Supplementary Information

Solvent Viscosity Facilitates Replication and Ribozyme Catalysis from an RNA Duplex in a Model Prebiotic Process

Christine He^{1,a}, Adriana Lozoya-Colinas^{2,a}, Isaac Gállego², Martha A. Grover¹, and Nicholas V. Hud2,*

¹ School of Chemical & Biomolecular Engineering, Georgia Institute of Technology, Atlanta, GA, 30332-0100, United States

 2 School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, GA, 30332-0400, United States

^a The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors

* To whom correspondence should be addressed. Tel: +1 (404) 385-1162; Email: hud@chemistry.gatech.edu

Present Address: Christine He, Innovative Genomics Institute, University of California, Berkeley, 2151 Berkeley Way, Berkeley, CA, 94704, United States

Isaac Gállego, Medical Research Council Lab of Molecular Biology, Francis Crick Avenue, Cambridge Biomedical Campus, Cambridge CB2 0QH, UK. 01223 267000

This file includes:

Supplementary Materials and Methods DNA/RNA sequences Supplementary Figures S1-S5 Supplementary Tables S1-S3 Supplementary References

Supplementary Materials and Methods

Preparing RNA samples in glycholine. Glycholine was prepared from recrystallized choline chloride and glycerol as described previously (1). All oligonucleotides were purchased from Integrated DNA Technologies or GE Dharmacon, and resuspended in 18.2 MΩ/cm water (Barnstead Nanopure™). For preparation and transfer of viscous samples, we used mass rather than volume as a unit to express the amount of glycholine used, and hence we express our concentrations in molal (*m*). To prepare RNA samples in glycholine, the appropriate amount of aqueous RNA stock was added to a weighed amount of glycholine and mixed until homogeneous. Water was then removed by vacuum centrifugation for at least 12 hours. All RNA samples in aqueous buffer contain 0.1 M NaCl, 20 mM Tris, pH 7.5.

In vitro transcription of RNA templates. The sequences for all RNA templates were taken from the plasmid pBluescript II SK(-), a phagemid excised from lambda ZAPII (GenBank accession: X52330.1). This plasmid was chosen as a model of a mixed, arbitrary sequence (i.e., not designed to have a particular sequence or base pair composition). A 43 nt *trans*-acting hammerhead ribozyme enzyme sequence was placed into the middle of the RNA template sequences. All sequences are listed below in DNA/RNA Sequences, with the hammerhead ribozyme region highlighted.

The 650 DNA template for in vitro transcription (IVT_gblock) was ordered as a gBlock from IDT. PCR was performed (NEB Q5 High Fidelity DNA polymerase) using two sets of primers (IVT_S_FWD and IVT_S_REV, and IVT_A_FWD and IVT_A_REV) to produce separate dsDNA templates for in vitro transcription of sense and anti-sense RNA template strands. After PCR, amplicons were purified by PAGE in native conditions (0.5x TBE, 7.5% acrylamide). We refer to these amplicons as d96_S through d613_S, and d96_A through d613_A.

Double stranded RNA templates for replication were produced by co-transcription of both the sense and anti-sense DNA constructs in a one-pot reaction. 750 ng each of the sense and antisense plasmid constructs (for example, S-d96_HHRmid and AS-d96_HHRmid) were mixed and transcribed using the HiScribe T7 High Yield RNA Synthesis Kit (New England Biolabs). The transcription reaction was run at 37 °C overnight. After transcription, DNA templates were digested with Turbo DNase (ThermoFisher Scientific). Double stranded RNA templates were purified by PAGE in native conditions (0.5x TBE, 7.5% acrylamide), visualized by UV shadowing, and purified from the gel slab. Finally, the RNA templates were purified by ethanol precipitation.

