Synthesis and applications of RNAs with position-selective labelling and mosaic composition

Yu Liu1, Erik Holmstrom2, Jinwei Zhang3, Ping Yu4, Jinbu Wang1, Marzena A. Dyba4, De Chen5, Jinfa Ying6, Stephen Lockett5, David J. Nesbitt4, Adrian R. Ferré-D’Amare3, Rui Sousa7, Jason R. Stagno1 & Yun-Xing Wang1

Knowledge of the structure and dynamics of RNA molecules is critical to understanding their many biological functions. Furthermore, synthetic RNAs have applications as therapeutics and molecular sensors. Both research and technological applications of RNA would be dramatically enhanced by methods that enable incorporation of modified or labelled nucleotides into specifically designated positions or regions of RNA. However, the synthesis of tens of milligrams of such RNAs using existing methods has been impossible. Here we develop a hybrid solid–liquid phase transcription method and automated robotic platform for the synthesis of RNAs with position-selective labelling. We demonstrate its use by successfully preparing various iso-dose- or fluoroscentially labelled versions of the 71-nucleotide aptamer domain of an adenine riboswitch for nuclear magnetic resonance spectroscopy or single-molecule Förster resonance energy transfer, respectively. Those RNAs include molecules that were selectively isotope-labelled in specific loops, linkers, a helix, several discrete positions, or a single internal position, as well as RNA molecules that were fluoroscentially labelled in and near kissing loops. These selectively labelled RNAs have the same fold as those transcribed using conventional methods, but they greatly simplify the interpretation of NMR spectra. The single-position isotope- or fluoroscentially labelled RNA samples reveal multiple conformational states of the adenine riboswitch. Lastly, we describe a robotic platform and the operation that automates this technology. Our selective labelling method may be useful for studying RNA structure and dynamics and for making RNA sensors for a variety of applications including cell-biological studies, substance detection, and disease diagnostics.

The ability to synthesize RNAs with selectively labelled regions or positions is important for the following reasons. First, NMR is a powerful technique for probing RNA structure and dynamics, but it is limited by severe signal overlap in the spectra of RNA. This could be overcome by region- or position-specific iso-dose labelling of RNA so that spectral signals from critical residues could be observed without interference from the rest of the molecule. Second, Förster resonance energy transfer (FRET) experiments measure distances between fluorophores introduced site-specifically into macromolecules and have been used to study RNA structure and dynamics. Optimal placement of FRET pair fluorophores usually requires incorporation at RNA chain-internal positions and is different with current methods of synthesis, particularly with large RNAs. Third, position-specific incorporation of nucleotides derivatized with heavy or anomalously scattering atoms into RNA may aid crystallographic phase determination. Fourth, RNAs synthesized with modified nucleotides to enhance functionality or RNase resistance can exhibit enhanced in vivo efficacy. Lastly, RNA aptamers, labelled with fluorophores at detection-sensitive positions, could have broad applications as sensors because of their high affinity and specificity for a variety of substrates, including molecular cancer markers.

Solid-phase chemical synthesis is limited to short RNAs (≤60 nucleotides (nt)) and iso-dose-labelled versions of these chemicals are not commercially available. Solution-phase transcription with bacteriophage T7 or SP6 RNA polymerases (RNAPs) allows synthesis of longer RNAs, but is limited to either uniformly labelled samples or those labelled by base type. Incorporation of modified nucleotides by T7 RNAP is limited by poor processivity during initiation. However, once the RNA has been extended to at least 10 nt, elongation is highly processive. Elongation can be stalled by omitting the nucleoside 5′-triphosphate (NTP) required for transcription beyond a specified template position, and can be restarted by addition of the missing NTP(s). This pause–restart method enables synthesis of RNAs with region- or position-specific modifications by using solid-phase-coupled DNA templates and limiting combinations of NTPs, so that elongation can be stopped at specific positions and restarted with a new NTP mix (Fig. 1). A similar approach has been attempted for in vitro peptide synthesis. The pause–restart method has been used for mechanistic studies, although the synthesis of RNA using this method on a preparative scale has not been realized.

To achieve this, we first coupled 5′-biotinylated DNA templates to streptavidin-agarose beads (Fig. 1a) and then incubated them with T7 RNAP and an NTP mix lacking cytidine 5′-triphosphate (CTP)16,17. This causes elongation to stall at +14, where the first CTP would otherwise be incorporated. After extensive washing with buffer, elongation was resumed by addition of a new NTP mix containing CTP. RNA synthesis can be repeatedly paused and resumed in this way to incorporate modified or labelled NTPs at desired locations (Fig. 1b). Once transcription extends past the position where specific labelling is desired, the reaction is terminated and re-initiation prevented by addition of an NTP mix without GTP, addition of heparin, which prevents reinitiation, and a complete NTP mix, or quenching at 4 °C. The entire process can be repeated several times with the same bead-bound template to increase RNA yield. The probability of incorporating a wrong nucleotide occurs 1 in 20,000 (ref. 16). Development of the method—dubbed PLOR for position-selective labelling of RNA—is described in detail in Extended Data Figs 1 and 2. As the method requires multiple washing and NTP addition steps, we have developed a fully automated platform for performing PLOR (Fig. 1c; see Methods and Extended Data Fig. 2 for details of design and operations of the platform). We illustrate applications of this method using the 71-nt aptamer domain of an adenine riboswitch, riboA71 (Fig. 2a), which changes conformation upon binding adenine. The sequence of riboA71 was from *Vibrio vulnificus* with the first stem sequence modified to enhance the transcription yield. This adenine riboswitch controls gene expression through translational activation. The class of adenine riboswitches has been studied using various methods.
The crystal structures of the adenine-bound form of riboA71 (Protein Data Bank (PDB) accession numbers 1Y26 and 4TZX) show loops 1 and 2 in a kissing loop interaction. Selective labelling of these loops would allow study of their structure and dynamics in solution. Using PLOR, we synthesized RNAs in which only loop 1 (Lp1-CN) or loop 2 (Lp2-CN) was labelled with $^{13}$C/$^{15}$N isotope so that only signals from those regions were observed in heteronuclear correlation spectra (Fig. 2b, c). These RNA molecules fold in the same way as those synthesized using conventional in vitro transcription, as illustrated by superimposing their spectra on that of the fully $^{13}$C/$^{15}$N-RNA (Extended Data Figs 4 and 5). These samples could be used for detection of loop–loop interactions by recording various heteronuclear-edited/filtered nuclear Overhauser effect (NOE) spectra. We further demonstrated the utility of PLOR by $^{13}$C/$^{15}$N-labelling both loops (Lp1+2-CN) (Fig. 2d), showing that the spectrum of Lp1+2-CN is consistent with the superimposition of the Lp1-CN and Lp2-CN spectra.

We also used PLOR to $^{13}$C/$^{15}$N-label the linkers that connect adjacent duplexes: linker 2 (Lk2-CN) and stem 1 plus linker 1 (S1+Lk1-CN) (Fig. 2a). The use of selectively labelled RNAs greatly simplified the spectra (Fig. 2e, f), but overlaps were still present, as evidenced by the overlapping signals of A4, 5, 7, and 9 in the C8H8 region (Fig. 2f). Such overlaps are common for RNA duplexes because of similar chemical environments around the detected nuclei in the A-type duplex conformation. To overcome this, we used PLOR to generate a sample (S1+Lk1-H) in which stem 1 and linker 1 were protonated, and the rest of the molecule was deuterated (this RNA was not $^{13}$C/$^{15}$N-labelled). By recording a two-dimensional NOE spectrum of S1+Lk1-H, and performing an NOE-walk in the H1‘–H6/H8 region (Fig. 2g, top panel), the overlaps were partly resolved and proton chemical shifts assigned. Without partial deuteration, interpretation of this region of the spectrum would be impossible (Fig. 2g, bottom panel). With the increasing availability of high-field NMR spectrometers, the use of selectively protonated/deuterated RNA samples is one of the most effective strategies for RNA structure determination.

