Supplementary Information for ‘In-ice evolution of RNA polymerase ribozyme activity’

Authors: James Attwater, Aniela Wochner and Philipp Holliger

Table of Contents

Supplementary Materials and Methods

CBT library generation and manipulation by StEP. A library of R18 variants was generated by mutagenic PCR (6 cycles of 94°C for 30 s, 50°C for 30 s, 72°C for 120 s) using primer 5'-biotin-biotin-GATCGAGATCTCGATCCCGCGAAATTAATACG ACTCACTATAGGACAACC-3' and a ten-fold excess of primer 5'-GGAGCCGAAGCT CC-3' in the presence of 400 μM each of dATP, dGTP, dCTP, dUTP, 8-oxo-2'-deoxyguanosine-5'-triphosphate (TriLink) and 2'-deoxy-P-nucleoside-5'-triphosphate (TriLink), and the products were bound to beads to deplete the wild-type sequences by elution in 0.1 M NaOH, and amplified to generate ribozyme genes for selection with positions mutagenised at an average rate of 4.4% (mostly transitions). Recombination is thought to be a critical process during RNA evolution[33], and we introduced such opportunities through StEP shuffling[34] of libraries during selection. Hot-start PCR was used (f.c.: 1 μM each primer, 0.1 M tetramethyl ammonium chloride, 1x Taq buffer, 50 μM each dNTP, 0.1 U/μl SuperTaq (HT Biotechnology), 0.3 pmol PCR product in 20 μl; 120 cycles of 94°C for 30 s, 65°C for 1 s), and genes were further amplified by standard PCR. In test runs, the resulting library contained 10% parent genes and 90% recombination products (with an estimated 3% crossing-over chance per backbone position and 0.3-1% mutation rate as determined by sequencing).

Sequencing of full-length extension products. Comparative fidelity was assessed by sequencing complete BioFITCU10-A extension products synthesized by tC19Y (Supplementary Table S3) on the templates described in Supplementary Table S2, using primer/template separation before gel resolution as described above for the reactions in Supplementary Fig. S8a avoiding the use of competing oligonucleotide contaminating the sequence pool. These bands were excised from the gel, and RNA molecules were eluted, precipitated, and poly-A tailed and reverse transcribed, before cloning and sequencing as described previously[11].

Sequencing of total extension products. BioloopA10 primers were extended upon template I-10 by tC9Y under standard conditions (17˚C, 120 h), and were then heated (95˚C 20 min) in 6 M urea with CompI-19 to separate extended primers and templates. The primers were then bound to MyOne Streptavidin C1 microbeads, heated again in denaturing buffer with CompI-19, poly-A tailed, and reverse-transcribed (SuperScript III Reverse Trancriptase, Invitrogen); the resulting cDNA

was released from the beads with RNaseH, amplified and endowed with terminal tags to allow high-throughput sequencing using a MiSeq (Illumina).

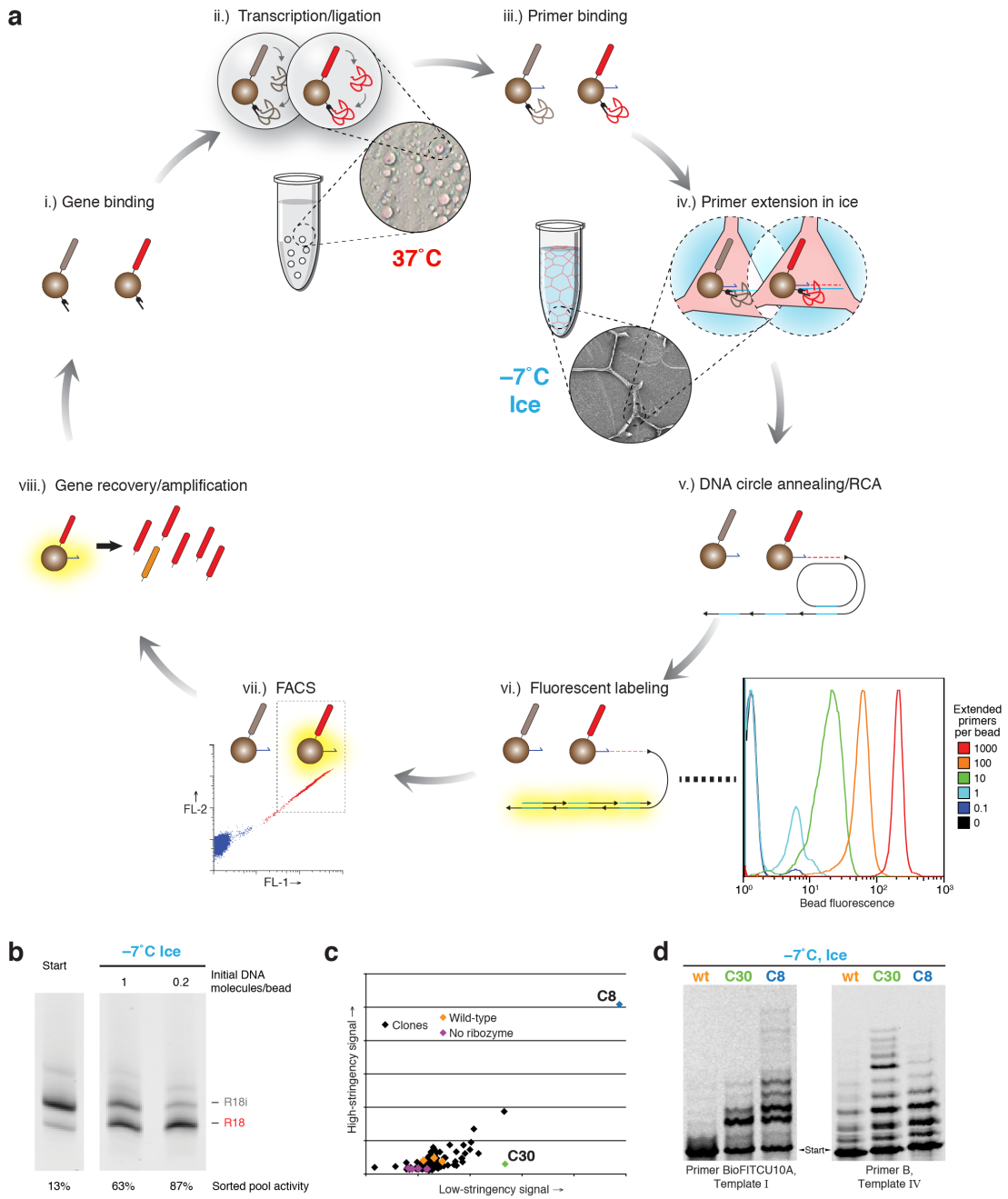
The resulting data were quality filtered only to remove sequences with ambiguous calls, and trimmed to remove tag, primer and poly-A sequences, before assignment of deletions, yielding 1.17×10^5 positions sequenced. The template comprises ten 11-base repeats, and separate tallies of errors and successful incorporations were constructed for sequences of different lengths according to the repeat number reached, allowing investigation of biases in length-selection sequencing (Supplementary Fig. S7a). A further separate tally of errors and successful incorporations in the penultimate and final base of each sequence was assembled (see Fig. 4d). Sequences ending in the one or two correct A incorporations at the relevant position in each repeat would see them removed during poly-A sequence trimming, so such resulting sequences were evenly assigned between the +0A, +1A and +2A possibilities. There was no evidence of ribozyme-mediated insertions, as a very similar level of insertions (~0.025%) was observed in an Illumina-sequenced control RNA prepared similarly, which also yielded a total 0.51% background rate of errors, the individual components of which were subtracted from those in the extension data to reveal ribozyme-induced error rates. Sequencing of both the extension products and a mock RNA full-length extension product yielded a low rate of long (3-10 nt) deletions; longer deletions cannot be identified due to the repeating nature of the template, which may have contributed to deletion emergence during the rtPCRs. Only the first and last residues in contiguous stretches of deleted residues are included in ribozyme fidelity analysis; the intervening residues are omitted from analysis, as they are likely to not reflect ribozyme-induced deletion tendencies.

