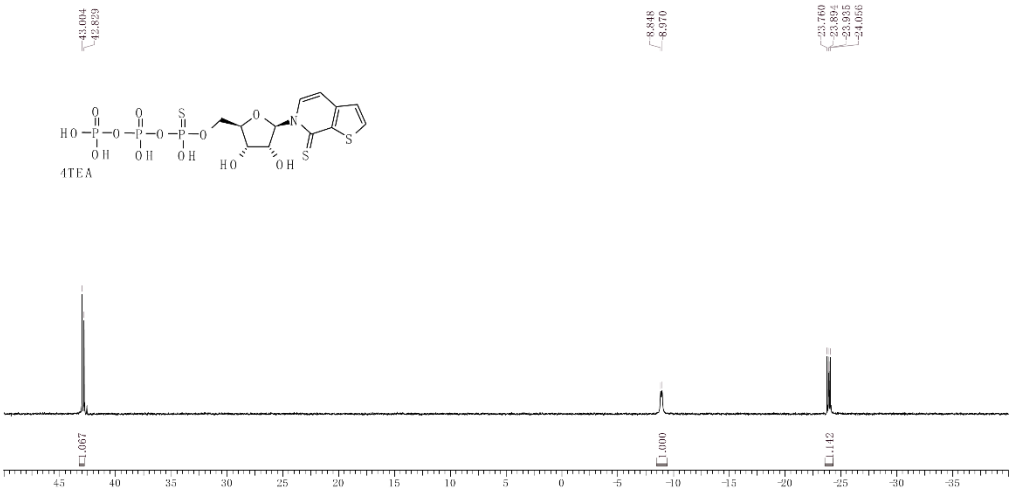
^{1}H NMR: (400 MHz D_2O)

δ ppm 8.58 (d, *J*=7.2 Hz, 1H) 8.04 (d, *J*=5.2 Hz, 1H) 7.57 (d, *J*=6.8 Hz, 1H) 7.45 (d, *J*=5.2 Hz, 1H) 7.09 (d, *J*=2.0 Hz, 1H) 4.38-4.50 (m, 5H)



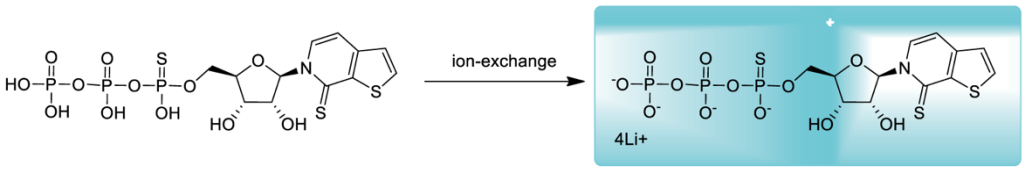
^{31}P NMR: (162 MHz D_2O)

δ ppm 42.9 -8.91, -23.9



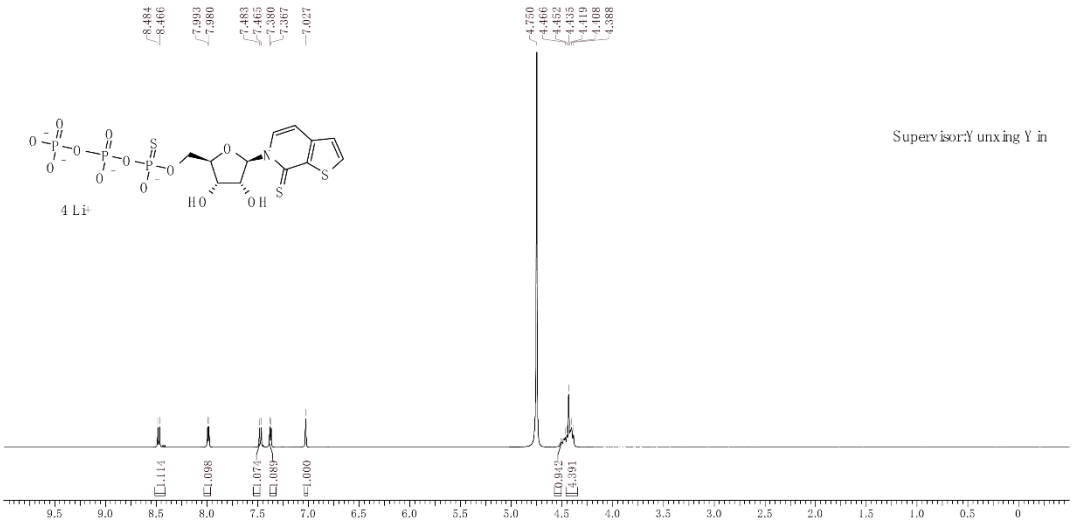
General procedure for preparation of target

Compound 4 (1.50 g, 2.70 mmol, 1.00 *eq*) was dissolved in de-ionized H2O (5.00 mL) passed through the resin (Li+, Dowex 50WX8 100mesh). The target compound was showed spot at TLC, when the spot disappears, stop added water. The water was collected and lyophilized. Target (0.81 g, 1.46 mmol, 54.0% yield) was obtained as a yellow solid.