Thermal stability analysis of r604_HHRmid. UV absorbance was used to measure the thermal stability of a r604 duplex template (604 bp RNA duplex with the HH sequence near the end). Samples containing 900 μ *m* of the r604 duplex in glycholine and aqueous buffer were prepared. UV measurements were performed on 1 mm quartz cuvettes in a temperature-controlled UV-Vis spectrophotometer (Cary 50 UV-Vis spectrophotometer). Heating and cooling traces were generated for each sample by recording spectra (220-350 nm) from 20 to 90 °C at intervals of 1 °C. The hybridization state of the r604 template was measured using the absorbance at 260 nm (Figure S1). Melting temperature (T_m) values (Figure S1, dashed lines) were estimated as described previously (2). Consistent with previous studies of DNA/RNA in viscous eutectic solvents (1,35), glycholine depresses the thermal stability of r604_HHR compared to aqueous buffer ($T_{m, glycoline}$) $= 47$ versus °C vs. T_{m, water} = 57 °C).

Estimated secondary structure formation on r613 template strands. The Mfold web server (6) was used to estimate the relative stability of secondary structure formation on the sense vs. antisense strand (Figure S3). Mfold only makes predictions for aqueous buffer, and is not without limitations in aqueous buffer. Thus, Mfold was considered the best method for providing a qualitative estimate of the relative stability of secondary structures on the template strands. The Mfold predictions estimate that overall structure formation on the sense template strand is 12 kcal/mol more favorable than the anti-sense template strand (by difference of predicted ΔG values). This prediction is consistent with our observation that S4' binding to the sense template strand is significantly lower than S4 binding to the anti-sense template strand (Figure 2D).

Cleavage assay by the HH ribozyme. Samples were prepared as 2 mg aliquots of glycholine containing either the minimal HH ribozyme or the template at a concentration of 2×10^{-6} molal and FAM-labeled substrate in a 4:1 molar ratio with the minimal HH ribozyme or the template. Samples were then heated to 80ºC for 2 min and cooled to 20ºC at a constant rate over 20 min. RNA samples were diluted with water, and 10x T4 RNA ligase 2 buffer was added (to a concentration of 1x) for a final reaction volume of 20 µl. **T4 RNA ligase 2 buffer was selected as the buffer for the cleavage assays to maintain the same conditions used for the RNA ligation reactions. Specifically, at 1x concentration the T4 RNA ligase 2 buffer contains 2 mM MgCl2 and 50 mM Tris (pH 7.5), 1 mM DTT, 400 µM ATP.** After incubating for 1 h at 20ºC, the reaction was quenched by adding EDTA to a final concentration of 2.5mM and adjusting pH to 5 with a sodium acetate buffer. Then, 6 µl of each reaction mixture was combined with 6 µl of 2x loading dye. The resulting sample was heated to 85ºC for 2 min and then placed on ice before loading into a 20% denaturing polyacrylamide gel (8 M urea, 1x TBE buffer). Prior to sample loading, gels were pre-run for >30 min at 14 W and 50 V/cm. After sample loading, gels were run at 20 W for around 1 h and imaged for FAM fluorescence before staining with SYBR Gold.

DNA/RNA sequences

HH ribozyme: GCAGGUACAUCCAGCUGAUGAGUCCCAAAUAGGACGAAAUGCC HH ribozyme substrate: /56-FAM/GGCAUCCUGGAUUCCACUGC

1 GCAUC UUUUA CUUUC ACCAG CGUUU CUGAG UUGCU CUU<mark>GC AGGUA CAUCC AGCUG</mark>

Template sequences with HH ribozyme sequence highlighted in each.

r96 template (sense strand):

Figure S1. Heating and cooling trace of r604 template in aqueous buffer and glycholine. T_m values are marked by dashed lines. The difference between the heating and cooling traces in glycholine is consistent with the trapping of the strands of r604 in their single stranded state.