The overall error spectrum (Fig. 4c) represents an average of the data in each length bin (Supplementary Fig. S7a), weighted by the abundance of bin sequences seen by denaturing PAGE and the average length of sequences within each bin (Fig. 4a) (overall average length of extended primers on template I-10: 26.81 nt). The averaged fidelity of the final two bases in each sequence was calculated similarly without weighting for sequence length. Overall fidelities were calculated as geometric means of the fidelities for each of the four bases. The chance of each error type being extended (Supplementary Fig. S7d) was estimated by first subtracting the internal position error rates from the final two base error rates (as not all errors in the final two would have been responsible for termination), and using the resulting rate to

calculate the expected number of terminating errors in the final two bases of an average extension product - allowing comparison to the expected number of internal errors across the whole average extension product (26.81 nt length). These two numbers represent errors that cause termination and errors that don't, allowing a simple estimate of the termination probability to be obtained for each.

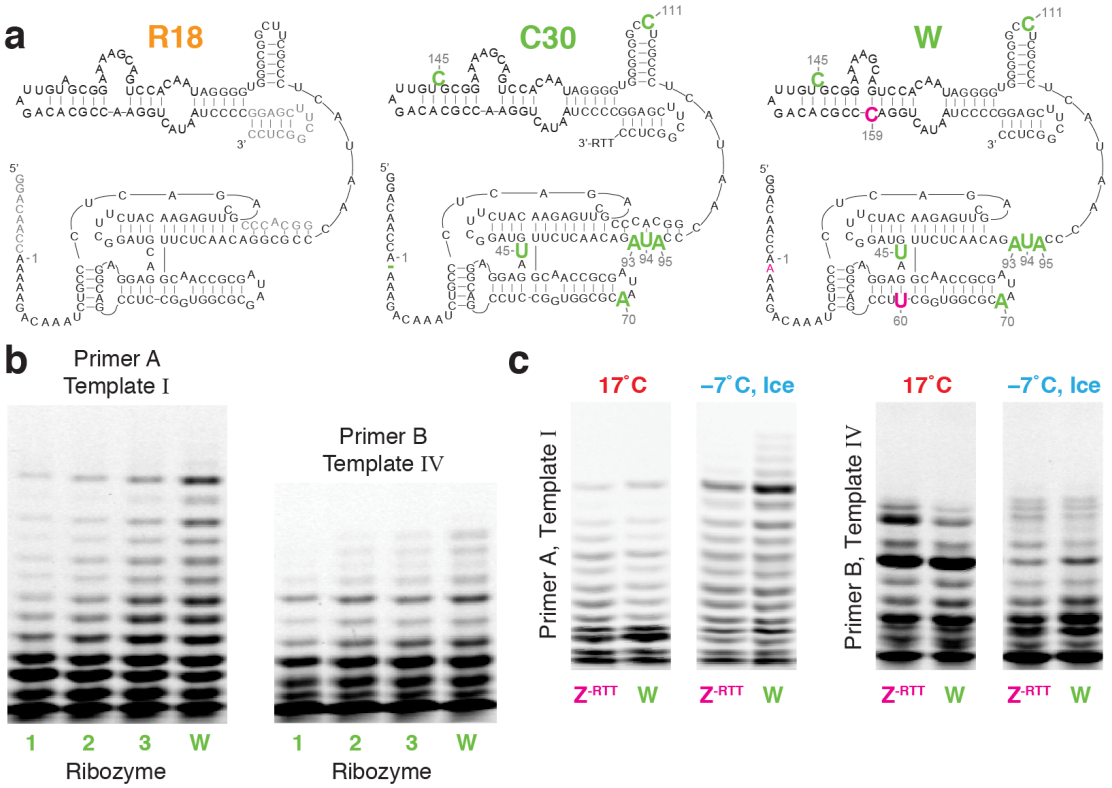
Template selection. The first two rounds of the template selection in ice were carried out as at 17˚C essentially as described previously[11], except that extension reactions contained only 1 mM of each NTP, and reactions were frozen at −25˚C before incubation at −7˚C as normal (Round 1: 423 h incubation; Round 2: 452 h incubation). The output pool from the second round of selection was then subjected to mutagenic PCR to introduce further variation: 5 cycles of [94˚C 30 s, 45˚C 30 s, 72˚C 120 s], followed by 15 cycles of [94˚C 30 s, 50˚C 30 s, 72˚C 120 s], using 0.5 μM each of primers U10D and T7C19, in the presence of 400 μM each of dATP, dGTP, dCTP, dUTP as well as mutagenic dNTPs: 8-oxo-2'-deoxyguanosine-5'-triphosphate (TriLink) and 2'-deoxy-P-nucleoside-5'-triphosphate (TriLink), followed by 'recovery' PCR with only standard dNTPs. The resulting RNA template library was used in two half-scale selection extensions in ice, one using 0.2 M $MgCl_2$, and one using 0.2 M $MgSO_4$ (742 h incubation at −7˚C). As no clear differences were observed in the outputs of these two rounds, they were combined and used in a final half-sized selection extension (0.2 M $MgSO_4$, 140 h, −7˚C). Purification and recovery for rounds 3 and 4 used a similar protocol to round 2 but using non-mutagenic PCR. To examine the progress of the selection, the output pools were cloned and sequenced, or transcribed and gel-purified to allow assessment of pool template activity (Supplementary Fig. S8a).