While all labelling schemes described could, in theory, be performed using segmental labelling by ligations, such a method is impractical for placing isotope-labelled residues in multiple discrete positions. This can, however, be readily achieved using PLOR. Figure 2h shows the spectrum of another sample generated by PLOR, 4nt-CN, where only residues at positions A21, C38, U39, and G60 were $^{13}$C/$^{15}$N-labelled.

RNAs may exist in multiple conformations, and while NMR is suitable for detecting such conformers, interpretation of data is often complicated by crowded signals. Coexistence of conformers can be unambiguously identified by monitoring a single residue at a critical structural location. It has been shown that riboA71 thermodynamic data are best interpreted by a four-state model, and that its conformation is affected by ligand and Mg$^{2+}$. Inspection of the crystal structure of riboA71 reveals the U39 N3–H3 imino group forming a hydrogen bond with N3 of the adenine ligand (Fig. 3a). We used PLOR to incorporate a single $^{1}$H/$^{13}$C/$^{15}$N-labelled uridine-5′-triphosphate (UTP) at this position (U39-CN). In the absence of adenine we detected multiple weak cross-peaks for both the H6–C6 and the N3–H3 imino groups in the $^{13}$C and $^{15}$N heteronuclear single-quantum correlation signals of U39, respectively, indicative of multiple conformations (Fig. 3b, c). At a 1:1 RNA:adenine ratio we detected four distinct cross-peaks (Fig. 3b, c), implying coexistence of at least four conformations of the adenine-binding pocket, in agreement with the four-state model.

This type of information would be difficult to obtain unambiguously using a uniformly labelled sample, where extensive overlap and cross-peaks from other residues would obscure the weak cross-peaks associated with alternative conformations of the complex. Nevertheless, it is noteworthy to point out that a slightly different multi-state model has also been proposed on the basis of a study of an adenine riboswitch aptamer domain of a slightly different sequence using NMR and a uniformly labelled RNA sample.
Fig. 4a). smFRET measurements with this RNA in 0 (loop 2; Cy5 was placed at the end, rather than in the middle, of loop 2) and fluorescent Cy3 at U24 (loop 1), and Cy5 at either U65 (stem 1) or C55 (loop 2; Cy5 was placed at the end, rather than in the middle, of loop 2) to avoid steric interference with respect to loop–loop interactions; Fig. 4a). smFRET measurements with this RNA in 0 μM adenine yield an $E_{\text{FRET}}$ histogram (Fig. 4b) that can be approximated as the sum of two broad Gaussian curves, indicating the presence of at least two populations of molecules with distinct loop–loop interactions. Approximately 20% of the population is in a high-FRET (≈0.92) conformation consistent with the U24–C55 distance (<20 Å) observed in the crystal structure. Upon addition of ligand, this conformation becomes increasingly populated (Fig. 4b, c). To further evaluate the utility of PLOR for incorporation of multiple modifications at specific sites, we imaged single U24Cy3–C55Cy5-B molecules (Fig. 4d), and not only confirmed the presence of the three RNA modifications, but also showed from the individual time traces (Fig. 4e) that adenine is not required to form the high-FRET conformation.
PLOR allows efficient synthesis of milligram quantities of RNA (Figs 2 and 3, Extended Data Fig. 6 and Extended Data Table 2). Both the His-tagged T7 RNAP and the DNA templates are reusable, and the same batch of bead-attached DNA templates was used for all samples generated here. The DNA templates can be produced routinely and economically in micromole quantities by PCR; all reagents are commercially available, and the technology is automated using a robotic platform. Extension of PLOR for other applications is easily envisioned. For example, variants of T7 RNAP have been developed that allow incorporation of nucleotides with modified ribose groups into RNA which can, among other useful characteristics, provide resistance to RNases for in vivo applications5,30. PLOR may also be useful for X-ray crystallography by incorporating heavy-atom-derivatized nucleotides into the RNA at one or more desired positions for phase determination of X-ray diffraction data.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 4 November 2014; accepted 25 February 2015.

1. Serganov, A. et al. Structural basis for discriminative regulation of gene expression by adenine- and guanine-sensing mRNAs. Chem. Biol. 11, 1729–1741 (2004).

2. Geiger, A., Burgstaller, P., von der Eltz, H., Roeder, A. & Famulok, M. RNA aptamers that bind l-arginine with sub-micromolar dissociation constants and high enantioselectivity. Nucleic Acids Res. 24, 1029–1036 (1996).

3. Mariol, T. et al. Aptamers: molecular tools for analytical applications. Anal. Bioanal. Chem. 390, 589–1007 (2008).

4. Shangguan, D. et al. Cell-specific aptamer probes for membrane protein elucidation in cancer cells. J. Proteome Res. 7, 2133–2139 (2008).

5. Wilson, C. & Keefe, A. D. Building oligonucleotide therapeutics using non-natural chemistries. Curr. Opin. Chem. Biol. 10, 607–614 (2006).

6. Usman, N., Ogilvie, K. K., Jiang, M. Y. & Cedergren, R. J. Automated chemical synthesis of long oligonucleotides using 2′-O-silylated ribonucleoside 3′-O-phosphoramidites on a controlled-pore glass support: synthesis of a 43-nucleotide sequence similar to the 3′-half molecule of an Escherichia coli formyl-methionine tRNA. J. Am. Chem. Soc. 109, 7845–7854 (1987).

7. Scaringe, S. A., Wincott, F. E. & Caruthers, M. H. Novel RNA-synthesis method using 5′-O-silyl-2′-O-orthorster protection groups. J. Am. Chem. Soc. 120, 11820–11821 (1998).

8. Milligan, J. F., Groebe, D. R., Witherell, G. W. & Uhlenbeck, O. C. Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. Nucleic Acids Res. 15, 8783–8798 (1987).

9. Lu, K., Miyazaki, Y. & Summers, M. F. Isotope labeling strategies for NMR studies of RNA. J. Biomol. NMR 46, 113–125 (2010).

10. Lyakhov, D. L. et al. Pausing and termination by bacteriophage T7 RNA polymerase. J. Mol. Biol. 280, 201–213 (1998).

11. Scaringe, S. A., Wincott, F. E. & Caruthers, M. H. Novel RNA synthesis method using 5′-O-silyl-2′-O-orthorster protection groups. J. Am. Chem. Soc. 120, 11820–11821 (1998).

12. Shangguan, D. 

13. Nudler, E., Gusarov, I. & Bar-Nahum, G. Methods of walking with the RNA elongating T7 RNA polymerase. Methods Enzymol. 371, 170–179 (2003).

14. Pavlov, M. Y., Freistroffer, D. V. & Ehrenberg, M. Synthesis of region-labelled proteins for NMR studies by in vitro translation of column-coupled mRNAs. Biochimie 79, 415–422 (1997).

15. Marble, H. A. & Davis, R. H. RNA transcription from immobilized DNA templates. Biotechnol. Prog. 11, 393–396 (1995).

16. Huang, J., Briege, L. G. & Sousa, R. Misincorporation by wild-type and mutant T7 RNA polymerases: identification of interactions that reduce misincorporation rates by stabilizing the catatonically incompetent open conformation. Biochemistry 39, 11571–11580 (2000).

17. Makarova, O. V., Makarov, E. M., Sousa, R. & Dreyfus, M. Transcribing of Escherichia coli genes with mutant T7 RNA polymerases: stability of lacZ mRNA inversely correlates with polymerase speed. Proc. Natl Acad. Sci. USA 92, 12250–12254 (1995).

18. Mentessana, P. E., Chin-Bow, S. T., Sousa, R. & McAllister, W. T. Characterization of halted T7 RNA polymerase elongation complexes reveals multiple factors that contribute to stability. J. Mol. Biol. 302, 1049–1062 (2000).

19. Leiply, D. & Draper, D. E. Dependence of RNA tertiary structural stability on Mg2+ concentration: interpretation of the Hill equation and coefficient. Biochemistry 49, 1843–1853 (2010).

