Supplementary Information

# A simple physical mechanism enables homeostasis in primitive cells

Aaron E. Engelhart, Katarzyna Adamala, and Jack W. Szostak*

*correspondence to: szostak@molbio.mgh.harvard.edu

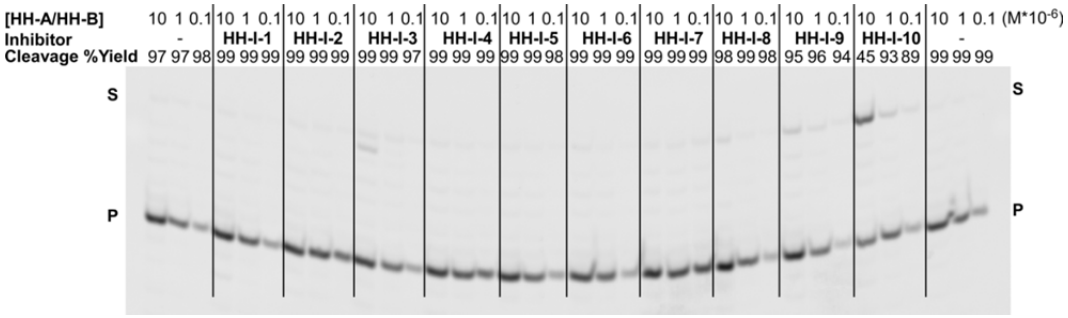
**Supplementary Information Figure 8.** Control ethanol precipitation demonstrating identical recovery efficiency for **HH-A** and its ribozyme cleavage product.

**Supplementary Information Figure 9.** Electropherogram showing **HH-A** and cleavage product alongside RNA size marker.

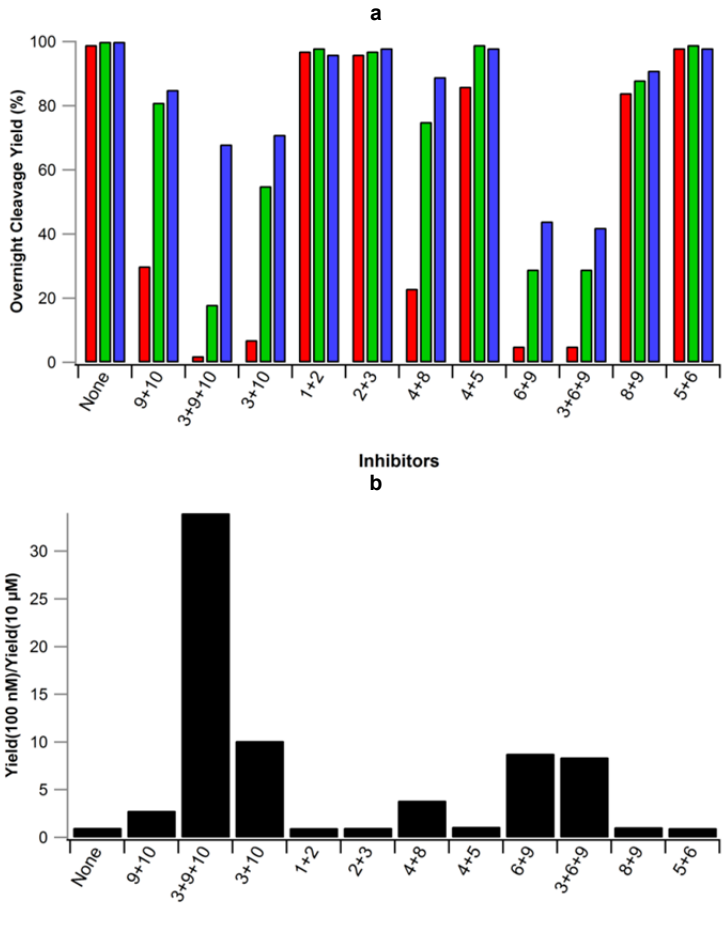
**Supplementary Information Figure 10.** Repurification of vesicles demonstrating oligonucleotide contents are retained on the experimental timescale.