4 **target**

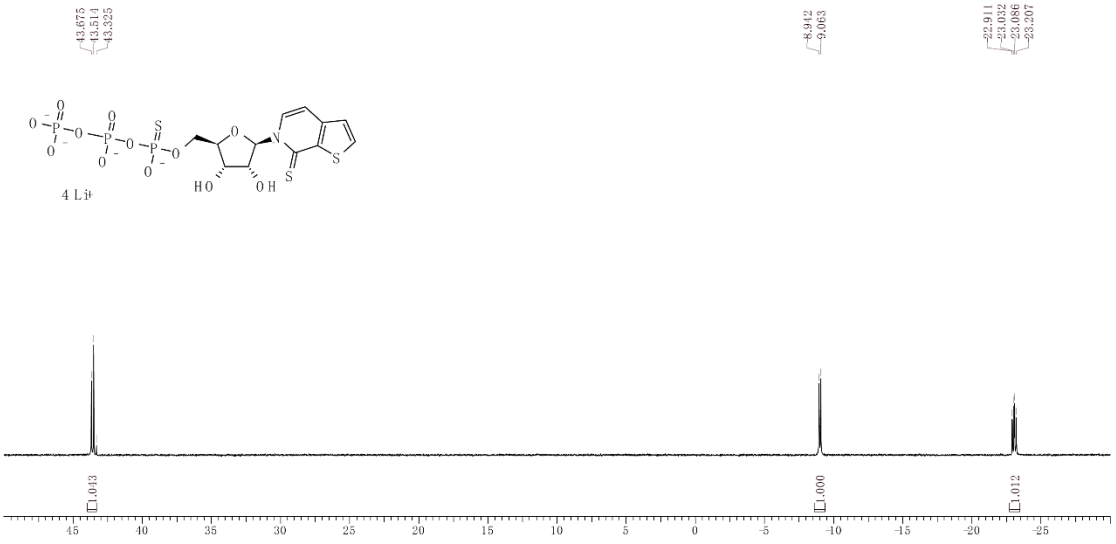
^{1}H NMR: (400 MHz D_2O)



δ ppm 8.47 (d, *J*=7.2 Hz, 1H) 7.99 (d, *J*=5.2 Hz, 1H) 7.47 (d, *J*=6.8 Hz, 1H) 7.37 (d, *J*=5.2 Hz, 1H) 7.03 (s, 1H) 4.39-4.47 (m, 5H)

^{31}P NMR: (162 MHz D_2O)

δ ppm 43.5 -9.00, -23.1



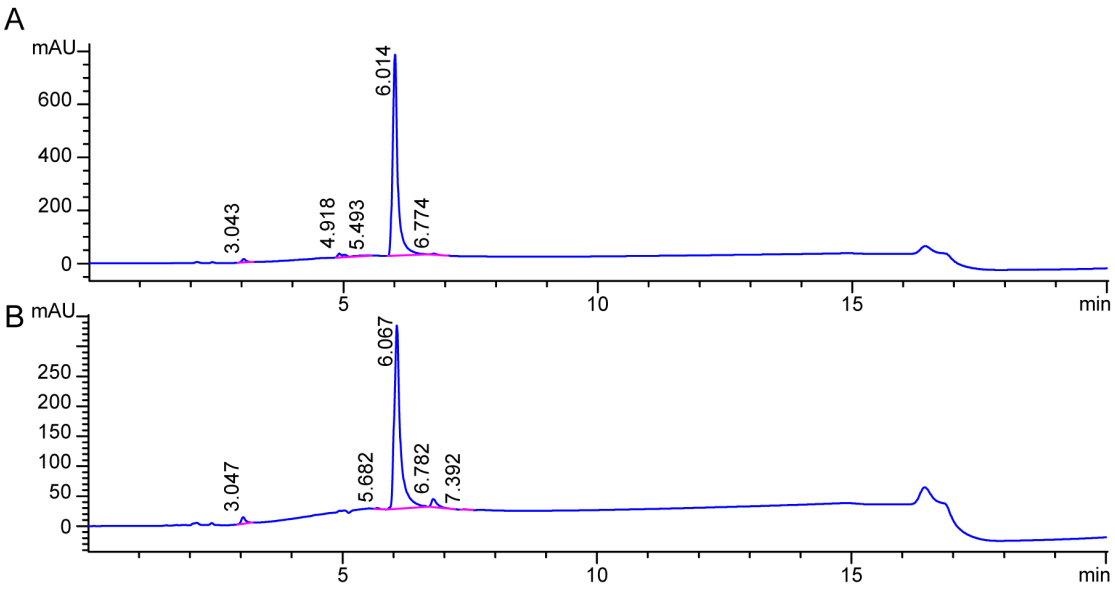


Figure S1. High resolution HPLC profiles for (**A**) R_p and (**B**) S_p diastereomers of chemically synthesized rTPT3$_{\alpha S}$TP.