20. Delfosse, V. et al. Riboswitch structure: an internal residue mimicking the purine ligand. Nucleic Acids Res. 38, 2057–2068 (2010).

21. Noeske, J. et al. An intermolecular base triple as the basis of ligand specificity and affinity in the guanine- and adenine-sensing riboswitch RNAs. Proc. Natl Acad. Sci. USA 102, 1372–1377 (2005).

22. Lee, M. K., Gal, M., Frydman, L. & Varani, G. Real-time multidimensional NMR follows RNA folding with second resolution. Proc. Natl Acad. Sci. USA 107, 9192–9197 (2010).

23. Lemay, J. F., Penedo, J. C., Tremblay, R., Lilley, D. M. & Lafontaine, D. A. Folding of the adenosine riboswitch. Chem. Biol. 13, 857–868 (2006).

24. Dalgarno, P. A. et al. Single-molecule chemical denaturation of riboswitches. Nucleic Acids Res. 41, 4253–4265 (2013).

25. Rieder, R., Lang, K., Graber, D. & Micura, R. Ligand-induced folding of the adenosine deaminase A-riboswitch and implications on riboswitch translational control. ChemBioChem 8, 896–902 (2007).

26. Zhang, J. & Ferré-D’Amare, A. R. Dramatic improvement of crystals of large RNAs by cation replacement and dehydration. Structure 22, 1363–1371 (2014).

27. Zhang, Q., Steizer, A. C., Fisher, C. K. & Al-Hashimi, H. M. Visualizing spatially correlated dynamics that directs RNA conformational transitions. Nature 450, 1263–1267 (2007).

28. Leiply, D. & Draper, D. E. Effects of Mg2+ on the free energy landscape for folding a purine riboswitch RNA. Biochemistry 50, 2790–2799 (2011).
29. Ferre-D’Amare, A. R. & Doudna, J. A. Use of cis- and trans-ribozymes to remove 5′ and 3′ heterogeneities from milligrams of in vitro transcribed RNA. Nucleic Acids Res. 24, 977–978 (1996).

30. Pinheiro, V. B. & Holliger, P. Towards XNA nanotechnology: new materials from synthetic genetic polymers. Trends Biotechnol. 32, 321–328 (2014).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank D. E. Draper, A. Bax, A. Byrd, M. Summers, A. Rein, and J. Strathern for discussions. This work was supported in part by the Intramural Research Programs of the National Cancer Institute, the National Institute of Diabetes, Digestive and Kidney Diseases, the National Heart, Lung and Blood Institute; by the Intramural Antiviral Target Program (IATAP) of the Office of the Director, National Institutes of Health; by the 2013 Director’s Challenge Innovation Award of the National Institutes of Health; by the fund from the National Cancer Institute under contract number HHSN261200800001E; by the National Science Foundation (CHE1266416, PHYS1125844); by the National Institutes of Health Molecular Biophysics Training Program (T32 GM-065103); and by the W. M. Keck Foundation, and the National Institute of Standards and Technology. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government.

Author Contributions Y.L. and P.Y. performed RNA synthesis and NMR experiments; E.H. and D.J.N. designed and performed smFRET experiments; J.W. contributed to chemical shift assignments; J.Z. and A.R.F. helped to design the DNA template using PCR, crystallized and determined the three-dimensional structure of the PLOR-generated riboA71; J.Y., M.A.D., D.C., and S.L. helped to characterize RNA; R.S. provided critical advice about T7-enzymatic synthesis and revised the manuscript; J.R.S. helped to characterize RNAs and revised the manuscript; Y.L., E.H., J.Z., J.R.S., and Y.-X.W. prepared figures; Y.-X.W. designed PLOR and the automated platform and wrote the manuscript. All authors discussed the results.

Author Information Structure coordinates have been deposited in Protein Data Bank under accession number 4XNR. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to Y.-X.W. (wangyunx@mail.nih.gov).
METHODS

No statistical methods were used to predetermine sample size.

DNA templates. Initial DNA templates used for riboA71 synthesis by the PLOR method were purchased (Aventia Bioscience). The 5′-biotinylated non-coding strand, 5′-TCTGATTTACCGTCTGCTATAAGCACCTACATATGGGAGATATATCTCAATGCATATCGTTTGGAGCTTACACAGCTTCAATACCTCATTAGTCCAAA, was annealed to its complementary strand (the coding strand) in buffer A (40 mM Tris-HCl, 100 mM K$_2$SO$_4$, 6 mM MgSO$_4$, pH 8.0). An 18-nt spacer before the T7 RNAP promoter (underlined) was inserted to reduce crowding on the bead surface and potential steric hindrance to the T7 RNAP. Double-stranded DNA templates were also generated by touchdown PCR (TD-PCR) to abrogate non-specific products. The commercially synthesized DNA strands were used as templates in the TD-PCR. The forward primer contained biotin at its 5′-end and the reverse primer contained two 2′-O-methylguanosine (mG) residues to reduce non-templated nucleotide addition to transcripts. The primer sequences were: 5′-/Bio/TCTGATTTACCGTCTGCTATAAGCACCTACATATGGGAGATATATCTCAATGCATATCGTTTGGAGCTTACACAGCTTCAATACCTCATTAGTCCAAA and 5′-mGmGmGGAATATATCAGTCmGmGTTAGGTAGAATTAAAGGAGAGCAGTACTACGACGATTATC (Integrated DNA Technologies), and 5′-9mGmGGAATATATCAGTC (Integrated DNA Technologies). The reagents used in the TD-PCR reactions were as follows: 0.1 μM template, 10 μM primers, 200 μM dNTPs, Taq DNA polymerase, and PCR buffer (50 mM KCl, 10 mM Tris-HCl, 2 mM MgCl$_2$, pH 8.4). The cycling program of the TD-PCR involved two separate phases: the touchdown phase and the PCR phase. During the touchdown phase, the annealing temperature was reduced from 75 °C to 45 °C over 20 successive cycles. Each touchdown cycle, therefore, began with melting at 95 °C (5 min for first cycle), followed by annealing (temperature varies) for 45 s and elongation at 72 °C for 1 min. The PCR phase was a generic amplification stage of 25 cycles of 95 °C for 30 s, annealing at 50 °C for 30 s, and elongation at 72 °C for 1 min. The DNA, 5′-TCTGATTTACCGTCTGCTATAAGCACCTACATATGGGAGATATATCTCAATGCATATCGTTTGGAGCTTACACAGCTTCAATACCTCATTAGTCCAAAAT CCGCCGAGCAGCCTGTCAGCTGCGGAGCTTAGCTGAAACTGTCCTTGGTGAAACCACTTAATACCCGGAAGGTGGTGGTGGGAC CCTCCG-3′, and its complementary strands were used as templates in TD-PCR for turnip virus (TCV) PCR. The sequences of the primers used in the TD-PCR for TCV PCR were, 5′-Bio/TCTGATTTACCGTCTGCTATAAGCACCTACATATGGGAGATATATCTCAATACCTCATTAGTCCAAAACCGCCGAGCAGCCTGTCAGCTGCGGAGCTTAGCTGAAACTGTCCTTGGTGAAACCACTTAATACCCGGAAGGTGGTGGTGGGAC CCTCCG-3′, and its complementary strands were used as templates in TD-PCR for TCV PCR.