 Figure S2. (Left) Full images of agarose gels from which portions are shown in Figure 2B. Densitometry analysis was used to quantify the binding of fluorescently tagged oligonucleotides to the RNA template in Figure 2B. Values used for the plot shown as Figure 2C in main text are listed in Table S1. (Right) Gel image of control experiments with double and single stranded template, the labeled oligonucleotides used, and the 2-hour timepoint from the gel shown at left to demonstrate that the oligonucleotides bind to the single stranded template.

A

Figure S3. Predicted secondary structure formation within the r604 sense (A) and anti-sense (B) strands. Listed below each structure prediction is the predicted ΔG of folding (kcal/mol) based on program parameters of 37 °C with 0.1 M NaCl salt.

Figure S4. Replication of RNA templates in different viscous solvents. Samples containing the r613 template duplex and oligonucleotides S1-S10 and S1'-S10' were heated and cooled, and then

ligated. Glycholine refers to glycerol and choline chloride in a 4:1 molar ratio, while reline refers to urea and choline chloride in a 2:1 molar ratio.

Figure S5. Denaturing polyacrylamide gel showing the results of heating, cooling, and ligating the r613 template duplex, oligonucleotides S1-S10 and S1'-S10', and a mixture of oligonucleotide substrates with random sequences (32 bp each) in glycholine. The Cy3 and Cy5 images track synthesis of the RNA sense and anti-sense strands, respectively.

Figure S6. Densitometry analysis of FAM-labeled HH substrate for Lanes 1 and 6 of gel shown as Figure 4A in main text. The inset of the bottom panel is a zoom-in of the cleaved HH substrate (7 nt) peak.

Figure S7. Densitometry analysis of FAM-labeled HH substrate in Lanes 1 and 4 of gel shown as Figure 4B in main text. The inset of the bottom panel is a zoom-in of the Cleaved HH substrate (7 nt) peak.

Figure S8. Replicate of experiments shown in Figures 4A and 4B. (Left) A denaturing polyacrylamide gel shows that an RNA duplex can cleave the HH substrate. (Right) Cleavage of the HH substrate is observed from products of viscosity-mediated replication with r613 and oligonucleotides S1-S10 and S1'-S10'.

Figure S9. Hammerhead cleavage by the sense r613 strand synthesized from the anti-sense r613 strand using oligonucleotide substrates S1-S10 in glycholine. PAGE analysis confirms that the newly synthesized sense strand exhibits hammerhead cleavage.

Supplementary Tables

Table S1: Densitometry analysis of template-bound fluorescence for Cy3S4 and Cy5S4' (values and standard deviations are based on three identical experiments)

Table S2: Densitometry analysis of all replication products from r613 template (values and standard deviations are based on three identical experiments)

Table S3: Yield of maximum length replication products from r96, r128, r192, and r613 templates (values are averages from two identical experiments for each template)

Supplementary References

- 1. He, C., Gallego, I., Laughlin, B., Grover, M.A. and Hud, N.V. (2017) A viscous solvent enables information transfer from gene-length nucleic acids in a model prebiotic replication cycle. *Nat. Chem.*, **9**, 318-324.
- 2. Mergny, J.L. and Lacroix, L. (2003) Analysis of thermal melting curves. *Oligonucleotides*, **13**, 515-537.
- 3. Mamajanov, I., Engelhart, A.E., Bean, H.D. and Hud, N.V. (2010) DNA and RNA in anhydrous media: duplex, triplex, and G-quadruplex secondary structures in a deep eutectic solvent. *Angew. Chem. Int. Ed.*, **49**, 6310-6314.
- 4. Lannan, F.M., Mamajanov, I. and Hud, N.V. (2012) Human telomere sequence DNA in water-free and high-viscosity solvents: G-quadruplex folding governed by Kramers rate theory. *J. Am. Chem. Soc.*, **134**, 15324-15330.
- 5. Gállego, I., Grover, M.A. and Hud, N.V. (2015) Folding and imaging of DNA nanostructures in anhydrous and hydrated deep-eutectic solvents. *Angew. Chem. Int. Ed.*, **54**, 6765-6769.
- 6. Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31**, 3406-3415.