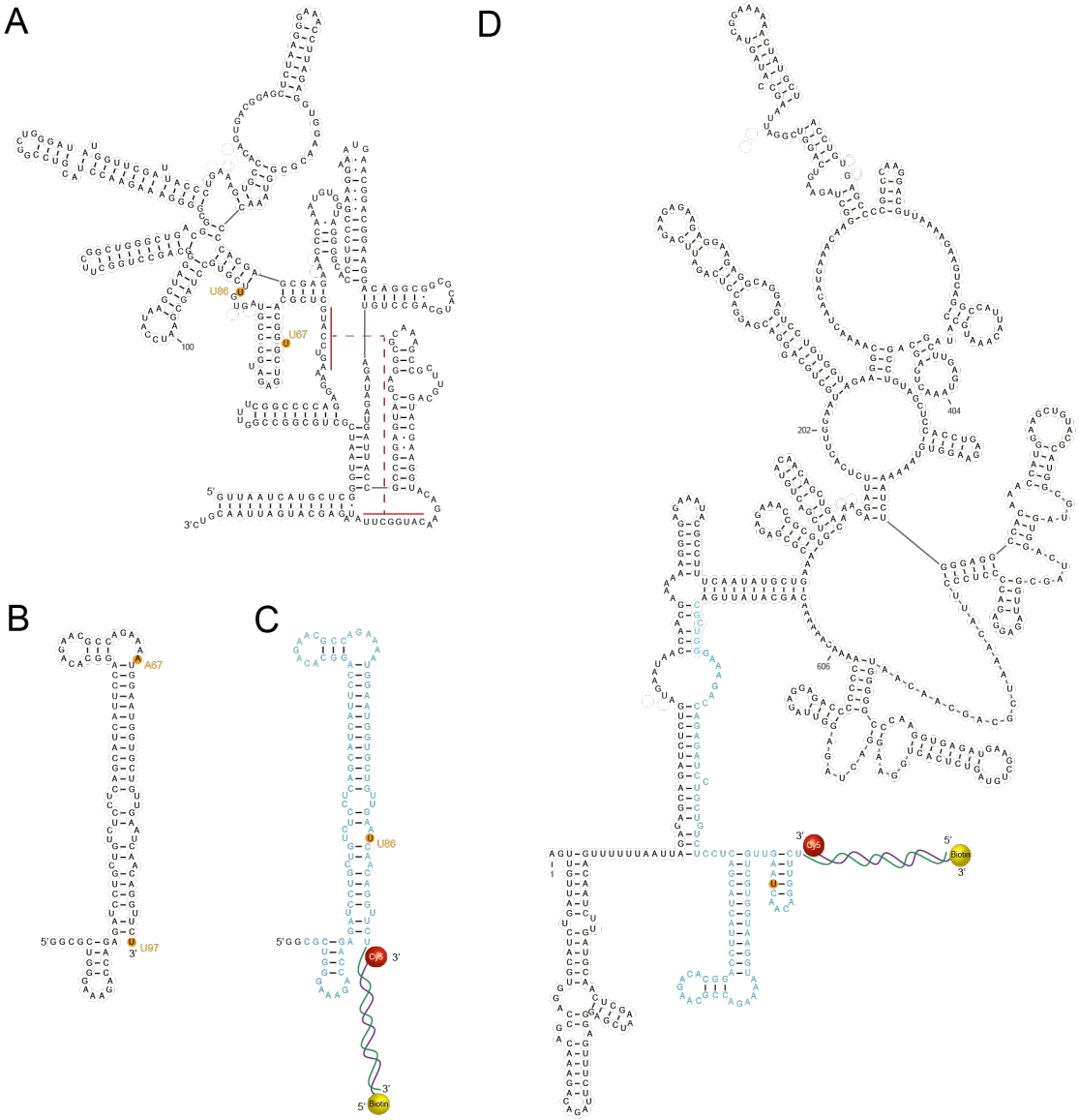


Figure S2. Secondary structures of RNAs used in this study. (**A**) Secondary structure of RNase P from *Bacillus stearothermophilus*. The labeling sites (U67, U86) are indicated with orange background. (**B**) Secondary structure of DENV 3'SL. The labeling sites (A67, U97) are indicated with orange background. (**C**) DENV 3'SL_{30} and (**D**) DENV-$mini_{30}$ used for smFRET experiments are fused with a 30-nts tail at their 3' ends (green). A complementary 30-nts single-stranded DNA with a Cy5 and a biotin labeled at its 3' and 5' ends, respectively, is used to introduce Cy5 and biotin labels by annealing. The sCy3 labeling sites on 3'SL_{30} or DENV-$mini_{30}$ are indicated with orange background. The sequence corresponding to 3'SL is highlighted in blue.