Supplementary Figures

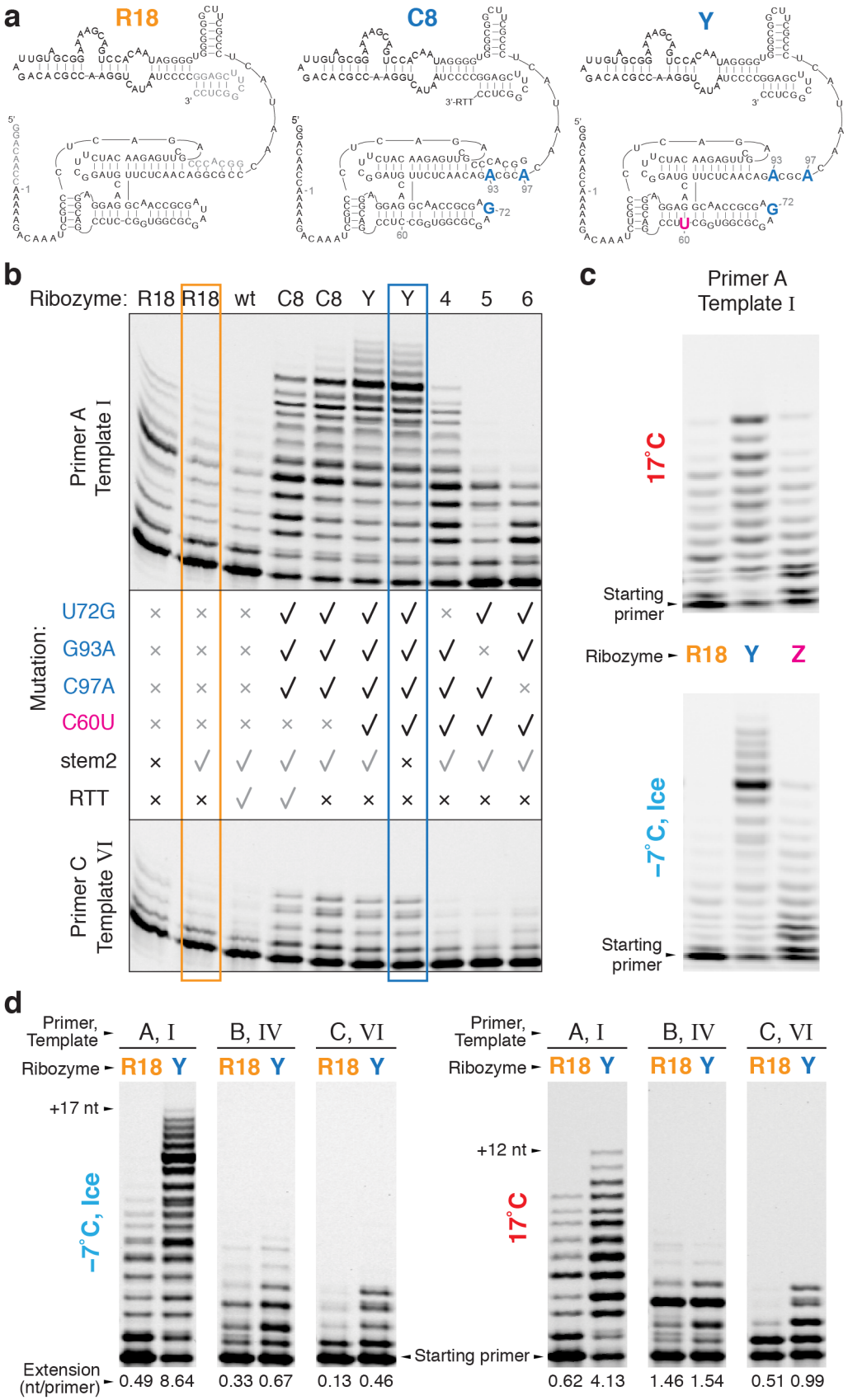


Supplementary Figure S1 I In-ice selection strategy and screening. a, CBT (compartmentalized bead-tagging) selection[11] was modified to allow the selection of polymerase ribozyme activity in ice. (i) A library of ribozyme genes is bound to streptavidin-coated microbeads at a density of up to one gene per bead. (ii) Ribozymes are transcribed and ligated to bead-bound RNA hairpins inside the aqueous compartments of a water-in-oil emulsion (inset diameter ~0.15 mm),

generating clonal bead-bound ribozyme populations. (iii) Clonal ribozyme beads are recovered from emulsion and further decorated with primer/template duplexes after which the beads are frozen (at −25˚C) in diluted ribozyme extension buffer and incubated at −7˚C to allow eutectic phase formation (iv). Ribozymes remain ligated to the beads during the incubation, enhancing ribozyme activity by promoting primer extension *in cis*. (v) After thawing and bead recovery, extended primers are used to prime rolling circle amplification (RCA) on a DNA minicircle labelling beads with DNA concatemers. (vi) These are converted to double-stranded DNA and stained with Picogreen®, allowing detection of even single extended primers. (vii) Fluorescent beads can then be isolated using flow cytometry. (viii) The genes bound to such beads, encoding ice-active ribozymes, are then recovered using PCR. Inset panel (vi): Histograms of fluorescence of single-bead gated FACS events after RCA [minicircle: DNAcirc+5], deriving from bead populations bound to varying densities of BioU10-Aext, mimicking fully-extended RNA primers. **b,** Test selection of in-ice CBT (132 hours, +7 minicircle; 10% PAGE, SYBR-Gold stained) using R18 genes mixed with genes encoding an inactivated R18 variant with a 21-nucleotide insertion (R18i[11]), and bound to beads at different densities prior to selection. Sorted pool activity is described by the percentage of R18 genes in the recovery PCR pool. Lower gene densities on beads resulted in higher recovery, by reducing stochastic 'hitchhiking' of inactive genes bound to the same bead. **c,** Scatter plot of relative ribozyme polymerase activity as determined by RPA after eight rounds of in-ice selection, using high-stringency (P3) and low-stringency (P2) probes. **d,** Denaturing PAGE of primer extension by selected ribozyme clones C30 and C8 and wild-type (wt) on the selection primer/template duplex sequence (BioFITCU10A/I) and an unrelated primer/template duplex (B/IV) (−7˚C ice, 0.5 mM each NTP, 162 hours).