| Name | Sequence (shown aligned to HH-B) | Complementary Residues in HH-B | $\Delta G_{298K}$ (kcal/mol) | $\log(K_{Dcalc\ 298K})$ |
|---|---|---|---|---|
| **HH-I-1** | 5′-UCAGU-3′<br>3′-GCGCGGAGUAGUCAGCUCGG-5′ | 7-11 | -4.22 | -3.1 |
| **HH-I-2** | 5′-**UG**U**C**U**U**-3′<br>3′-GCGCGGAGUAGUCAGCUCGG-5′ | 13-18 | -4.54 | -3.3 |
| **HH-I-3** | 5′-UCGAG-3′<br>3′-GCGCGGAGUAGUCAGCUCGG-5′ | 3-7 | -5.21 | -3.8 |
| **HH-I-4** | 5′-AUCAGU-3′<br>3′-GCGCGGAGUAGUCAGCUCGG-5′ | 7-12 | -5.71 | -4.2 |
| **HH-I-5** | 5′-**UG**C**U**UC-3′<br>3′-GCGCGGAGUAGUCAGCUCGG-5′ | 13-18 | -5.92 | -4.4 |
| **HH-I-6** | 5′-CUCAUC-3′<br>3′-GCGCGGAGUAGUCAGCUCGG-5′ | 10-15 | -7.04 | -5.2 |
| **HH-I-7** | 5′-AGUCGA-3′<br>3′-GCGCGGAGUAGUCAGCUCGG-5′ | 4-9 | -7.31 | -5.4 |
| **HH-I-8** | 5′-GCCUC-3′<br>3′-GCGCGGAGUAGUCAGCUCGG-5′ | 13-17 | -7.90 | -5.8 |
| **HH-I-9** | 5′-CGCGC-3′<br>3′-GCGCGGAGUAGUCAGCUCGG-5′ | 16-20 | -8.34 | -6.1 |
| **HH-I-10** | 5′-UCAUCAG-3′<br>3′-GCGCGGAGUAGUCAGCUCGG-5′ | 8-14 | -8.92 | -6.6 |
| **HH-A-5pStem** | 5′-CGCGCC-3′<br>3′-GCGCGGAGUAGUCAGCUCGG-5′ | 15-20 | -12.06 | -8.9 |
| **HH-A-3pStem** | 5′-UCGAGC-3′<br>3′-GCGCGGAGUAGUCAGCUCGG-5′ | 3-8 | -9.17 | -6.8 |

**Supplementary Information Table 1. Predicted thermodynamics for association of** HH-B **with inhibitor sequences and** HH-A **fragments.** The above series of 5-7 nt partial complements to **HH-B** were selected to represent a range of binding affinities for **HH-B** (ca. 250 nM-800 μM $K_{Dcalc}$). In some cases, inhibitors form G-U mismatches when hybridized to **HH-B**; the U residues involved are shown in bold. Affinities of **HH-I-1** through **HH-I-10**, **HH-A-5pStem**, and **HH-A-3pStem** for **HH-B** were calculated using MELTING 5.1.0[18,19].