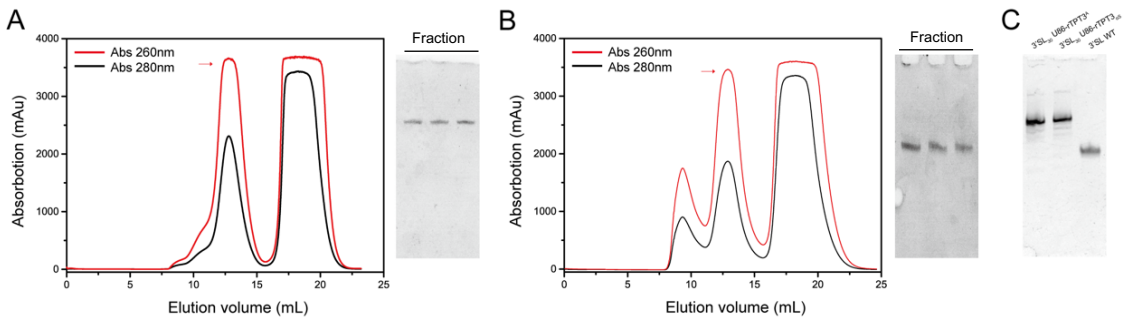


Figure S3. Purification of rTPT3$_{\alpha S}$- or rTPT3^{A}-modified RNAs. (**A-B**) Native purification of RNA transcripts of the doubly rTPT3$_{\alpha S}$-modified RNase P RNA with Superdex 200 column (**A**) and the doubly rTPT3$_{\alpha S}$-modified 3'SL RNA with Superdex 75 column (**B**) by Size exclusion chromatography. Peaks containing target RNA are indicated with red arrows. The native PAGE of fractions of purified RNAs are shown on the right. (**C**) Native PAGE of purified rTPT3$_{\alpha S}$- or rTPT3^{A}-modified 3'SL$_{30}$ at the site of U86. The wild type 3'SL RNA is included (lane 3) as reference.

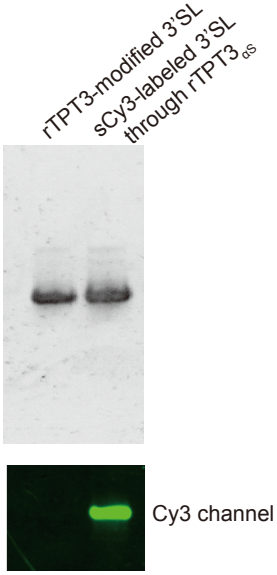


Figure S4. Reactivity of rTPT3-modified and rTPT3$_{\alpha S}$-modified RNAs with monomaleimido sCy3. PAGE analysis of rTPT3-modified and sCy3-labeled 3'SLs.

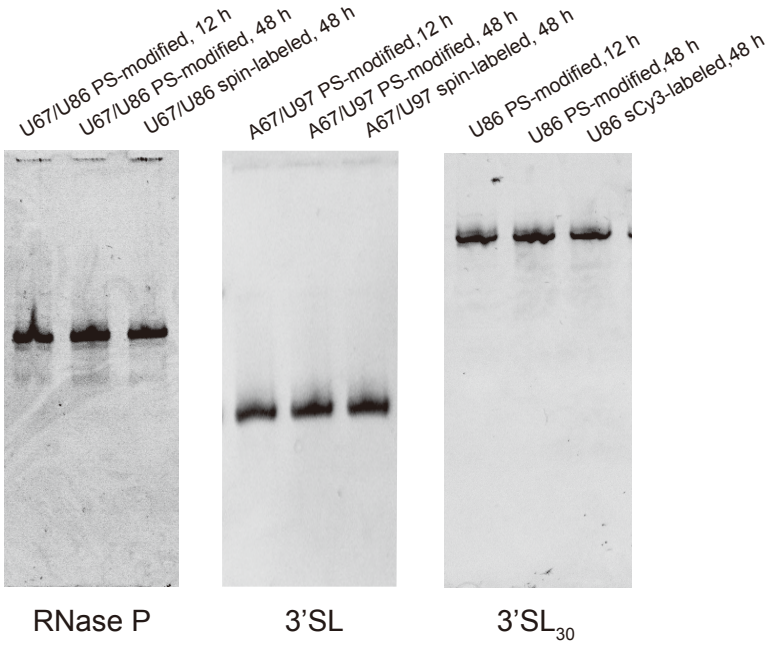


Figure S5. Stability test of rTPT3$_{\alpha S}$-modified and probe-labeled RNAs after storage at 4°C for 12 and 48 hours. Native PAGE of RNase P U67/U86 (rTPT3$_{\alpha S}$-modified and spin-labeled), 3'SL A67/U97 (rTPT3$_{\alpha S}$-modified and spin-labeled) and 3'SL$_{30}$ (rTPT3$_{\alpha S}$-modified and sCy3-labeled) are shown on left, middle and right panels, respectively.