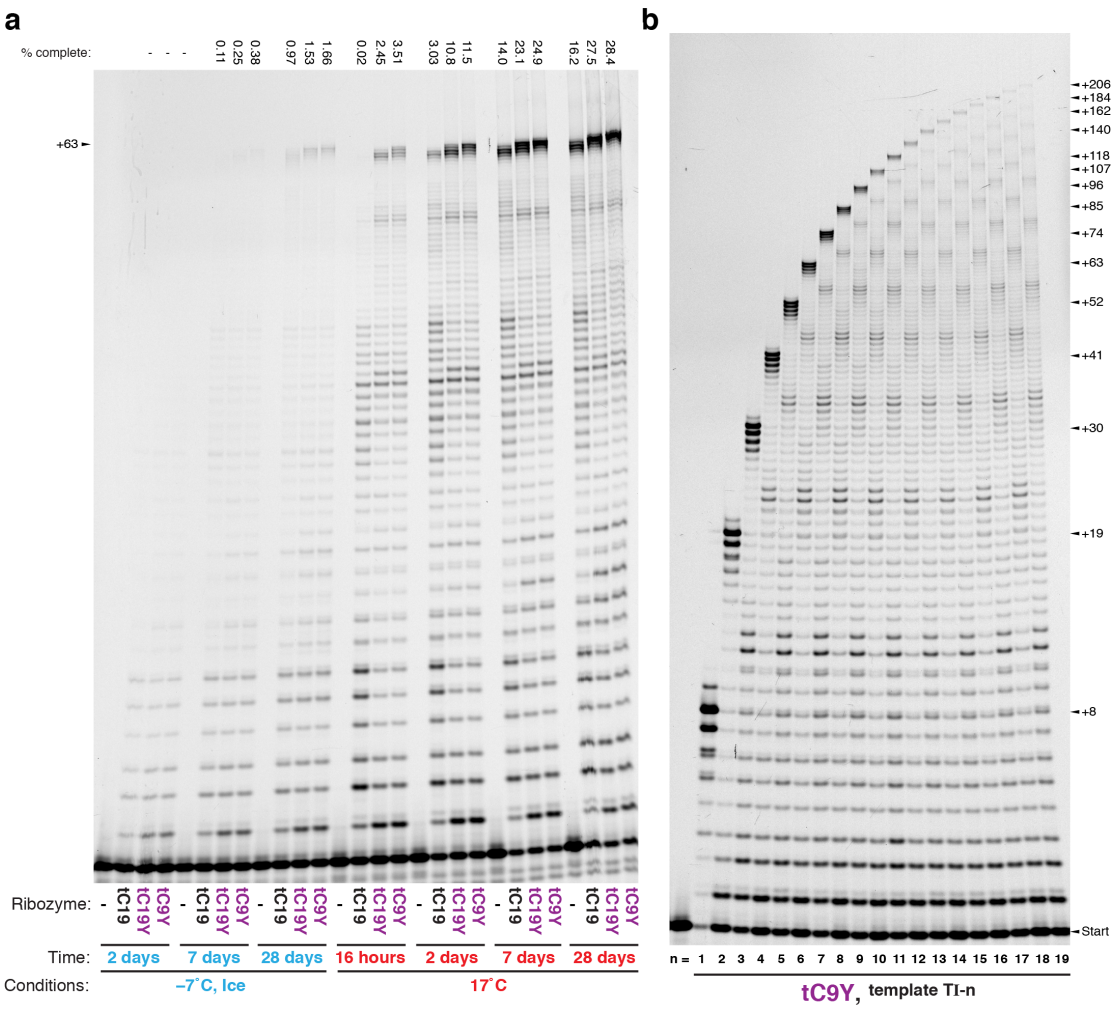


Supplementary Figure S2 I Engineering of the W ribozyme. a, Secondary structures of the wild-type R18 ribozyme construct (residues not mutagenized in the starting library are shown in grey); the isolated C30 ribozyme (ice-selected mutations in green); and W (mutations derived from the previously described ribozyme Z in magenta) (3'-RTT: a 3' run-through transcript sequence present in the selection). Two of the eight mutations in C30 (G93A, G95A) had previously been identified in Z. **b,** Reverting a deletion (A2) in C30 and including the two other mutations (C60U, A159C) from Z yielded the ribozyme 'W': shown are primer extension activities of ribozymes C30, W and intermediates (urea PAGE, –7˚C in ice, 6.5 days). 1: C30, no 3'-RTT; 2: C30 (A159C), no 3'-RTT; 3: C30 (-Δ2A, A159C), no 3'-RTT; W: C30 (-Δ2A, C60U, A159C), no 3'-RTT. **c,** Comparison of primer extension activity of ribozymes Z^{-RTT} and W at ambient temperature and in ice at –7˚C (8 days). W shares four mutations (C60U, G93A, G95A, A159C) with Z^{-RTT} but outperforms it in ice due to the presence of five more ice-selected mutations (C45U, G70A, C94U, U111C, A145C).