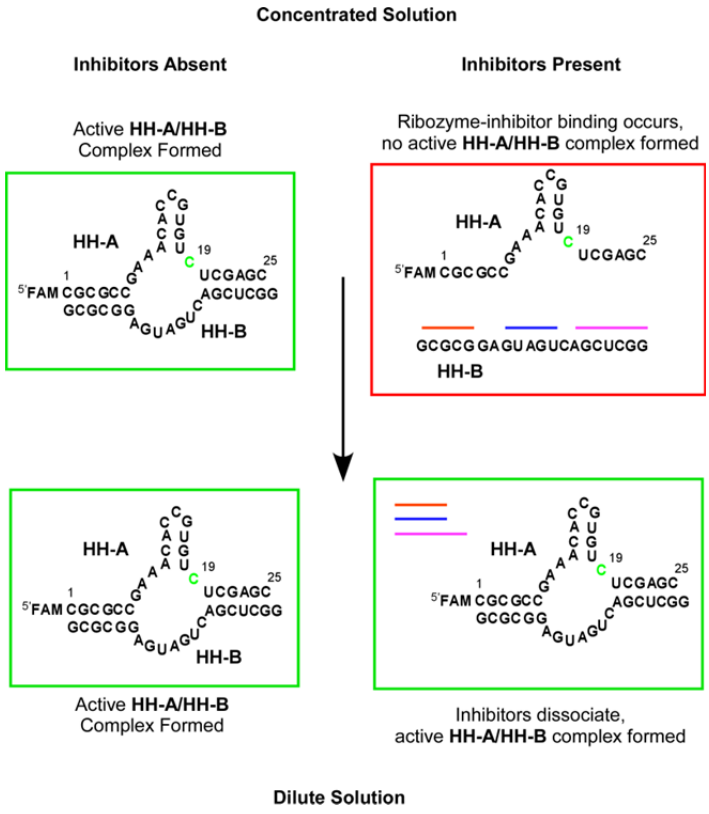

**Supplementary Information Figure 1. Concentration dependence of cleavage of** HH-A/HH-B-**inhibitor complexes**. Reactions were incubated overnight in 100mM sodium HEPES pH 7.5, 10 mM $MgCl_2$, 1 mM sodium EDTA pH 8, at the specified concentration of **HH-A** and **HH-B** (10 μM, 1 μM, or 100 nM), with 100 eq. of inhibitor (i.e., 1 mM, 100 μM, or 10 μM) present. Substrate (**HH-A**) is denoted as **S;** cleavage product is denoted as **P**. Lines separating reaction series do not denote separate gels; rather, they are intended to guide the eye.




**c**

| [HH] \ Inhibitors | 9+10 | 3+9+10 | 3+10 | 1+2 | 2+3 | 4+8 | 4+5 | 6+9 | 3+6+9 | 8+9 | 5+6 | None |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 10 µM | 30 | 2 | 7 | 97 | 96 | 23 | 86 | 5 | 5 | 84 | 98 | 99 |
| 1 µM | 81 | 18 | 55 | 98 | 97 | 75 | 99 | 29 | 29 | 88 | 99 | 100 |
| 100 nM | 85 | 68 | 71 | 96 | 98 | 89 | 98 | 44 | 42 | 91 | 98 | 100 |
| Yield(100 nM)/ Yield(10 µM) | 2.8 | 34 | 10 | 1.0 | 1.0 | 3.9 | 1.1 | 8.8 | 8.4 | 1.1 | 1.0 | 1.0 |

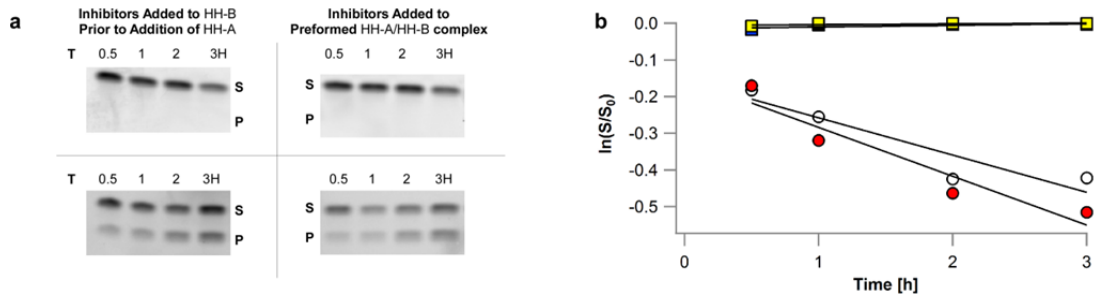
**Supplementary Information Figure 2. Cleavage yields of** HH-A/HH-B**-inhibitor mixtures.** Reactions were incubated overnight in 100 mM sodium HEPES pH 7.5, 10 mM $MgCl_2$, 1 mM sodium EDTA pH 8, at the specified concentration of both **HH-A** and **HH-B**, with 100 eq. (i.e., 1 mM, 100 µM, or 10 µM) of each inhibitor strand noted. **a)** Cleavage yields vs. **HH-A/HH-B** concentration: red=10 µM, green=1 µM, blue=100 nM. **b)** Ratio of 10 µM to 100nM cleavage yields. **c)** Data in panels **a** and **b** in tabular format.




**Supplementary Information Figure 3. Schematic illustration of effects of dilution on ribozyme assembly in the absence (left) and presence (right) of inhibitors** HH-I-3/9/10. In the absence of inhibitors, the catalytically active **HH-A-HH-B** complex is stable over the concentration range studied and is not impacted greatly by dilution from 10 μM to 0.1 μM. (Left) **HH-B,** when at high concentration and in the presence of an excess of complementary inhibitors, is sequestered in a complex with these strands; it does not participate in the **HH-A-HH-B** complex. (Right, top) Upon dilution, the weakly-binding inhibitors dissociate from **HH-B,** allowing it to interact with **HH-A,** forming the catalytically active **HH-A-HH-B** complex. (Right, bottom) The cytosine residue 5′ to the cleavage site is shown in red; the 5′-fluorescein end label on **HH-A** is denoted with a green star.