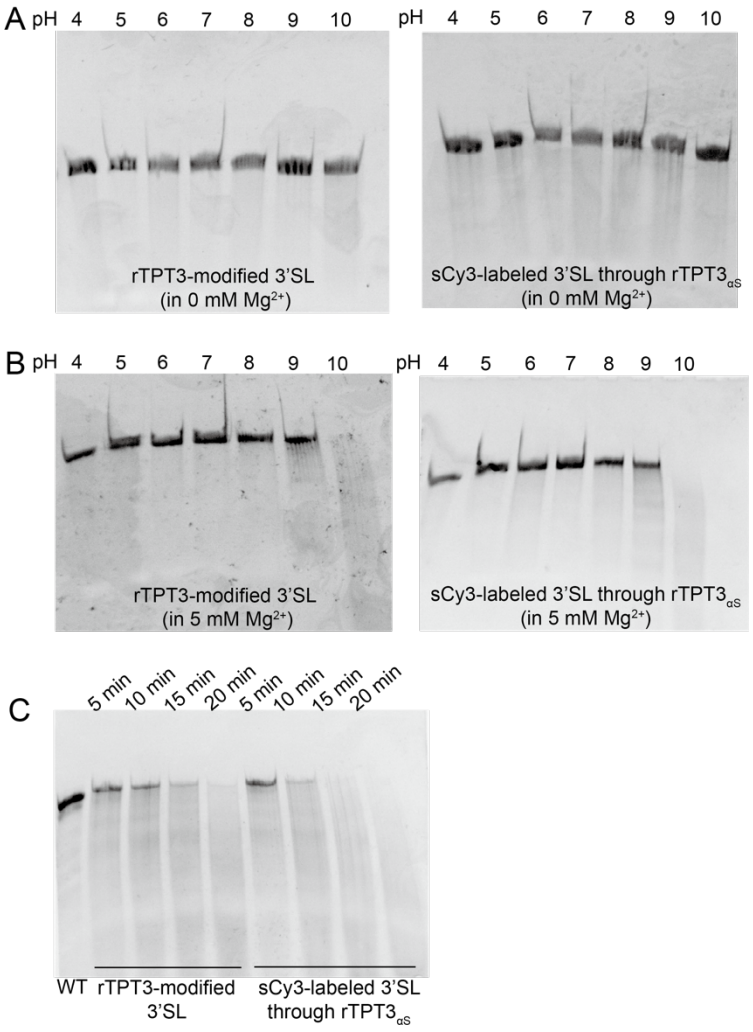


Figure S6. Hydrolytic stability of rTPT3-modified 3'SL at the site of A67 and sCy3-labeled 3'SL through rTPT3$_{\alpha S}$ at the site of A67 under different buffer conditions. (**A-B**) Denatured PAGE analysis of rTPT3-modified and sCy3-labeled 3'SLs at different pHs in the absence (**A**) and presence of 5 mM Mg^{2+} after heated at 65°C for 10 mins (**B**). (**C**) Denatured PAGE analysis of rTPT3-modified and sCy3-labeled 3'SL in the buffer containing 5 mM Mg^{2+} and at pH 9. Samples were collected at 5, 10, 15 and 20 minutes after heated at 65 °C.

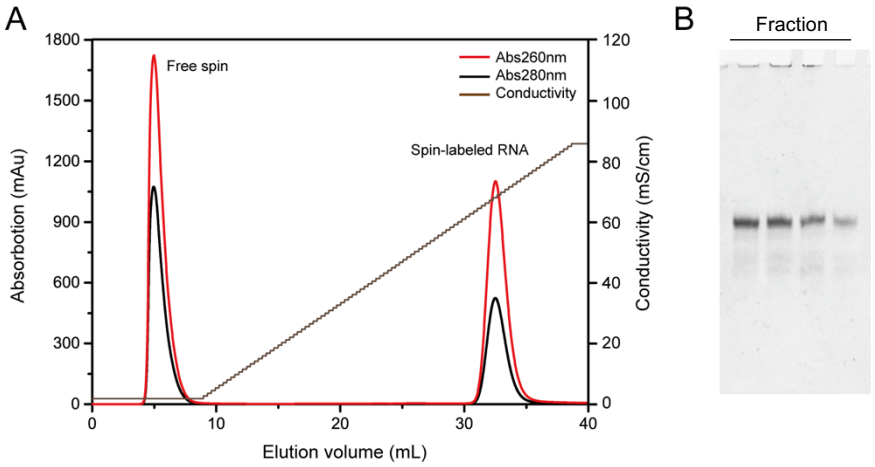


Figure S7. Purification of spin-labeled RNA samples under nondenaturing conditions. (**A**) The elution profiles of doubly Nanogold-labeled RNase P (U67/U86) by anion exchange chromatography. Peaks corresponding to free spin, the spin-labeled RNAs and the conductivity curve are indicated. (**B**) Native PAGE analysis of fractions of purified double spin labeled RNase P (U67/U86) from anion exchange chromatography.