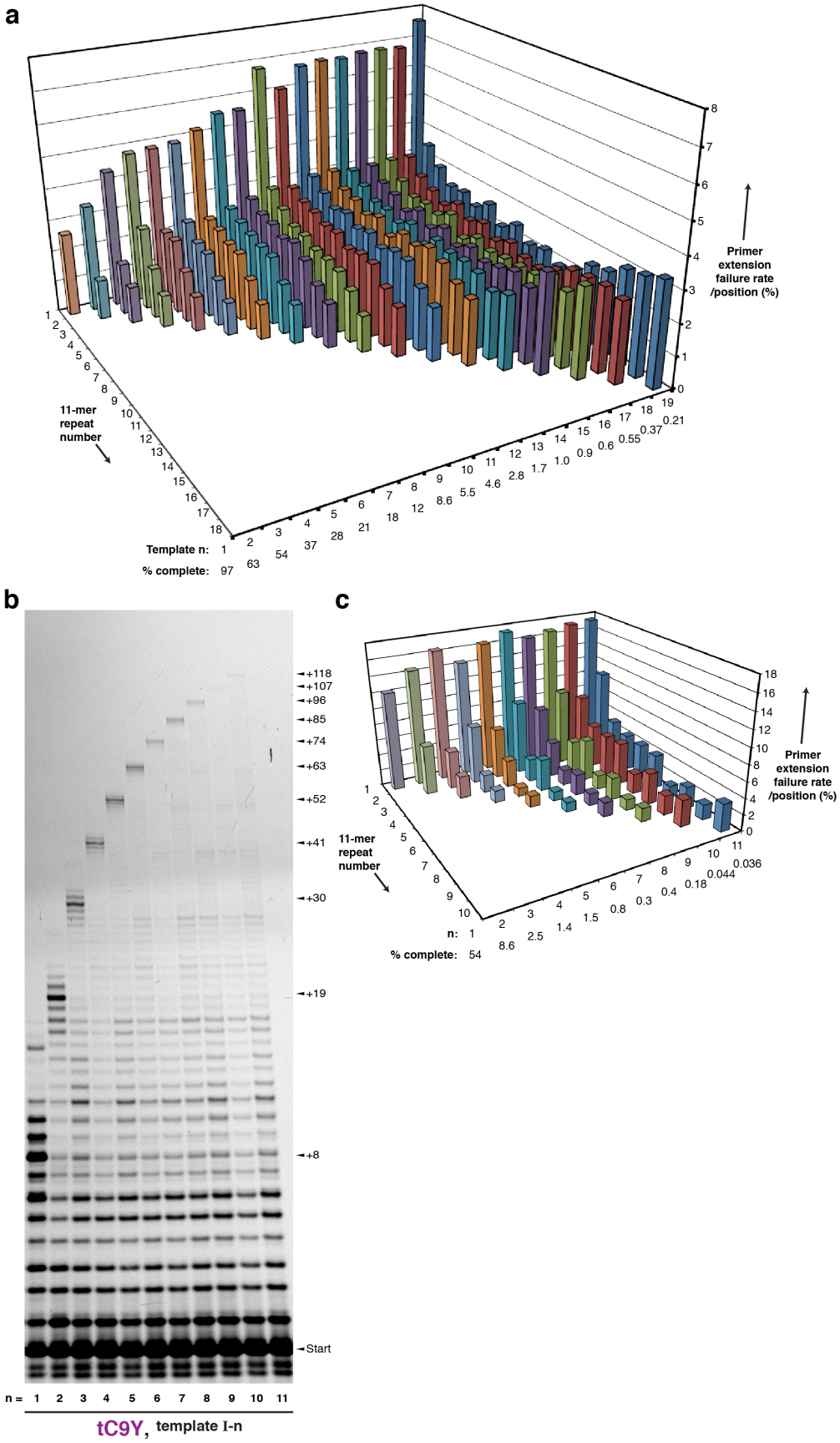


Supplementary Figure S3 I Engineering of the Y ribozyme. a, Secondary structures of the wild-type R18 ribozyme construct (residues not mutagenized in the starting library are shown in grey); the isolated ribozyme C8 (ice-selected mutations

in blue); and ribozyme Y (Z-derived mutation (C60U) in magenta). **b,** Engineering of Y from clone C8. Primer extension activity in ice (–7˚C, 0.5 mM each NTP, 7 days) was used to judge the contributions of mutations and elements, including a mutation (C60U) from Z, the short 'stem2' RNA that completes R18, and the 3' run-through transcript (RTT) on ribozyme activity. Each of the three ice-selected mutations from C8 was important for the Y phenotype (ribozymes 4-6). Further combinations of mutations from Z and Y often proved incompatible, suggesting that Y and Z represent functionally distinct (non-additive) phenotypes. **c,** Comparison of ribozyme activities of R18, Z and Y at -7˚C in ice and at 17˚C (8 days). **d,** Comparison of ribozyme activities of R18 and Y at -7˚C in ice and at 17˚C upon a number of different primer/template duplexes (left panels: -7˚C in ice, 0.5 mM each NTP, 9 days; right panels: 17˚C, 4 mM each NTP, 41.5 hours). The average number of nucleotides added per primer is indicated below each lane.



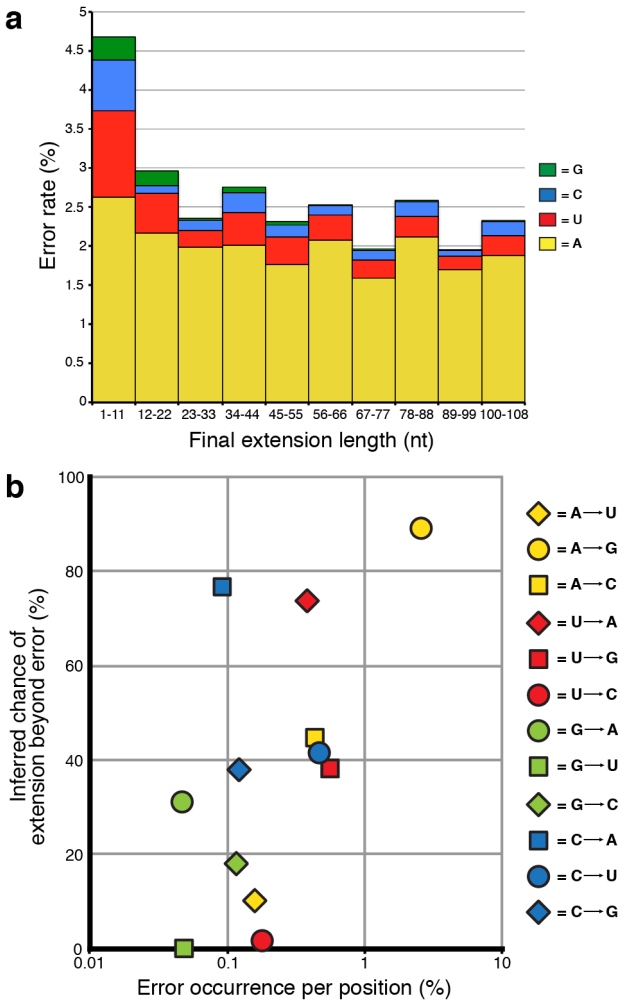
Supplementary Figure S4 I Time course of long RNA syntheses. a, Comparison of ss_{C19}-mediated primer extension (BioFITC-A / template I-6, 17˚C ($MgCl_2$ extension buffer, red) and -7˚C in ice ($MgSO_4$ extension buffer, blue)). tC9Y (Fig. 4b) has a 4-nt longer A linker than tC19Y between the 5' ss_{C19} sequence and the ribozyme, and the tC19 ribozyme[11] lacks the C60U, U72G and C97A mutations in tC19Y. The percentages of primers extended to within 19 nucleotides of the end of the template, i.e. the region where the hybridised ss_{C19} sequence prevents further extension, are indicated at the top. **b,** Primer extension activity of tC9Y upon the I-**n** series of templates in a shorter incubation (BioFITC-A, 17˚C, 60 hours).