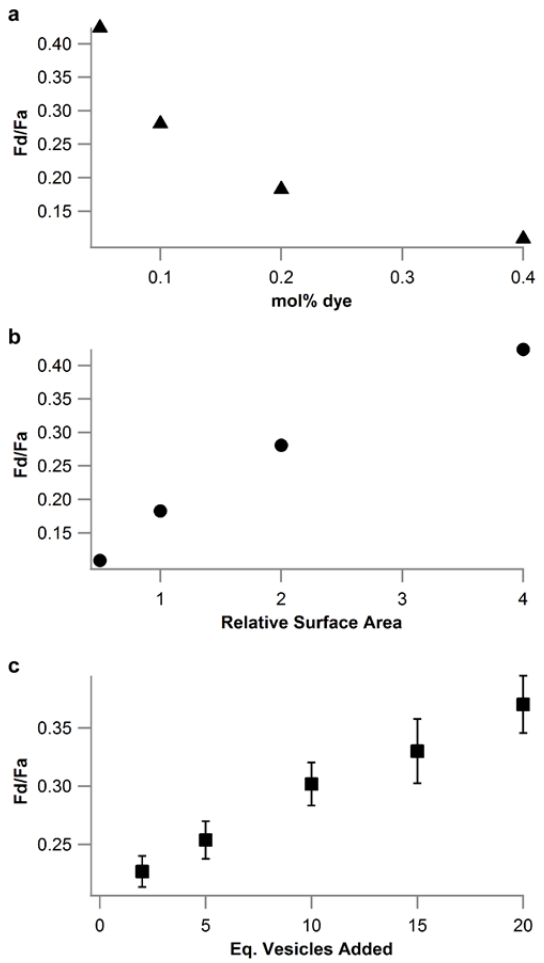
| Partial Complements Present | Apparent rate constant ($hr^{-1}$) 10 μM **HH-A/HH-B** | 1 μM **HH-A/HH-B** | 0.1 μM **HH-A/HH-B** | 1 μM **HH-A/HH-B** In Vesicles No Growth | 1 μM **HH-A/HH-B** In Vesicles After Growth+Volume Equilibration | $d([\mathbf{P}])/dt_{t=0,\ post\text{-}growth}$ / $d([\mathbf{P}])/dt_{t=0,\ pre\text{-}growth}$ (**P**=ribozyme product) |
|---|---|---|---|---|---|---|
| None | 0.65±0.02 | 0.52±0.03 | 0.32±0.03 | 0.23±0.02 | 0.11±0.02 | 0.096 |
| 100 eq. **HH-I-3** | 0.48±0.03 | 0.26±0.02 | 0.21±0.03 | | | |
| 100 eq. **HH-I-9** | 0.24±0.01 | 0.21±0.01 | 0.19±0.02 | | | |
| 100 eq. **HH-I-10** | 0.030±0.001 | 0.07±0.003 | 0.19±0.001 | | | |
| 100 eq. each **HH-I-3/9/10** | <0.005 | 0.016±0.003 | 0.17±0.01 | 0.026±0.008 | 0.12±0.004 | 0.92 |
| 300 eq. **r($N_6$)** | <0.005 | 0.071±0.016 | 0.10±0.05 | 0.027±0.002 | 0.099±0.005 | 0.73 |
| 200 eq. **r($N_6$)** | 0.010±0.003 | 0.089±0.004 | 0.19±0.01 | | | |
| 100 eq. **r($N_6$)** | 0.098±0.005 | 0.18±0.003 | 0.21±0.004 | | | |
| 30 eq. **r($N_6$)** | 0.33±0.02 | 0.28±0.008 | 0.27±0.02 | | | |
| 200 eq. **r($N_6$)** + 100 eq. **r($N_5$)** | 0.0057±0.001 | 0.059±0.001 | 0.17±0.007 | 0.043±0.004 | 0.11±0.004 | 0.51 |
| 300 eq. **r($N_5$)** | 0.067±0.002 | 0.12±0.01 | 0.21±0.03 | | | |

**Supplementary Information Table 2. Cleavage rates for hammerhead ribozyme systems in presence and absence of inhibitor oligonucleotides.** Errors for encapsulated systems without inhibitor, with **HH-I-3/9/10,** and with 300 eq. **r($N_6$)** are S.E.M., N=4; encapsulated systems with 200 eq. **r($N_6$)**, 100 eq. **r($N_5$)** are extreme values, N=2. Errors for unencapsulated systems without inhibitor, with **HH-I-3/9/10,** and with 300 eq. **r($N_6$)** are S.E.M, N=3; Errors for other unencapsulated single-inhibitor systems and random inhibitor systems are S.D. from linear fit of $\ln(S/S_0)$ vs. time in hours. Rates were determined over 2-4 h time intervals.