Solid-phase attachment of DNA templates. For this study, commercial agarose beads (30–165 μm diameter), coated with either streptavidin or neutravidin (Thermo Fisher Scientific), were used as the solid-phase support. In our hands, the neutravidin-coated agarose beads appeared to have a better retention of the biotinylated DNA templates and, therefore, were used as the solid-phase support in most cases. The double-stranded DNA template (0.8 μmol) was incubated with neutravidin-coated agarose beads (8 ml in buffer A for 3 days at 4 °C). The bead-associated templates were washed by repeated rinsing and passing through a Pierce Centrifuge Column with a 30 μm average pore size (Thermo) to remove non-bound DNA templates, and were stored at 4 °C. Approximately 80% of the DNA template was attached to the beads on the basis of 20–30 μg/50 μl of template. The procedure of preparing the 15N-H1-TCV (fluorescence labelled TCV)-bound DNA templates was functionalized with Cy3 by adding a Cy3-NHS ester and 0.5 volumes of 0.3 M sodium bicarbonate buffer (pH 8.3), followed by incubation for 1 h at 37 °C. The DNA-attached agarose beads were stored at 4 °C or −20 °C for long-term storage, and the DNA-attached agarose beads were rinsed at least five times and stored at 4 °C for future use. The final RNA products were gel-purified using 15% polyacrylamide gel electrophoresis (PAGE)-urea gels before being used for NMR data collection.

The procedure of preparing the 5′-NH-TCV (5′-labelled at 5′-end, shown in Extended Data Fig. 6a) is the same as described earlier. The total reaction volume used in the 5′-NH-TCV was 30 ml, and the concentrations of T7, DNA, ATP, GTP, and CTP in the initiation were 20 μM, 20 μM, 0.64 mM, 4.8 mM, and 0.32 mM, respectively. In the first elongation cycle, 20 μM NTP, 20 μM NTP-GTP and 60 μM NTP-UTP were added in the reaction. In the termination step, 0.4 mM ATP, 0.380 mM CTP, 0.5 μM GTP, and 0.360 mM UTP were added.

NMR experiments. All NMR samples in this study were in a buffer containing 10 mM KH$_2$PO$_4$, 30 mM KCl, 2 mM MgCl$_2$, pH 6.8. Adenine (5 mM) was added to all riboA71 NMR samples except U39-CN. Two-dimensional 15N-HOESY, 15N-ROSY, and 15N-ROSY spectra were collected at 25 °C for riboA71 and 15 °C for the 104-nt TCV RNA on Bruker spectrometers operating at a proton frequency of 850 MHz or 700 MHz and equipped with a triple-resonance cryo-probe. All spectra were processed and analysed with nmrDraw$^40$ and NMR View$^41$ (One Moon Scientific). The concentrations for Lp1-CN, Lp2-CN, Lk2-CN, S1-CN, Lk1-CN, 4-nt-CN, S1-CN, Lk1-H, U39-CN, 71-n-CN, 5′-NH-TCV, and 15N-104-nt-TCV (fully 15N-labelled TCV) were 0.5, 0.3, 0.2, 0.35, 0.4, 0.3, 0.4, 0.35, 0.5, 0.4, and 0.5 mM, respectively and the sample volumes were ~250 μl in Shimex tubes.

Fluorescent samples generated by PLOR. Five fluorescent samples were produced by the PLOR method: U24Cy3–C55Cy5, U24Cy3–C55Cy5-B, U24Cy3–U65Cy5, U24Cy3–U65Cy5-B and U24A55–U65A55–B (Extended Data Table 3). As an example, a detailed description of the U24Cy3–C55Cy5 synthesis is provided. DNA and T7 concentrations used in the preparation were 5 μM in 3 ml transcription buffer (buffer B). Then, NTPs were added as follows: 0.48 mM ATP, 0.48 mM GTP and 48 μM UTP (initiation); 10 μM ATP, 10 μM CTP and 10 μM UTP (elongation, cycle 1); 5 μM ATP and 5 μM GTP (cycle 2); 5 μM UTP (cycle 3); 5 μM ATP and 5 μM 5-aminooxallyl UTP (TriLink) (cycle 4); 5 μM ATP, 30 μM GTP, and 30 μM UTP (cycle 5); 15 μM ATP, 15 μM CTP, and 15 μM UTP (cycle 6); 5 μM ATP and 10 μM GTP (cycle 7); 10 μM CTP and 10 μM UTP (cycle 8); 15 μM ATP and 5 μM Cy5-UTP (cycle 9, at 37 °C); 5 μM ATP, 5 μM GTP, and 15 μM UTP (cycle 10); 10 μM ATP, 20 μM CTP, and 25 μM UTP (termination). Cy5-UTP was from GE Healthcare. The RNA transcript was lyophilized and then dissolved in sterile distilled H$_2$O. U24, a 24-mer oligonucleotide containing an aminoallyl group, was functionalized with Cy3 by adding a Cy3-NHS ester and 0.5 volumes of 0.3 M sodium bicarbonate buffer (pH 8.3), followed by incubation for 1 h at 37 °C, 12 h at 25 °C, and then 1 h at 37 °C. All incubations were performed in the dark. Homogeneity of doubly labelled riboA71 samples was assessed via 15N denaturing PAGE and analytical high-performance liquid chromatography before being used in free-diffusion smFRET experiments.

smFRET experimental set-up. A home-built microscope system was used to acquire all of the single-molecule fluorescence data, as recently described elsewhere$^{42}$. Briefly, alternating excitation 38 of the donor (Cy3) and acceptor (Cy5) was used to excite the sample to cover a 200–635 nm (20 MHz, 90 ps) range, respectively. The interleaved pulses are directed into the back aperture of a 1.2 numerical aperture water objective, which is used in a standard epifluorescence configuration, both to focus the excitation source down to a ~250 nm diffraction limited spot and to collect the emitted fluorescence photons. The collected fluorescence is then directed out of the side-port of the microscope and focused through a 50 μm confocal pinhole, which spatially defines the detection volume. The photons are then separated by both colour (donor, acceptor) and polarization (horizontal, vertical) before being directed onto one of four single-photon avalanche photodiodes, where the arrival time of each detected photon is recorded.

Free-diffusion smFRET. After being deemed suitable, the fluorescently labelled RNA samples were diluted to a final concentration of ~100 pM in single-molecule imaging buffer (10 mM HEPES, 50 mM NaCl, 1 mM Mg$^{2+}$, 0–1 mM Aed, 2 mM Trolox, 5 mM 3,4-protopcatachac acid (PCA), 100 mM protocatechu
Quantification of experimental yields. The experimental yields were measured by quantifying gel band intensities. Transcription reaction (0.2–1 μl) was loaded onto a 12% mini-PAGE gel. For quantification, 0.5, 1.0, and 1.5 pmol of pure Riboa71 samples were run as a standard. The gels were stained by Sybr Gold (Life Technologies) for 10 s, and then visualized using a Gel DocTM EZ imager (Bio-Rad). The stained gel band intensities were quantified using Image Lab 4.1 (Bio-Rad), and were used to determine the experimental yield of each synthesis product, given as a percentage relative to the amount of template used in the synthesis (Extended Data Table 2).

Estimating synthesis efficiencies. The transcription yields depend on both the reaction cycle efficiencies and the total number of cycles in the synthesis. For simplification, the average termination efficiency was taken as equal to the average elongation efficiency. The apparent efficiencies of initiation and elongation, therefore, were estimated using the equation

\[ \text{Yield} = I \times E^\text{C} \]  

where \( I \) and \( E \) are the percentage efficiencies of initiation and elongation, respectively, and \( n \) is the total number of cycles used for synthesis. As there are two unknown variables, \( I \) and \( E \), a second equation was needed. Because the same procedures, equipment, and reagents were used to make each NMR sample, one would expect similar values for \( I \) and \( E \) among different samples. Therefore, we used a crude method to calculate the efficiencies by setting \( I \) and \( E \) of one sample equal to the \( I \) and \( E \) of another ‘reference’ sample, and then solving for each variable. The same reference sample, Si+Lk1-CN, was used in each calculation. For example, the Lp1-CN sample was generated using a five-cycle synthesis with an experimental yield of 30.3%. Similarly, Si+Lk1-CN was generated using a one-cycle synthesis with a yield of 40.3%. Using equation (1), then, \( I \) and \( E \) for Lp1-CN were estimated as 43.3% and 93.1%, respectively. Using this method for all Riboa71 samples, the average values for \( I \) and \( E \) were 46.9% and 87.8%, respectively (Extended Data Table 2). The ‘theoretical’ yields for each sample were calculated using their individual efficiencies.