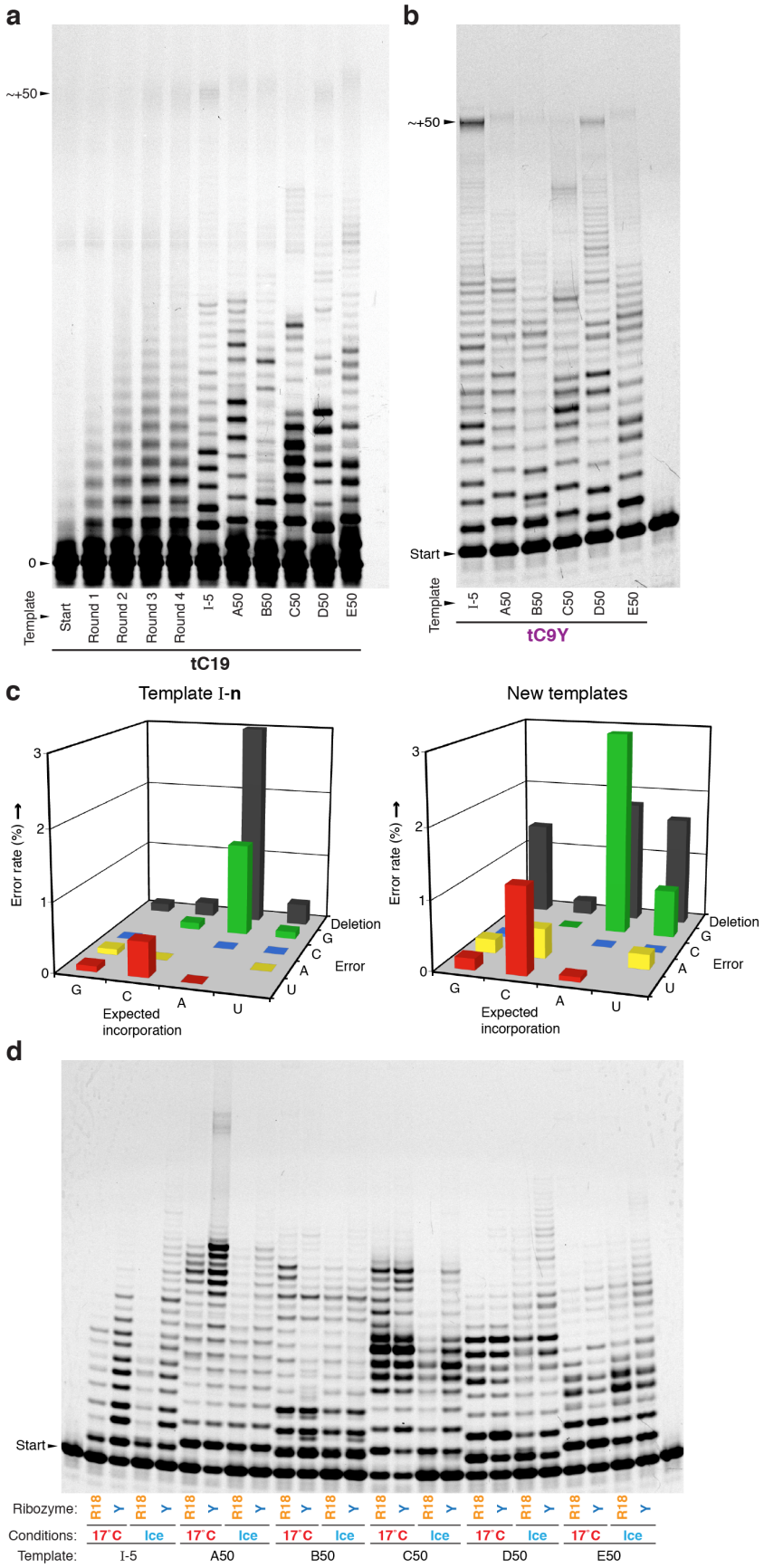
Supplementary Figure S5 I Quantification of termination parameters in long extensions. a, Quantification of the extension by tC9Y on the I-**n** repeat template

series in Fig. 4a. The '% complete' indicated for each template **n** represents the percentage of primers that were extended full length: to within 19 nucleotides of the end of the template (where the hybridised ss_{C19} sequence prevents further extension). For the intermediate 11-mer repeats in the extension, we calculated the fraction of primers that had failed to progress beyond each, and derived an average extension failure (termination parameter) per nucleotide. Plotting these reveals the pattern of termination as tC9Y synthesizes from the first repeat to the last, across a range of template lengths. A high apparent termination rate in the first 11-mer repeat present in all templates may represent ss_{C19}-independent extension of some primers by tC9Y. Between 15-25% of primers remained unextended. **b,** Denaturing PAGE depicting extension of primer BioFITC-A by tC9Y at -7˚C in ice upon the I-**n** series of templates (28 days, $MgSO_4$ extension buffer). **c,** Quantification of termination parameters of in-ice extension (b). Higher termination is also observed, across several of the initial 11-mer repeats. ss_{C19}-independent extension by Y yields longer products in ice than at 17˚C (Fig. 3c), and could explain this termination persistence into the synthesis. Between 70-80% of primers remained unextended in ice. Together this indicates that primers are only engaged inefficiently by tC9Y for ss_{C19}-mediated extension in ice.

Supplementary Figure S6 | Influence of core mutations on long ss$_{C19}$-mediated extensions. Denaturing PAGE depicting extension by the indicated ribozymes of primer A10 upon some members of the I-**n** series of templates at 17˚C for 7 days; average lengths of extended primers are indicated above each lane. Below each lane, estimates of the primer termination rate per position are given, calculated as described in Supplementary Fig. S5a using data from all the repeats in each lane, except the final two containing extension products and the first.