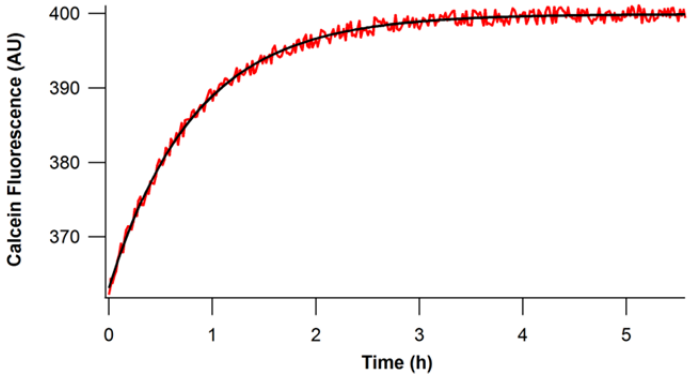



**Supplementary Information Figure 4. Inactive Inhibitor**-HH-B **complexes capable of activation by dilution are formed regardless of order of addition.** In typical reactions, **HH-B** and inhibitors were mixed with all reaction components except **HH-A** and $MgCl_2$, resulting in formation of a **HH-B**-inhibitor complex; **HH-A** was then added, and a portion of the reaction containing 10 μM **HH-A/HH-B** was diluted to produce 0.1 μM samples, then $MgCl_2$ was added to each sample, resulting in a ribozyme that was inactive at 10 μM **HH-A/HH-B** concentration and active at 0.1 μM **HH-A/HH-B** concentration**.** Electropherograms are shown at left in panel **a**. Substrate (**HH-A**) is denoted as **S**, cleavage product is denoted as **P**. Kinetic plots are shown in panel **b.** Calculated $k_{\mathbf{HH\text{-}A=HH\text{-}B}=(10\ \mu M)} < 0.005\ h^{-1}$ (blue squares, overlapped by preformed **HH-A-HH-B** complex series described below, denoted by yellow squares), $k_{\mathbf{HH\text{-}A=HH\text{-}B}=(0.1\ \mu M)} = 0.10 \pm 0.03\ h^{-1}$ (white circles) (N=1, error is SD from curve fit).