Discussion of PLOR efficiency. The efficiency of the PLOR method depends heavily on a number of factors, such as the number of cycles involved in achieving a desired labelling scheme (equation (1)), the inherent initiation and elongation efficiencies for a given RNA sequence, the purity of NTPs, the efficiency of the washing/filtration cycles, the T7 RNAP quality, and the efficiency of DNA-bead attachment. For any given commercially bought NTP, the presence of other NTPs can be as high as 5%. Moreover, it is expected that some degree of cross-contamination among nucleotides occurs during synthesis as a result of carryover from one reaction cycle to the next owing to inefficient washing/filtration during solid-phase extraction. Such impurity and cross-contamination in the reaction mixture in any given cycle may result in non-synchronized transcription, where initiation or elongation may pause at non-designated positions, depending on the availability of various NTPs. In addition, sufficient quantities of all four NTPs due to contamination may lead to full-length transcripts prematurely (that is, before the designed termination phase) as well as re-initiation of transcription by recycled T7, again resulting in non-synchronized transcription. During the initiation and elongation phases, abortive and full-length transcripts generated as a result of non-synchronized transcription are removed during the washing/filtration cycles and thus do not affect the purity of the full-length products collected at the end of the termination phase. However, highly purified NTPs and a more robust solid-phase extraction may alleviate non-synchronized transcription, thereby increasing the yield of the correctly labelled, full-length product. Leaking of DNA templates from the agarose beads may also result in lower transcription yields. The leaking is possibly due to a high off-rate of biotin binding to streptavidin beads and may be mitigated by using glass beads as the solid phase or using mutant streptavidinagarose beads with a much reduced off-rate.

Since it is an enzyme-catalysed reaction, there are several inherent limitations. In principle, for a given RNA sequence, the yield of the PLOR method is directly correlated to the yield of the regular in vitro transcription synthesis. The specific limitations of PLOR is that only three types of residue are allowed in the first 13 residues. This would probably limit the ability to label residues in the first 13 nt in the 5’ end of a sequence.

Automated RNA synthesis platform. Since the PLOR process consists of repetitive steps and cycles, we designed and constructed an automated platform to perform the synthesis. This prototype RNA synthesizer (Extended Data Fig. 2) utilizes the modular liquid delivery system made by Zinsser Analytical, and is the first instrument to perform automated RNA synthesis using PLOR. The platform consists of several modules, including a rotating reaction chamber, a solid-phase extraction cube (SPEC), a pair of robotic arms, reaction vessels, four long-arm liquid transfer tips and four syringes, a six-way valve for liquid handling, and seven stations. The reaction chamber is located at the incubation station (1-station), where the reaction vessels containing reagents incubate at either 37 °C (initiation)

---

**RESEARCH LETTER**

dioxygenase (PDCD), pH 7.5). The PCA/PDCD oxygen10, and the blinking suppres-
sant, TROLOX46, were used together to enhance the photo-stability of cyanine-
derived fluorophores (that is, Cy3 and Cy5).

Surface-immobilized sfFRET. The selectively incorporated biotin moiety per-

mits surface immobilization of individual fluorescently labelled nucleic acids, as

recently described elsewhere47. Briefly, the piezoelectric scanning stage of the experi-

mental apparatus provides the ability to raster-scan an area of the sample

(for example 15 μm × 15 μm), resulting in a false-colour image of the surface.

The stages were designed to position the individual fluorescence features

within the focus of the microscope objective, allowing continuous excitation

and collection of fluorescence from a single molecule. FRET values are derived

from the ratio of acceptor (A) fluorescence to the sum of both donor (D) and

acceptor (A) fluorescence intensities (that is, \( I_A / (I_D + I_A) \)). These fluorescence

intensities are background corrected and adjusted for various experimental arte-

facts such as acceptor direct excitation, donor cross-talk on the acceptor channel,

and differential quantum yields of the two fluorophores48.

Crystallographic analysis. Unmodified Riboa71 RNA synthesized using PLOR

(Extended Data Table 1) was crystallized as previously described48. Briefly, for

crystallization by vapour diffusion, a solution containing 700 μM RNA, 50 mM

KOAc pH 6.8, 100 mM MgCl₂, 1 mM spermine, and 5 mM adenosine was mixed 1:1

(v/v) with a reservoir solution comprising 50 mM Tris-HCl pH 8.5, 100 mM KCl,

10 mM MgCl₂, and 30% polyethylene glycol (PEG) 400. Crystals grew to maximal

dimensions of 250 μm × 50 μm × 50 μm in 1–2 weeks at 4 °C and were directly

flash-frozen by plunging into liquid nitrogen. The structure of the PLOR-gener-

ated riboswitch-adenine complex was solved by molecular replacement with

PHASER using a published structure49 (PDB accession number 4TZX) as a search

model. Initial solutions were subjected to manual rebuilding50 interspersed with

iterative rounds of rigid-body, simulated-annealing, and individual isotropic

B-factor refinement using PHENIX51. Refinement statistics are summarized in

Extended Data Table 4. The coordinates and structure factors have been deposited in

Protein Data Bank under accession number 4XR.

Optimizing the transcription reaction. Fundamental to our method of the pre-

parative T7-catalysed reactions is a novel stoichiometry concentration of reac-
tants. This novel scheme is counter-intuitive to the conventional in vitro

transcription, where multiple millimolar concentrations of NTPs are used.

Various reaction conditions, including buffer conditions, NTP concentrations,

reaction time, incubation temperatures, and T7 RNAP concentrations were tested in

Lp2-CN synthesis to optimize the yields. Among all factors, the effects of NTP

concentrations on yields. The initiation phase requires a substantially higher

concentration, however, affected yields dramatically

where the reaction vessels containing reagents incubate at either 37

°C for 10 min, followed by addition of NTPs dissolved in the transcription buffer and

further incubation at 37 °C for 15 min, resulted in a 20% increase of the final product

with incubating all reagents simultaneously at 37 °C for 15 min. Incubating the reaction at 37 °C gave higher yields than incubating at 25 °C, 35 °C, or 40 °C, as expected. Incubating the initiation reaction mixture at 37 °C for 10–20 min was sufficient for initiation for >30 min resulted in decreased yields by as

much as 50%. Types of anion had little effect on the reaction yield. KCl/MgCl₂ and

K₃SO₄/MgSO₄ were tested and the product yields were very similar (Extended Data Fig. 3b, panel II). Mg²⁺ concentration, however, affected yields dramatically and was optimal at 6 mM (Extended Data Fig. 3b, panel III). This differs from the standard transcription reaction where 20–26 mM Mg²⁺ is optimal. A series of K₃SO₄ concentrations (0, 20, 40, 60, 80, and 100 mM) was tested. Yields were similar when K₃SO₄ concentrations were between 20 and 100 mM. Transcription initiated with 100 mM K₃SO₄ had a higher yield than having no K₃SO₄ present during initiation (Extended Data Fig. 3b, panel IV). A series of different values of pH for the initiation reaction for >30 min resulted in decreased yields by as

much as 50%. The optimal pH for the initiation was between 7.5 and 8.0. We tested the impact of NTP concentrations on yields. The initiation phase requires a substantially higher

NTP concentration (~2 mM) than the elongation phase, but this concentration is still relatively low compared with the standard transcription reactions where NTP concentrations are typically 4–6 mM. The optimal NTP concentrations for elonga-
tion using PLOR are in the micromolar range (Extended Data Table 1). This is advantageous as it means that the expensive labelled NTPs, which are used only

during elongation, can be used at low concentrations to achieve a high percentage

of their incorporation into the RNA. Higher NTP concentrations during elonga-
tion led to reduced yields (Extended Data Table 3, panel II). The ratio of T7RNAP:DNA was tested for its effect on yield. The yield at 1:1 was similar to

that at 2:1 and higher than that at 0.5:1. Higher ratios resulted in lower yields.