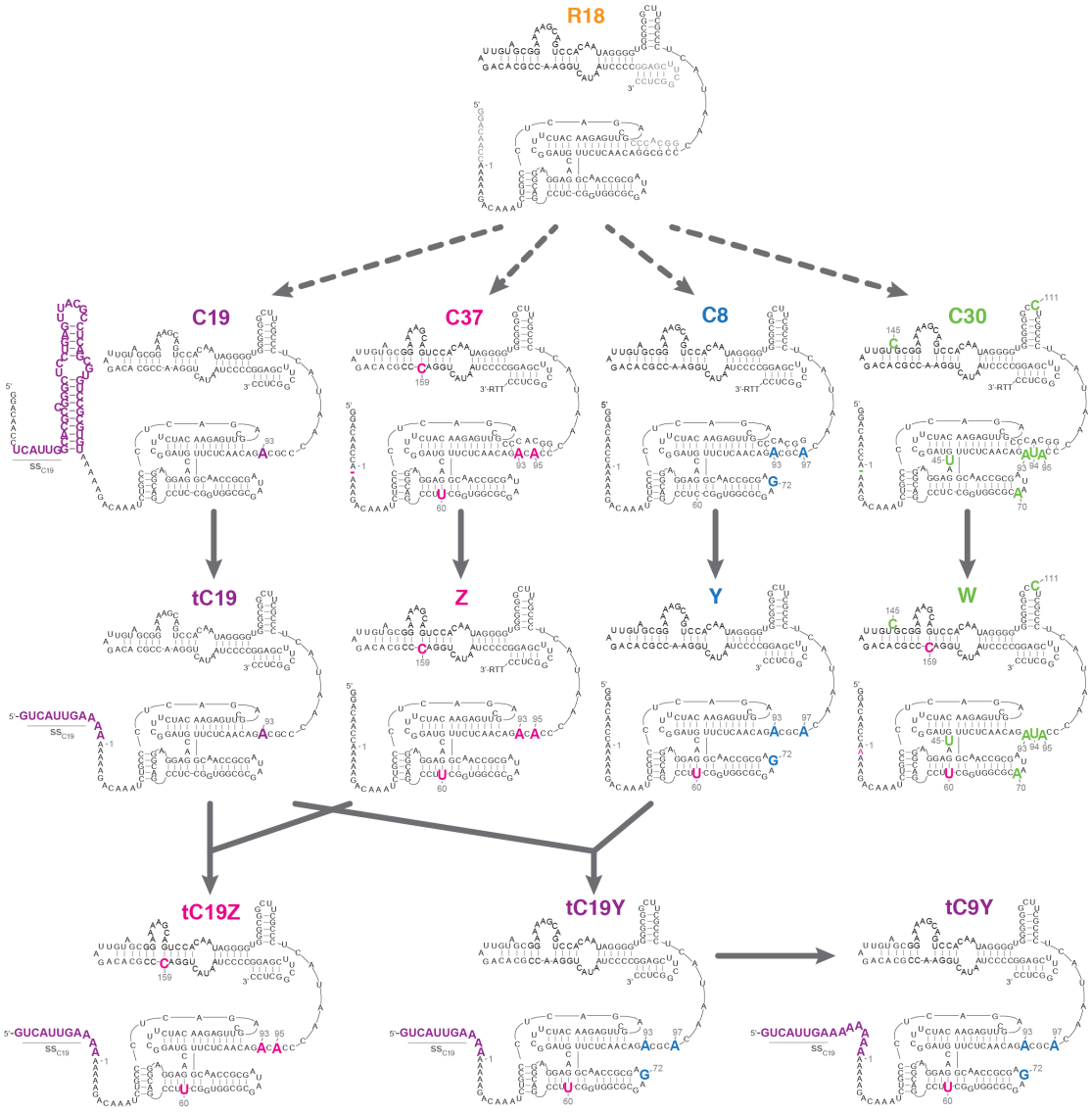


Supplementary Figure S7 I High-throughput sequencing of total tC9Y extension products. a, The overall error rates within sequences of varying lengths; each column is subdivided into the relative error rates at each of the four bases. The error spectrum of full-length extension products (97.7% fidelity, length 100-108) is broadly similar to those of shorter sequences, with the exception of the shortest products (potentially enriched in ss_{C19}-independent extension products). **b,** Using the average extended primer length on template I-10 (26.81) and the error rates associated with the final and internal positions (Fig. 4d), the chances of each error type being extended can be inferred, providing an estimate of their influence upon termination; these are plotted against the abundance of each error type. Most error species can undergo further extension, such as A to G transitions, allowing their representation in final extension products and the benchmarking of fidelity effects through sequencing of such products: 80% of all errors, including deletions, are found at internal positions in sequences. However, a couple of the rarer errors, such as G to U transversions, appear to effectively terminate ribozyme-catalyzed extension, as we found no evidence of such errors occurring within sequences.



Supplementary Figure S8 | Extension of in-ice selected templates. a, Template activity of the polyclonal pool in ice (primer BioFITCU10-A, tC19, –7˚C in ice,

17.5 days, $MgSO_4$ extension buffer): the starting template library (Start), the template libraries during successive rounds of template selection in ice (Rounds 1-4) compared to the template I-5, and five in-ice selected templates (A50-E50) cloned from the Round 4 template pool. Extension product separation did not use competing oligonucleotides; instead, primers were bound to MyOne Streptavidin C1 microbeads, which were heated (60˚C, 1 min) twice in denaturing buffer (9.3 M Urea, 10 mM Tris pH 7.4, 1 mM EDTA, 0.1% Tween-20) to remove template molecules, allowing clear resolution by urea-PAGE after primer elution (95˚C 5 min, in 95% formamide, 10 mM EDTA). **b,** Template activity of in-ice selected templates (A50-E50) with ribozyme tC9Y (primer A, -7˚C Ice, 25 days, $MgSO_4$ extension buffer). **c,** Fidelity of tC19Y on template I-**n** compared to average fidelity on in-ice selected templates (A50, B50, D50, E50) calculated from full-length extension products (Supplementary Table S2). The fidelity values for templates A50, B50, D50, E50 were between 96.7% and 97.5%, slightly lower than the 98.3% observed on template I-n template. We also observed a broadened pattern of errors (compared to the I-n template), although errors arising from G·U / U·G wobble pairs and deletions remain the dominant substitutions. **d,** Template activity of in-ice selected templates (A50-E50) and template I-5 at –7˚C in ice and at 17˚C (primer A, 25 days) in the absence of template tethering. The tC19 ribozyme used in the template selection is closely related to the wild-type R18 ribozyme, which also performs well on the evolved templates.



Supplementary Figure S9 | Provenance of ribozymes used in this and previous studies. Dashed arrows to the first tier of clones indicate evolution from the parental R18 ribozyme using CBT; solid arrows indicate engineered derivatives. The experiments yielding the ribozymes on the left are described in ref. 11.

Supplementary Table S1. Parameters used for CBT selection in ice.

Round	Starting gene density /bead	Primer	Template	Extension time (h)	[1]Minicircle stringency	[2]Library size	[3]Gene recovery	Post-recovery	Pool polyclonal activity vs. wt (Starting pool: 11%)
1	1	BioU10-A	I	186	+7	5.2×10^7	2.72%		15%
2	0.2	BioU10-A	I	550	+7	1.5×10^7	35.35%		26%
3	0.2	BioU10-A	I	172	+7	2.8×10^7	17.83%	StEP	26% (47% pre-StEP)
4	0.4	BioU10-A	I	325	0	4.7×10^7	1.06% High, 3.74% Low	StEP	24%
5	0.4	BioU10-A	II	256	0	2.3×10^7	0.548% High, 1.93% Low	Gel purification	36%
6	0.2	BioU10-A	I	115	0	1.5×10^7	0.35% High, 1.67% Low		56%
7	0.2	BioU10-A	I	124	0	1.1×10^7	0.235% High, 0.995% Low		69%
8	0.1	BioU10-A	I	62	0	4×10^6	0.171% High, 0.819% Low	Gel purification	53% (High) 84% (Low)
Screening		BioU10-A	I	334	P2 (+5) P3 (+3)	44 clones from each of the High and Low pools			

[1]Minicircle stringency denotes the overlap (in nucleotides) with the unextended primer, with a 10-12 nucleotide overlap required to primer RCA. [2]'Library size' was calculated as the total number of sorted single beads × the gene density per bead. [3]'Gene recovery' represents the number of positive sorted beads relative to the library size, as, regardless of the gene density, beads with the highest fluorescence likely carried a gene. In later, 'stringent' rounds, high-fluorescence beads were sorted into two gates: the brightest into a 'High' gate and the rest into a 'Low' gate (beads in each gate are given as % of total bead count), which were amplified separately and then combined at a 1:1 ratio. The polyclonal pool activity was judged by denaturing PAGE.