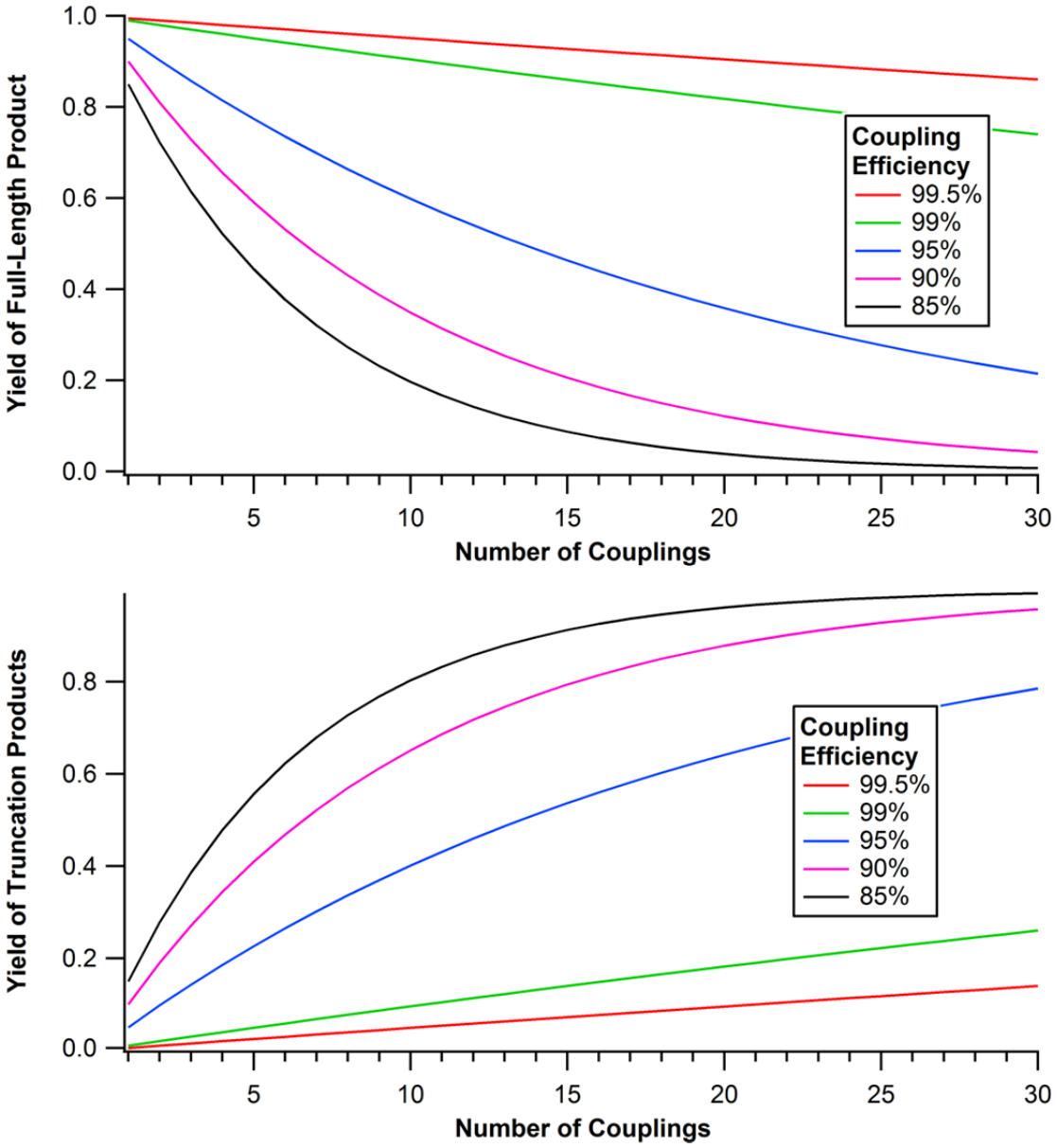
In order to demonstrate that even preformed **HH-A-HH-B** complexes could form inhibitor complexes, **HH-A** and **HH-B** were mixed with all reaction components except **HH-I-3/9/10** and $MgCl_2$. **HH-I-3/9/10** were then added, diluted (if applicable), and $MgCl_2$ was added. Electropherograms are shown at left in panel **a**. Substrate (**HH-A**) is denoted as **S**, cleavage product is denoted as **P**. Kinetic plots are shown in panel **b**. Calculated $k_{\mathbf{HH\text{-}A=HH\text{-}B}=(10\ \mu M)} < 0.005\ h^{-1}$ (yellow squares), $k_{\mathbf{HH\text{-}A=HH\text{-}B}=(0.1\ \mu M)} = 0.13 \pm 0.03\ h^{-1}$ (red circles) (N=1, error is SD from curve fit).




**Supplementary Information Figure 5. Calibration curve and growth of MA/GMO/DOPA vesicles in the presence of MA/GMO vesicles.** A calibration curve was generated for 100 nm vesicles containing lipid FRET dyes (panel **a**) and converted to a calibration curve for final relative surface area relative to initial surface area (panel **b**). 10% DOPA MA/GMO vesicles containing 0.2% FRET dyes were then mixed with pure MA/GMO vesicles and growth was monitored by FRET (panel **c**). The addition of 20 equivalents of pure MA/GMO vesicles to 10% DOPA MA/GMO vesicles resulted in a change of Fd/Fa equivalent to a ca. 3-fold increase in surface area, or a ca. 5-fold increase in volume, given spherical vesicles (error bars represent extreme values, N=2).




**Supplementary Information Figure 6. Volume growth of MA/GMO/DOPA vesicles.** MA/GMO/DOPA vesicles (100 mM total lipid concentration) in 250 mM tris-HCl, pH 8, containing 20 mM calcein and purified from unencapsulated calcein by size-exclusion chromatography, were mixed with 20 eq. MA/GMO vesicles of the same total lipid concentration. Volume equilibration of these vesicles was monitored by observing the fluorescence signal ($\lambda_{ex}$=415 nm, $\lambda_{em}$=600 nm) associated with encapsulated calcein dequenching as buffer and water equilibrated across the enlarged membrane. The rate of fluorescence recovery was single-exponential, with $k_{app}$=1.26±0.02 $h^{-1}$ (error represents extreme values, N=2).