The effect of reaction time in the termination step on yield is shown in Extended

Data Fig. 3b, panel VIII.

©2015 Macmillan Publishers Limited. All rights reserved
or at 25 °C (elongation or termination). There are two reagent-stations (R-stations): one attached to a cooling unit to keep reagents at 4 °C, and another at room temperature. The reagent-addition/filter station (RF-station) has two main purposes: (1) it is where the reaction reagents are transferred to the reaction vessels, and (2) it is where solid-phase extraction is performed through the addition of buffers and nitrogen gas, and is connected to a vacuum pump for the removal of liquid waste. The wash station (W-station) washes the tips/syringes between liquid transfers. The reaction chamber cover and the SPEC are docked at the cover station (C-station) and SPEC-station, respectively, when not in use. The operation for performing the PLOR RNA synthesis is controlled by a software program with customizable input files. Certain types of information require user input to the computer based on the individual protocol, such as volumes of the various NTPs, for each cycle, to be delivered into the reaction vessels, and the number of reaction cycles. Other variables that may be changed less frequently are set as default, for each cycle, to be delivered into the reaction vessels, and the number of reaction cycles. Other variables that may be changed less frequently are set as default, for each cycle, to be delivered into the reaction vessels, and the number of reaction cycles.

Here we describe the blocks for the automated synthesis of the Lp1-CN sample. Pre-incubation (initiation). After manual addition of the template-attached beads and T7 RNAP to the reaction vessels, which are housed in the reaction chamber at the I-station, the robotic arms seal the reaction chamber with the cover, and the chamber temperature is ramped to 37 °C with gentle rotation. After 10 min, the robotic arms remove the reactor cover and transfer it to the C-station; they then transfer the reaction vessels to the RF-station.

Reagent addition (initiation). The liquid-transfer system transfers the user-specified amounts of ATP, GTP, UTP, and transcription buffer individually from their respective stock-solution containers from the R-station (a cooling unit is used to keep all stock solutions and the transcription buffer at 4 °C) to the reaction vessels in the RF-station. The syringe tips are then rinsed three times with wash buffer at the W-station.

Incubation (initiation). The robotic arms return and seal the cover on the reaction chamber, and the reaction chamber begins gentle rotation at 37 °C for 15 min. Once the rotation stops, the robotic arms move the cover to the C-station and the reaction vessels to the RF-station. The temperature of the reactor chamber drops back to 25 °C.

Solid-phase extraction (SPE) (initiation). The robotic arm attaches the SPEC to the top of the reaction vessels, one at a time, at the RF-station. A nitrogen flow at 80 p.s.i. is applied to the SPEC for 25 s to remove the liquid phase from the reaction vessels. The robotic arm then returns the SPEC to the SPEC-station.

Rinsing (initiation). Wash buffer is added to the reaction vessels by the liquid-transfer system, and the tips are then rinsed three times at the W-station. The reaction vessels are placed in the reactor chamber by the robotic arm, rotated for ~1 min, and then returned to the RF-station.

The SPE (initiation) and rinsing (initiation) blocks are repeated five times, followed by one more SPE (initiation) block before proceeding to the elongation steps.

Reagent addition (elongation cycle 1). This block is the same as the reagent addition (initiation) block except that the NTPs added are ATP, CTP, and UTP (see Extended Data Table 1).

Incubation (elongation cycle 1). This block is the same as the incubation (initiation) block except that the incubation is performed at 25 °C for 10 min.

SPE (elongation cycle 1) and rinsing (elongation cycle 1). These blocks are the same as SPE (initiation) and rinsing (initiation) blocks, respectively.

Each of the remaining cycles includes the following blocks. Reagent addition (elongation), incubation (elongation), SPE (elongation) and rinsing (elongation). The NTP additions for each cycle are listed in Extended Data Table 1.

Reagent addition (termination). This block is the same as previous reagent addition blocks except that ATP, CTP, GTP, and UTP are added. Incubation (termination) is the same as incubation (elongation) except that the incubation time is 12 min.

SPE (termination) and rinsing (termination). These blocks are the same as the SPE (initiation) and rinsing (initiation) blocks, respectively, except that during SPE in the termination step, the liquid removed from the reaction vessels is collected in sample containers rather than being pumped to the waste container.

31. Korbie, D. J. & Mattick, J. S. Touchdown PCR for increased specificity and sensitivity in PCR amplification. Nature Protocols 3, 1452–1456 (2008).
32. Kao, C., Zheng, M. & Rüdisser, S. A simple and efficient method to reduce nontemplated nucleotide addition at the 3 terminus of RNAs transcribed by T7 RNA polymerase. RNA 5, 1268–1272 (1999).
33. Mukherjee, S., Brieba, L. G. & Sousa, R. Structural transitions mediating transcription initiation by T7 RNA polymerase. Cell 110, 81–91 (2002).
34. Zuo, X. et al. Solution structure of the cap-independent translational enhancer and ribosome-binding element in the 3′ UTR of turnip crinkle virus. Proc. Natl Acad. Sci. USA 107, 1385–1390 (2010).
35. Ying, J. et al. Measurement of $^1$H,$^15$N and $^{13}$C residual dipolar couplings in nucleic acids from TROSY intensities. J. Biomol. NMR 51, 89–103 (2011).
36. Delaglio, F. et al. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J. Biomol. NMR 6, 277–293 (1995).
37. Holmstrom, E. D. & Nesbitt, D. J. Single-molecule fluorescence resonance energy transfer studies of the human telomerase RNA pseudoknot: temperature-/urea-dependent folding kinetics and thermodynamics. J. Phys. Chem. B 118, 3853–3863 (2014).
38. Kapanidis, A. N. et al. Alternating-laser excitation of single molecules. Acc. Chem. Res. 38, 523–533 (2005).
39. Aitken, C. E., Marshall, R. A. & Puglisi, J. D. An oxygen scavenging system for improvement of dye stability in single-molecule fluorescence experiments. Biophys. J. 94, 1826–1835 (2008).
40. Cordes, T., Vogelsang, J. & Tinnefeld, P. On the mechanism of Trolox as antiblinking and antibleaching reagent. J. Am. Chem. Soc. 131, 5018–5019 (2009).
41. Fiore, J. L., Hodak, J. H., Pietsert, O., Downey, C. D. & Nesbitt, D. J. Monovalent and divalent promoted GAAA tetraloop-receptor tertiary interactions from freely diffusing single-molecule studies. Biophys. J. 95, 3892–3905 (2008).
42. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. Acta Crystallogr. D 66, 486–501 (2010).
43. Afonine, P. V. et al. Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr. D 68, 352–367 (2012).
44. Steinberg, G., Stromsberg, K., Thomas, L., Barker, D. & Zhao, C. Strategies for covalent attachment of DNA to beads. Biopolymers 73, 597–605 (2004).
45. Chivers, C. E. et al. A streptavidin variant with slower biotin dissociation and increased mechanostability. Nature Methods 7, 391–393 (2010).
Extended Data Figure 1 | Synthesis algorithm for the selective labelling of RNA using the PLOR method. All samples except for fully labelled 71-nt-CN, S1+Lk1-CN, S1+Lk1-H, and 15N-104-nt-TCV were synthesized following this algorithm. The initiation, elongation, and termination stages are shown in green, blue, and red, respectively. Various NTP combinations added during the elongation cycles depend on the desired labelling scheme.
Extended Data Figure 2 | Automated platform for PLOR synthesis.