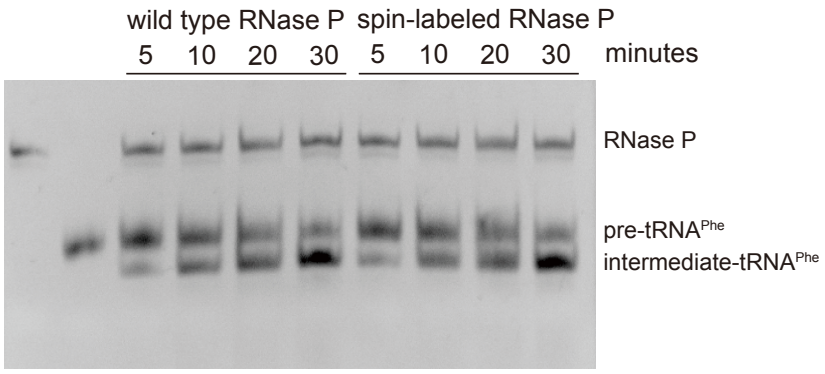


Figure S8. The kinetics of pre-tRNA processing by the wild type and spin-labeled RNase P RNAs. Samples were collected at 5, 10, 20 and 30 minutes after starting the reactions and denatured immediately for analysis.

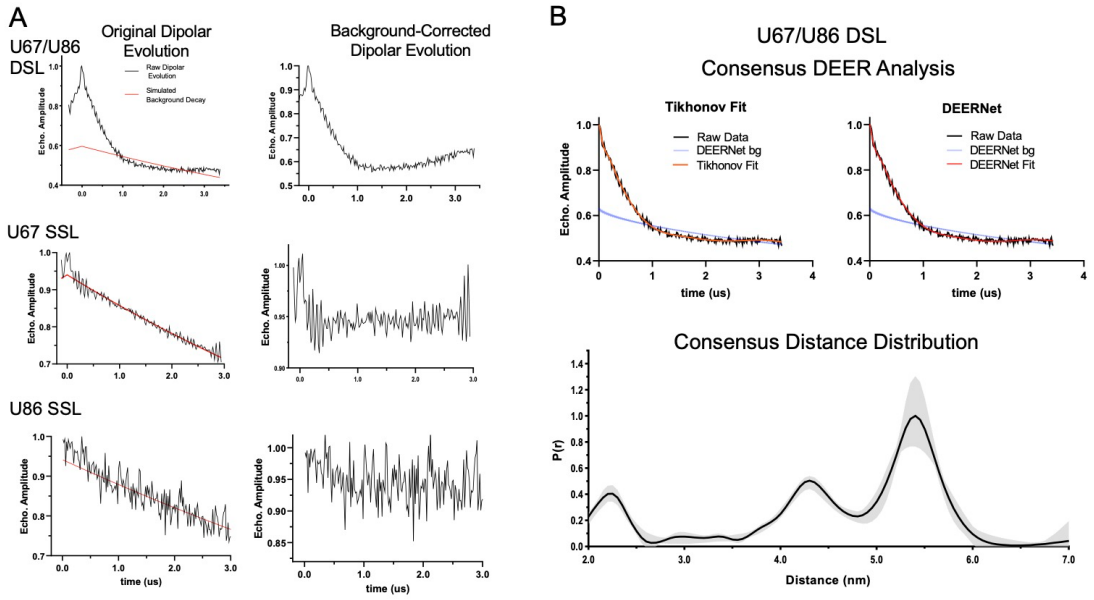


Figure S9. Additional PELDOR data and analysis. **(A)** Original PELDOR data. Shown on the right are raw dipolar evolution traces overlaid with the corresponding simulated background generated by LongDistances (1). Shown on the left are the background-corrected dipolar evolution traces. For the single-labeled U67 SSL and U866 SSL samples, no decay was observed in the background corrected traces, indicating a lack of inter-molecular dipolar coupling between the spin labels. **(B)** Analysis of the double-labeled U67/U86 DSL data by Consensus DEER Analyzer (CDA). CDA (2,3) carries out automatic data analysis using two fitting approaches, and was used here following a recently published community guideline (4) for reporting PELDOR data. Shown on the top are the CDA generated fitting using either Tikhonov regularization (left) or neural network (DEERNet) (right). Shown on the bottom is the CDA generated consensus distance distribution, with the shading indicating the uncertainty range. The CDA determined optimal regularization parameter (α) is 3.12, and the resulting distance distribution is highly similar to that obtained using LongDistances (Main Text Fig. 4D). The CDA analysis indicated that the measured data has a signal-to-noise ratio of 41.5 and a modulation-depth of 37%. These results support confidence on the reported distance distribution.