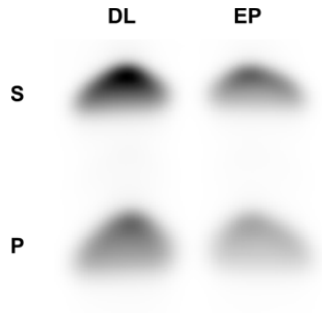

**Supplementary Information Figure 7. Yields of full-length RNA products (a) and truncation products (b) predicted based on stepwise coupling efficiency.** Even at 99.5% coupling efficiency, only 86% full-length product is produced in a 30-coupling reaction (Panel **a**). At 99% coupling efficiency, 73% full-length product is produced; at 95% coupling efficiency, 20% full-length product is produced; at 90% coupling efficiency, 3.8% full-length product is produced, and at 85% coupling efficiency, only 0.6% full-length product is produced.

Conversely, even at fairly high coupling efficiencies (95%), truncation products are the predominant reaction products in stepwise RNA synthesis beyond 14 nt (Panel **b**). At this coupling efficiency, a 30-coupling RNA synthesis will produce 79% truncation products. At 85% coupling efficiency, the synthesis of RNA larger than 5 nt will produce mostly truncation products. At this coupling efficiency, a 30-coupling RNA synthesis will produce >99% truncation products.

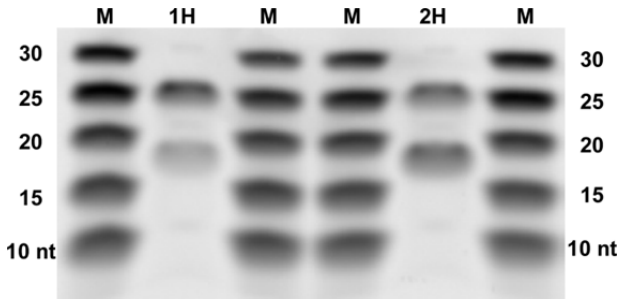
While the above data represent a simple arithmetic model of stepwise synthesis, the lower range of rates is comparable to or higher than those of the most highly evolved ribozyme polymerases known; the tC19 polymerase exhibits a termination rate of 7% (i.e., 93% extension efficiency), and the R18 polymerase exhibits a termination rate of 40% (60% coupling efficiency). tC19 can elongate a primer by 91+nt with 0.035% full-length yield, corresponding to 92% extension efficiency. Consistent with this, polymerization reactions catalyzed by these enzymes produce large amounts of truncation products relative to full-length product[13, 14]. Additionally, it is likely that nonspecific RNA degradation would contribute to the pool of short oligonucleotides, given that unconstrained RNA linkages exhibit a rate of transesterification-based degradation 2-3 orders of magnitude higher than those found in structured RNAs[15].

| Stepwise Efficiency | 30-coupling yield | 1-10 coupling yield over 30 couplings | [Short]/[Full Length] | [Short] if reaction generates 10μM full-length product (μM) |
|---|---|---|---|---|
| 99.5% | 86% | 4.9% | 0.06 | 0.6 |
| 99% | 73% | 9.6% | 0.13 | 1.3 |
| 95% | 20% | 40% | 2.0 | 20 |
| 90% | 3.8% | 65% | 17 | 170 |
| 85% | 0.6% | 80% | 120 | 1200 |