a, Diagram of the automated platform for PLOR synthesis depicting its various parts and stations. b, Top-view (left) and side-view (right) photographs of the in-house automated platform.
Extended Data Figure 3 | Optimization of experimental conditions for PLOR synthesis. a, The products of the PLOR (left) and the standard transcription methods (right). PLOR generates a pure full-length product with the desired labelling in the final step. b, Comparison of PLOR efficiency for Lp2-CN synthesis under various conditions: I, freshness of DNA-attached beads; II, anion specificity; III, [Mg²⁺]; IV, K₂SO₄ presence in the initiation; V, increasing NTPs (1X represents NTP amounts in Extended Data Table 1); VI, ratio of ATP/GTP:DNA; VII, ratio of UTP:DNA; and VIII, incubation time of the termination. The right lane contains pure RiboA71 as a control.
Extended Data Figure 4 | Comparison of NMR spectra of PLOR-generated and 71-nt-CN samples in adenine-bound form. Superposition of the $^1$H$^{13}$C-TROSY spectra of 71-nt-CN with Lp1-CN (a), Lp2-CN (b), Lp1+2-CN (c), Lk2-CN (d), S1+Lk1-CN (e), and 4nt-CN (f). These results indicate that the RNAs synthesized by PLOR have the same fold as 71-nt-CN generated using the traditional solution-based transcription method and are functional as evidenced by binding of the adenine ligand.
Extended Data Figure 5 | PLOR-generated riboA71 maintains both sequence and structural fidelity. a, Structural superposition of the PLOR-generated riboA71 (PDB accession number 4XNR) with that of the riboA71 generated using the regular in vitro transcription (PDB accession number 4TZX). The root mean squared deviation between all C1 atoms was 0.3 Å. b, Structural superposition of the PLOR-generated riboA71 (PDB accession number 4XNR) with that of the riboA71 (PDB accession number 1Y26). The root mean squared deviation between all C1 atoms was 0.4 Å. c, Sequences and secondary structures of the RNAs in b. The arrows denote nucleotide sequence differences between the two ribA71 sequences. d, Composite simulated anneal-omit 2|Fo| − |Fc| electron density of the riboA71 RNA structure (PDB accession number 4XNR) at 2.2 Å resolution calculated using the final model (1.0 s.d.). e, Portion of the electron density in c in the adenine-binding pocket unambiguously identifies the nucleobase identities of the PLOR-generated RNA, revealing no undesired sequence changes introduced by PLOR. f, Portion of the electron density in c of the G6·C66 base pair, which differs in identity between the two structures, suggesting that if there had been undesired sequence substitutions that resulted from using the PLOR method, they would have been readily detected in the crystallographic analysis.
Extended Data Figure 6 | Using PLOR to isotopically label a 104-nt RNA.

a, Secondary structure of the 104-nt structural element in the TCV genomic RNA. The $^{15}$N-H1-TCV was synthesized by the PLOR method, and the sequence in red is $^{15}$N-labelled in $^{15}$N-H1-TCV. 
b, $^{15}$N-TROSY spectrum of $^{15}$N-H1-TCV RNA.
c, Superposition of the $^1$H$^1$N-TROSY spectra of $^{15}$N-H1-TCV with $^{15}$N-104-nt-TCV$^{34}$, indicating that the PLOR-generated selectively labelled RNA has the same fold as that generated using the regular in vitro transcription.
**Extended Data Table 1 | PLOR recipes for the RiboA71 and TCV samples**

| Sample      | NTP addition                                                                 | \( C_{\text{DNA/Vol.}} \) |
|-------------|------------------------------------------------------------------------------|-----------------------------|
| Lp1-CN      | Initiation: 96ATP, 96GTP, 9.6UTP; Elongation (cycle 1): 2ATP, 2CTP, 2UTP;     | 20 µM, 40 mL                |
|             | (cycle 2): 13\(^{15}\)N-ATP, 13\(^{15}\)N-GTP; (cycle 3): 13\(^{15}\)N-ATP,    |                             |
|             | 2\(^{15}\)N-U-UTP; (cycle 4): 2\(^{15}\)N-GTP; Termination (cycle 5): 10ATP,  |                             |
|             | 11CTP, 7GTP, 17UTP                                                           |                             |
| Lp2-CN      | Elongation (cycle 2): 3ATP, 7GTP, 8UTP; (cycle 3): 3ATP, 3CTP, UTP; (cycle 4): | 20 µM, 50 mL                |
|             | ATP, 2GTP; (cycle 5): 2\(^{15}\)N-CTP, 2\(^{15}\)N-U-UTP; (cycle 6): 3\(^{15}\)N-ATP; |                             |
|             | (cycle 7): 2CTP, GTP, 3UTP; Termination (cycle 8): 2ATP, 4CTP, 5UTP           |                             |
| Lp1+2-CN    | Elongation (cycle 2): 13\(^{15}\)N-ATP, 13\(^{15}\)N-GTP; (cycle 3): 13\(^{15}\)N- | 20 µM, 50 mL                |
|             | ATP, 2\(^{15}\)N-U-UTP; (cycle 4): 2\(^{15}\)N-GTP; (cycle 5): ATP, 4GTP, 8UTP; |                             |
|             | (cycle 6): 3ATP, 3CTP, UTP; ATP, 2GTP; (cycle 8): 2\(^{15}\)N-CTP, 2\(^{15}\)N-U- |                             |
|             | UT; (cycle 9): 3\(^{15}\)N-ATP; (cycle 10): 2CTP, GTP, 3UTP; Termination      |                             |
|             | (cycle 11): 2ATP, 4CTP, 5UTP                                                 |                             |
| Lk2-CN      | Elongation (cycle 2): ATP, GTP; (cycle 3): ATP, 2UTP; (cycle 4): 5GTP, 3UTP; | 20 µM, 40 mL                |
|             | (cycle 5): ATP; (cycle 6): 11\(^{15}\)N-CTP, 12\(^{15}\)N-GTP, 4\(^{15}\)N-U |                             |
|             | UT; ATP, 2\(^{15}\)N-U-UTP; (cycle 7): 13\(^{15}\)N-ATP; (cycle 8): 2\(^{15}\)N- |                             |
|             | 2\(^{15}\)N-CTP; Termination (cycle 9): 8ATP, 8GTP, 3GTP, 10UTP               |                             |
| S1+Lk1-CN   | Initiation: 96\(^{15}\)N-ATP, 96\(^{15}\)N-GTP, 9.6\(^{15}\)N-U-UTP; Elongation/ | 20 µM, 20 mL                |
|             | Termination (cycle 1): 14ATP, 13CTP, 10GTP, 21UTP                            |                             |
| 4m-CN       | Elongation (cycle 2): 13\(^{15}\)N-ATP, GTP; (cycle 3): 2ATP, 6GTP, 8UTP;    | 20 µM, 40 mL                |
|             | (cycle 4): 13\(^{15}\)N-CTP, 13\(^{15}\)N-U-UTP; (cycle 5): 4ATP, 2CTP, 2GTP; |                             |
|             | (cycle 6): 3ATP, 2CTP, 5UTP; (cycle 7): 13\(^{15}\)N-GTP; Termination (cycle 8): |                             |
|             | 2ATP, 4CTP, 5UTP                                                             |                             |
| S1+Lk1-H    | Initiation: 96ATP, 96GTP, 9.6U TP; Elongation/Termination (cycle 1): 14\(^{15}\) | 20 µM, 20 mL                |
|             | H-ATP, 13\(^{15}\)H-CTP, 10\(^{15}\)H-GTP, 2\(^{15}\)H-UTP                      |                             |
| U39-CN      | Elongation (cycle 2): 3ATP, 7GTP, 8UTP; (cycle 3): 3ATP, 3CTP, 12\(^{15}\)N- | 20 µM, 30 mL                |
|             | U-UTP; (cycle 4): ATP, 2CTP, 2GTP; (cycle 5): 3ATP, 2CTP, 5UTP; (cycle 6): 2ATP, |                             |
|             | 2GTP, 3UTP; Termination (cycle 7): 4CTP, 2UTP                                |                             |
| RiboA71     | Elongation (cycle 2): 3ATP, 7GTP, 8UTP; (cycle 3): 3ATP, 3CTP, UTP; (cycle 4): | 20 µM, 30 mL                |
| (4XR)       | ATP, 2GTP; (cycle 5): 2CTP, 2UTP; (cycle 6): 3ATP, (cycle 7): 2CTP, GTP, 3UTP; |                             |
|             | Termination (cycle 8): 2ATP, 4CTP, 5UTP                                      |                             |
| \(^{15}\)N-H1-TCV | Initiation: 32\(^{15}\)N-ATP, 240\(^{15}\)N-GTP, 16\(^{15}\)N-CTP; Elongation | 20 µM, 30 mL                |
|             | (cycle 1): \(^{15}\)N-CTP, \(^{15}\)N-GTP, \(^{15}\)N-U-UTP; Elongation (cycle 2): |                             |
|             | 2ATP, 19CTP, 25GTP, 18UTP                                                    |                             |

*The initiation and elongation (cycle 1) information are not listed if they are the same as those used in the synthesis of Lp1-CN.