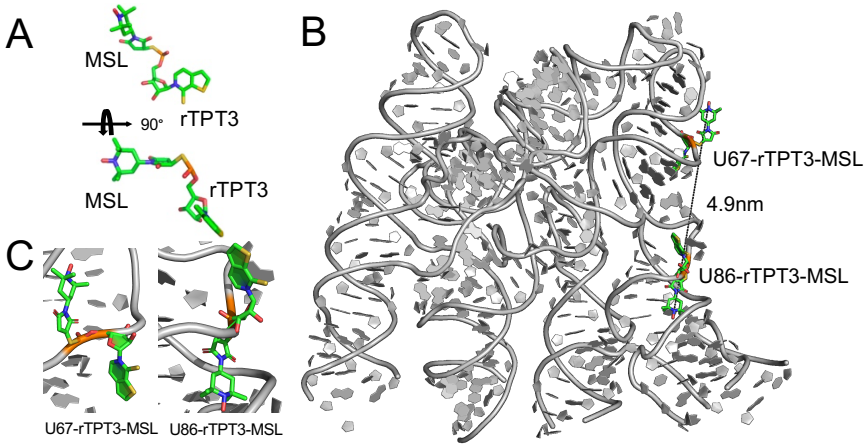


Figure S10. Inter-spin distance modeling for U67/U86 labeled RNase P. A pdb-format MSL label was generated with Chem3D, and modeled onto the RNase P structure (PDB ID 2A64) using ALLNOX (5). (**A**) The chemical structure of MSL-labeled R_p-rTPT3$_{\alpha S}$. (**B**) Representative snapshot of the double spin labeled RNase P. (**C**) The enlarged views of U67-rTPT3-MSL and U86-rTPT3-MSL in (**B**).

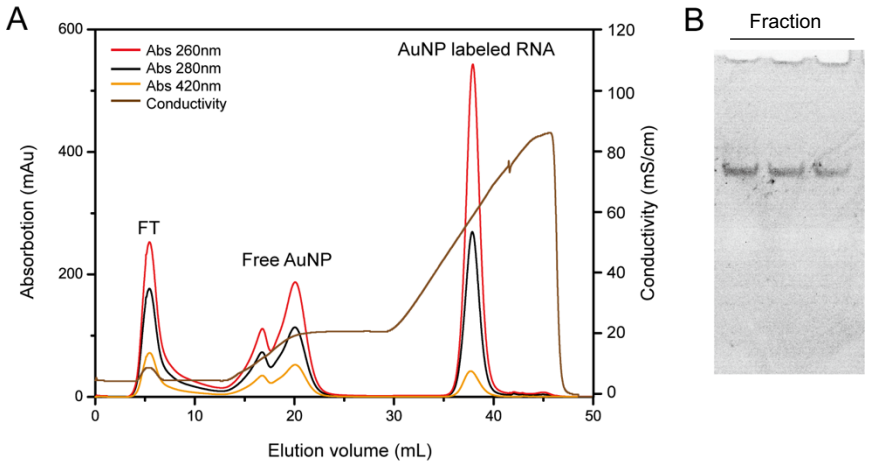


Figure S11. Purification of Nanogold-labeled RNA samples under nondenaturing conditions. (**A**) The elution profiles of doubly Nanogold-labeled 3′SL RNA by anion exchange chromatography. Peaks corresponding to the flow-through (FT), free Nanogold, the Nanogold-RNA conjugate and the conductivity curve are indicated. (**B**) Native PAGE analysis of fractions of purified doubly Nanogold-labeled RNA from anion exchange chromatography.

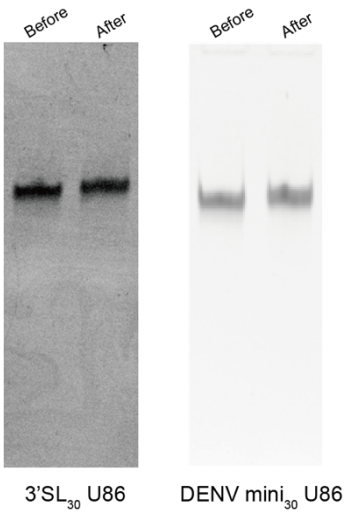


Figure S12. The native PAGE of fluorophore-labeled 3'SL_{30} and DENV-$mini_{30}$ RNAs (both at site of U86) before and after sCy3 labeling.