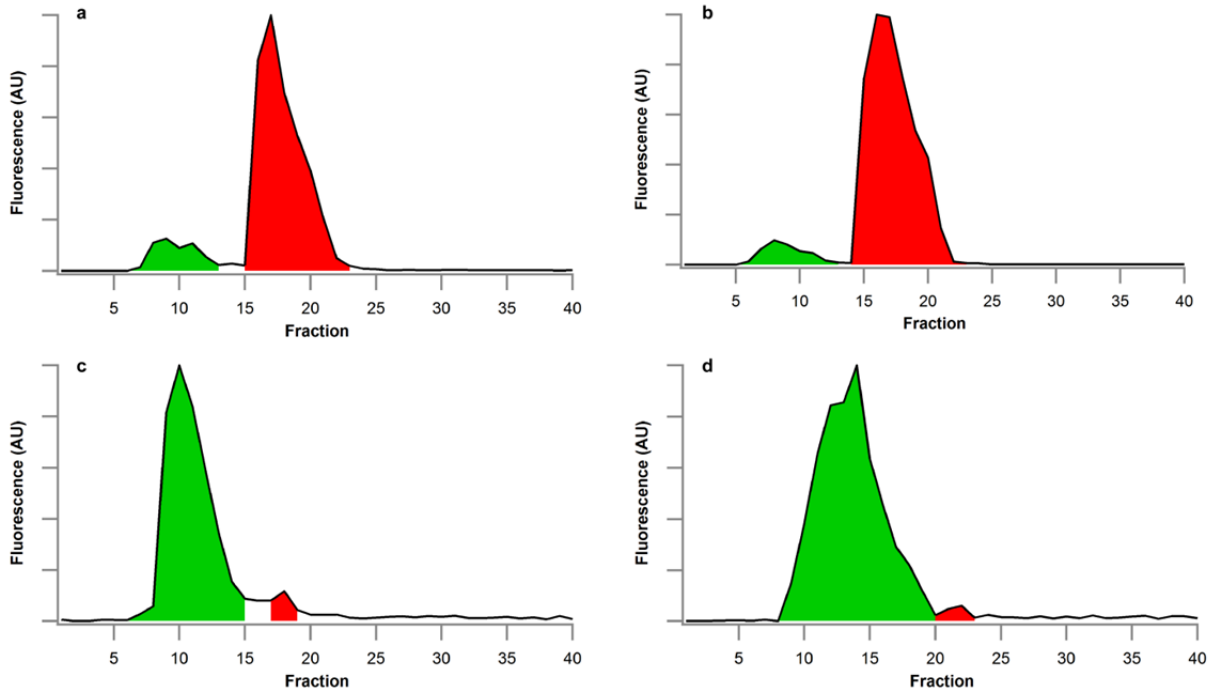
**Supplementary Information Table 3. Relative yields and concentrations of full-length and short RNAs predicted based on stepwise coupling efficiency.** Even at very high coupling efficiencies (ca. 95% and below), more short truncation products (1-10 couplings, i.e., 2-11 nt) are generated than full-length products (30 couplings, i.e., 31 nt). At more modest, but still high (85%), coupling efficiencies, short oligomers are expected to be produced in large (>100-fold) excess.




**Supplementary Information Figure 8. Control ethanol precipitation demonstrating identical recovery of HH-A and its ribozyme cleavage product.** Representative unencapsulated hammerhead cleavage reactions (10 μM **HH-A/HH-B,** 250 mM tris-HCl pH 8, 4 mM $MgCl_2$, 10 μL reaction volume) after 65 min reaction time were subjected to a) dilution into 10 μL loading buffer and loading 10 μL of the resulting mixture directly onto the gel (lane labeled **DL**), or b) mixing with 300 μL MA/GMO/DOPA vesicles (60 mM/30 mM/10 mM), ethanol precipitation and washing, followed by resuspension in 12 μL loading buffer and loading 10 μL onto the gel (lane labeled **EP**). The ribozyme cleavage product yield (ribozyme cleavage product labeled **P**, uncleaved **HH-A** ribozyme substrate labeled **S**) was identical in both lanes (46%).




**Supplementary Information Figure 9. Electropherogram showing HH-A and cleavage product alongside RNA size marker ladder.** A representative hammerhead cleavage reaction (10 μM **HH-A/HH-B,** 250 mM tris-HCl pH 8, 4 mM $MgCl_2$) was stopped at 1 H and 2 H by addition to 8 M urea, 1 X TBE containing 20 mM supplemental EDTA. Fluorescein end-labeled products (lanes labeled **1H** and **2H**) were analyzed by PAGE alongside a size marker ladder consisting of fluorescein-labeled 10, 15, 20, 25, and 30 nt ssRNA oligonucleotides (lanes labeled **M**), demonstrating the cleavage product is of the expected length (19 nt). Imaging was performed using the fluorescein channel of a Typhoon gel scanner as described in the Methods section. All inhibitors used in this study were of shorter length (5-7 nt) than the smallest band in the size marker ladder (10 nt).




**Supplementary Information Figure 10. Repurification of vesicles demonstrating oligonucleotide contents are retained on the experimental timescale.** Initial Sephadex 4B purification traces (monitored by fluorescein emission) of MA/GMO/DOPA vesicles containing 10 μM **HH-A** (**a**) or 100 μM **FAM-DNA-HH-I-3** (**b**) and repurification traces of the same vesicles mixed with 20 eq. MA/GMO vesicles, tumbled for 3 H, allowing for > 90-95% volume relaxation (Supplementary Information Figure 6), 4 mM $MgCl_2$ added, then tumbled for an additional 4 H (**c** for **HH-A**, **d** for **FAM-DNA-HH-I-3**). The repurification trace demonstrates ≥ 90-95% retention of both oligonucleotides on this timescale. In each trace, the green shaded region denotes the encapsulated oligonucleotide fraction, and the red shaded region denotes the unencapsulated oligonucleotide fraction.