†The template concentrations and total reaction volumes were listed for NMR and FRET sample synthesis.
Extended Data Table 2 | Synthesis efficiencies for the PLOR-generated NMR samples

| Sample    | Number of Elongation/termination cycles (n) | Efficiency * | Average Efficiency | Overall Yield |
|-----------|---------------------------------------------|--------------|--------------------|---------------|
|           |                                             | I  | E   | I  | E   | Calculated † | Experimental |
| Lp1-CN    | 5                                           | 43.3% | 93.1% | 24% | 30.3% |
| Lp2-CN    | 8                                           | 49.0% | 82.2% | 16.2% | 10.2% |
| Lp1+2-CN  | 11                                          | 47.0% | 85.6% | 11% | 8.5% |
| Lk2-CN    | 9                                           | 44.6% | 90.2% | 14.3% | 17.6% |
| S1+Lk1-CN | 1                                           | N/A  | N/A  | 40.4% | 40.3% |
| 4nt-CN    | 8                                           | 47.8% | 84.3% | 16.2% | 12.2% |
| S1+Lk1-H  | 1                                           | N/A  | N/A  | 40.4% | 38.5% |
| US9-CN    | 7                                           | 44.2% | 91.1% | 18.5% | 23.0% |
| 15N-H1-TCV| 2                                           | N/A  | N/A  | 35.5% ‡ | 32.0% |

*Efficiencies, I and E, for each sample were calculated using the equation Yield = I × E^n, where I and E are the initiation and elongation efficiencies, respectively, and n is the total number of cycles in the elongation and termination phases (see previous text on the calculations of synthesis efficiencies).

†These yields represent the theoretical yields for each sample based on average efficiencies, I and E, and the number of cycles.

‡The yield of 15N-H1-TCV is calculated by using I and E from the riboA71 system.
Extended Data Table 3 | PLOR recipes for the fluorescent-labelled RiboA71 samples

| Sample | NTP addition | [DNA], total vol. |
|--------|--------------|------------------|
| U24Cy3-C55Cy5 | Elongation (cycle 2): ATP, GTP; (cycle 3): UTP; (cycle 4): ATP, 5-aminoallyl-UTP; (cycle 5): ATP, 6GTP, 6UTP; (cycle 6): 3ATP, 3CTP, UTP; (cycle 7): ATP, 2GTP; (cycle 8): 2CTP, 2UTP; (cycle 9, 37°C): 3ATP, Cy5-CTP; (cycle 10): CTP, GTP, 3UTP; Termination (cycle 11): 2ATP, 4CTP, 6UTP | 5 μM, 3 mL |
| U24Cy3-C55Cy5-B | Elongation (cycle 2): ATP, GTP, (cycle 3): UTP; (cycle 4): ATP, 5-aminoallyl-UTP; (cycle 5): ATP, 6GTP, 6UTP; (cycle 6): 3ATP, 3CTP, UTP; (cycle 7): ATP, 2GTP; (cycle 8): 2CTP, 2UTP; (cycle 9, 37°C): 3ATP, Cy5-CTP; (cycle 10): CTP, GTP, 3UTP; (cycle 11): 2ATP, 3UTP; (cycle 12): CTP; Termination (cycle 13): 2UTP, 3 biotin-11-CTP | 5 μM, 3 mL |
| U24Cy3-U65Cy5 | Elongation (cycle 2): ATP, GTP; (cycle 3): UTP; (cycle 4): ATP, 5-aminoallyl-UTP; (cycle 5): ATP, 6GTP, 6UTP; (cycle 6): 3ATP, 3CTP, UTP; (cycle 7): ATP, 2GTP; (cycle 8): 2CTP, 2UTP; (cycle 9): ATP, GTP; (cycle 10): 2UTP; (cycle 11, 37°C): ATP, Cy5-CTP; Termination (cycle 12): 4CTP, 2UTP | 5 μM, 3 mL |
| U24Cy3-U65Cy5-B | Elongation (cycle 2): ATP, GTP; (cycle 3): UTP, (cycle 4): ATP, 5-aminoallyl-UTP; (cycle 5): ATP, 6GTP, 6UTP; (cycle 6): 3ATP, 3CTP, UTP; (cycle 7): ATP, 2GTP; (cycle 8): 2CTP, 2UTP; (cycle 9): ATP, GTP; (cycle 10): 2UTP; (cycle 11, 37°C): ATP, Cy5-CTP; Termination (cycle 13): 2UTP, 3 biotin-11-CTP | 5 μM, 3 mL |
| U24A555-U65A488-B | Elongation (cycle 2): ATP, GTP, (cycle 3): UTP, (cycle 4): ATP, 5-aminoallyl-UTP; (cycle 5): ATP, 6GTP, 6UTP; (cycle 6): 3ATP, 3CTP, UTP; (cycle 7): ATP, 2GTP; (cycle 8): 2CTP, 2UTP; (cycle 9): ATP, GTP; (cycle 10): 2UTP; (cycle 11, 37°C): ATP, Alexa488-UTP; (cycle 12): CTP; Termination (cycle 13): 2UTP, 3 biotin-11-CTP | 5 μM, 1 mL |

*Cy3- and Alexa Fluor 555-labelled samples were generated by adding Cy3-NHS ester (GE Healthcare) or Alexa Fluor 555-NHS ester (Invitrogen), respectively, to a solution containing the lyophilized transcript product dissolved in double-distilled H2O, followed by the addition of 0.5 volumes of 0.3 M sodium bicarbonate (pH 8.3). The reactions were incubated in the dark with gentle mixing for 1 h at 37 °C, 12 h at 25 °C, and then 1 h at 37 °C.

†The template concentrations and total reaction volumes were listed for NMR and FRET sample synthesis.
Extended Data Table 4 | Crystallographic statistics of data collection and refinement for the PLOR-generated riboA71 structure (PDB accession number 4XNR)

| **PDB Accession Code** | 4XNR |
|------------------------|------|
| **Data collection**    |      |
| Space group            | P2₁,2₁,2 |
| Cell dimensions        |      |
| a, b, c (Å)            | 49.5 154.7 25.2 |
| αβγ (°)                | 90, 90, 90 |
| Resolution (Å)         | 41.68-2.207 (2.286-2.207)* |
| Rmerge (%)             | 6.1 (103.7) |
| I/σI                   | 16.9 (1.7) |
| CC1/2                  | 0.999 (0.578) |
| Completeness (%)       | 100 (99) |
| Redundancy             | 6.6 (6.7) |
| **Refinement**         |      |
| Resolution (Å)         | 41.68-2.207 (2.286-2.207)* |
| No. reflections        | 10356 (1014) |
| Rmerge/Rfree (%)       | 20.4 (36.6)/22.0 (37.1) |
| No. atoms              | 1562 |
| RNA                    | 1502 |
| Protein                | n/a |
| Ion                    | 6 |
| Water                  | 64 |
| Mean B-factors (Å²)    | 65.8 |
| RNA                    | 66.0 |
| Protein                | n/a |
| Ligand/Ion             | 79.3 |
| Water                  | 57.8 |
| R.m.s. deviations      |      |
| Bond lengths (Å)       | 0.001 |
| Bond angles (°)        | 0.27 |

*Highest resolution shell in parenthesis.