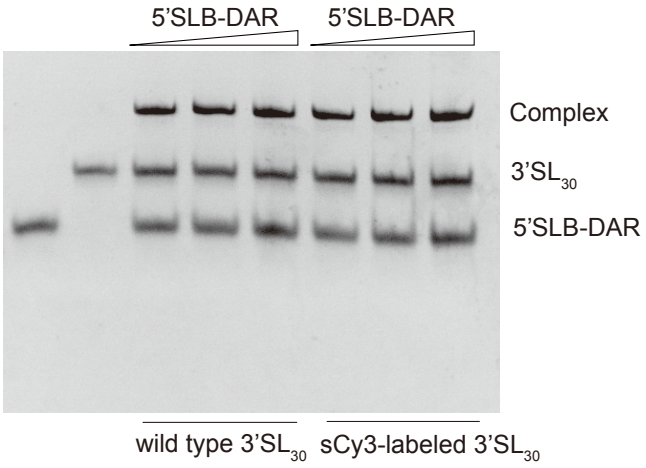


Figure S13. The RNA binding assay of wild type 3'SL_{30} or sCy3-labeled 3'SL_{30} with 5'SLB-DAR. The concentrations of wild type 3'SL_{30} or sCy3-labeled 3'SL_{30} are kept constant. The 5'SLB-DAR:3'SL_{30} molar ratio is 0.5:1, 1:1, 1.5:1, respectively.

Table S1. Overall structural parameters for 3'SL and RNase P by SAXS.

Sample	I_0[a]	Rg[a] (Å)	I_0[b]	R_g[b] (Å)	D_{max} (Å)	MW[c] (kDa)	MW[d] (kDa)
3'SL WT	0.13 ± 0.02	36.2 ± 0.4	0.13 ± 0.02	37.5 ± 0.4	132	32.10	32.30
3'SL U67/86	0.15 ± 0.02	37.0 ± 0.6	0.15 ± 0.02	37.8 ± 0.7	133		33.00
3'SL U67 SL	0.10 ± 0.01	37.3 ± 0.6	0.10 ± 0.01	38.4 ± 0.5	138	n/a	n/a
3'SL U97 SL	0.10 ± 0.01	39.1 ± 0.7	0.10 ± 0.01	40.1 ± 0.6	145		
3'SL U67/U97 DL	0.12 ± 0.02	39.4 ± 0.7	0.12 ± 0.02	40.6 ± 0.7	145		
Free AuNP	0.06 ± 0.01	8.6 ± 0.1	0.06 ± 0.01	8.7 ± 0.1	19		
RNase P WT	1.20 ± 0.08	50.8 ± 0.8	1.23 ± 0.08	52.1 ± 0.8	183	137.6	135.8
RNase P U67/86	1.22 ± 0.15	51.3 ± 0.6	1.27 ± 0.18	52.5 ± 0.7	185	137.6	138.7

[a] Derived from Guinier fitting; [b] derived from GNOM analysis; [c] MW: molecular weight predicted from sequences; [d] MW: molecular weight calculated based on the power law of volume of correlation

Table S2. Data for determination of RNA labeling efficiencies.

Spin-labeling efficiency

Sample	I_{2nd}(sample)	[NOX] (μM)	[RNA] (μM)	%Label
U67 single-labeled	38561210	95	195	47%
U86 single-labeled	38020502	89	190	48%
Double spin-labeled	47843364	198	197	50%

Nanogold-labeling efficiency

Sample	A_{260}	A_{420}	[Gold]/[RNA]	%Label
A67 single-Nanogold	9.105	0.891	54%	54%
U97 single-Nanogold	9.039	0.915	60%	60%
double- Nanogold	11.580	1.022	122%	61%

sCy3-labeling efficiency

Sample	A_{260}	A_{549}	$[A_{549}]/[A_{260}]$	%Label
3'SL_{30}	6.561	0.370	34%	34%
DENV-$mini_{30}$ RNA	12.692	0.108	40%	40%

References

1. Wingler, L.M., Elgeti, M., Hilger, D., Latorraca, N.R., Lerch, M.T., Staus, D.P., Dror, R.O., Kobilka, B.K., Hubbell, W.L. and Lefkowitz, R.J. (2019) Angiotensin Analogs with Divergent Bias Stabilize Distinct Receptor Conformations. *Cell*, **176**, 468-478 e411.
2. Worswick, S.G., Spencer, J.A., Jeschke, G. and Kuprov, I. (2018) Deep neural network processing of DEER data. *Sci Adv*, **4**, eaat5218.
3. Fabregas Ibanez, L., Jeschke, G. and Stoll, S. (2020) DeerLab: a comprehensive software package for analyzing dipolar electron paramagnetic resonance spectroscopy data. *Magn Reson (Gott)*, **1**, 209-224.
4. Schiemann, O., Heubach, C.A., Abdullin, D., Ackermann, K., Azarkh, M., Bagryanskaya, E.G., Drescher, M., Endeward, B., Freed, J.H., Galazzo, L. *et al.* (2021) Benchmark Test and Guidelines for DEER/PELDOR Experiments on Nitroxide-Labeled Biomolecules. *J Am Chem Soc*, **143**, 17875-17890.
5. Beasley, K.N., Sutch, B.T., Hatmal, M.M., Langen, R., Qin, P.Z. and Haworth, I.S. (2015) Computer Modeling of Spin Labels: NASNOX, PRONOX, and ALLNOX. *Methods Enzymol*, **563**, 569-593.