Supplementary Table S2. Error rates during RNA synthesis upon a range of templates.

Template I-n, tC19Y - full-length					
Correct base:		**G**	**C**	**A**	**U**
Total positions:		2424	988	1029	962
Errors at **17°C**:	G	–	1	14	1
	C	0	–	0	0
	A	2	0	–	0
	U	2	5	0	–
	Deletion	3	2	39	3
	Insertion	0	0	0	0
Positional fidelity (%):		99.7	99.2	94.8	99.6
Overall fidelity (%):		98.3			

Template I-n, tC19Y - full-length					
Correct base:		**G**	**C**	**A**	**U**
Total positions:		531	224	228	209
Errors at **–7°C in ice**:	G	–	0	10	1
	C	5	–	1	2
	A	3	0	–	4
	U	1	2	0	–
	Deletion	3	0	15	3
	Insertion	1	0	0	2
Positional fidelity (%):		97.6	99.1	88.6	94.3
Overall fidelity (%):		94.8			

Template A50, tC19Y - full-length					
Correct base:		**G**	**C**	**A**	**U**
Total positions:		702	121	150	120
Errors at **17°C**:	G	–	0	6	1
	C	0	–	0	0
	A	1	0	–	0
	U	1	4	0	–
	Deletion	6	1	1	1
	Insertion	0	0	0	0
Positional fidelity (%):		98.9	95.9	95.3	98.3
Overall fidelity (%):		97.1			

Template B50, tC19Y - full-length					
Correct base:		**G**	**C**	**A**	**U**
Total positions:		392	36	124	88
Errors at **17°C**:	G	–	0	7	1
	C	0	–	0	0
	A	1	0	–	0
	U	0	0	0	–
	Deletion	3	0	1	1
	Insertion	0	0	0	0
Positional fidelity (%):		99.0	100	93.5	97.7
Overall fidelity (%):		97.5			

Template D50, tC19Y - full-length					
Correct base:		**G**	**C**	**A**	**U**
Total positions:		580	58	232	116
Errors at **17°C**:	G	–	0	5	1
	C	1	–	0	0
	A	1	1	–	1
	U	1	1	0	–
	Deletion	1	0	3	2
	Insertion	0	0	0	0
Positional fidelity (%):		99.3	96.6	96.6	96.6
Overall fidelity (%):		97.2			

Template E50, tC19Y - full-length					
Correct base:		**G**	**C**	**A**	**U**
Total positions:		864	34	357	36
Errors at **17°C**:	G	–	0	5	0
	C	0	–	0	0
	A	2	0	–	0
	U	3	0	1	–
	Deletion	32	0	16	1
	Insertion	0	0	0	0
Positional fidelity (%):		95.7	100	93.8	97.2
Overall fidelity (%):		96.7			

Template I-10, 17°C, tC9Y - all primers					
Correct base:		**G**	**C**	**A**	**U**
Errors rates **in all products (%)**:	G	99.76	0.11	2.67	0.51
	C	0.11	99.24	0.45	0.15
	A	0.04	0.09	92.25	0.41
	U	0.01	0.44	0.13	98.57
	Deletion	0.08	0.11	4.50	0.36
Overall fidelity (%):		97.41			
Template I-10, 17°C, tC9Y - all primers					
Correct base:		**G**	**C**	**A**	**U**
Substitutions **at final two positions (%)**:	G	97.60	1.05	6.07	4.82
	C	1.31	94.77	3.43	2.35
	A	0.44	0.35	88.55	1.63
	U	0.65	3.82	1.94	91.20
Overall substitution fidelity (%):		92.93			
Template I-10, 17°C, tC9Y - all primers					
Correct base:		**G**	**C**	**A**	**U**
Errors rates **at preceding internal positions (%)**:	G	99.86	0.05	2.31	0.21
	C	0.02	99.56	0.20	0.00
	A	0.01	0.07	92.59	0.28
	U	0	0.19	0.02	99.10
	Deletion	0.10	0.13	4.88	0.40
Overall fidelity (%):		97.73			
Overall substitution fidelity (%):		99.21			

Fidelities are estimated by collating errors in the sequenced extension products listed below. The overall fidelity represents a geometric mean of the positional fidelities of incorporation opposite each base.

17˚C **Template I-n tC19Y** fidelities were judged from 31 fully extended products on template I-5, 17 on I-6, 17 on I-9, and 11 on I-10.

−7˚C ice **Template I-n tC19Y** fidelities were judged from 19 fully extended products on template I-6.

17˚C **Template A50 tC19Y** fidelities were judged from the central random sequence-derived segments of 31 extended products on Template A50.

17˚C **Template B50 tC19Y** fidelities were judged from the central random sequence-derived segments of 31 extended products on Template B50.

17˚C **Template D50 tC19Y** fidelities were judged from the central random sequence-derived segments of 31 extended products on Template D50.

17˚C **Template E50 tC19Y** fidelities were judged from the central random sequence-derived segments of 31 extended products on Template E50.

Calculation of 17˚C **Template I-10 tC9Y** weighted fidelities from all primers is described in Fig. 4, Supplementary Fig. S7 and Supplementary Materials and Methods.

Supplementary Table S3. Nucleic acids. Those sequences not listed with previous descriptions of protocols are listed below, along with any purification methods (GP = gel purified, RNE = purified using RNeasy (Qiagen)). RNA residues are listed in black, and DNA in gray; ss_{C19} sequences and binding sites are red, and primer-binding sites in templates are underlined.

Application	*Name*	*Source, purification*	*Sequence (5′→3′)*
Ribozymes	wt	IVT, RNE	GGACAACCAAAAAGACAAAUCUGCCCUCAGAGCUUGAGAACAUCUUCGGAUGCAGA GGAGGCAGCCUCCGGUGGCGCGAUAGCGCCAACGUUCUCAACAGGCGCCCAAUACU CCCGCUUCGGCGGGUGGGGAUAACACCUGACGAAAAGGCGAUGUUAGACACGCCAA GGUCAUAAUCCCCGGAGCUUCGGCUCCGCGGCCGCAAAAAAAAAAGGCUUACC
	R18	IVT, RNE	GGACAACCAAAAAGACAAAUCUGCCCUCAGAGCUUGAGAACAUCUUCGGAUGCAGA GGAGGCAGCCUCCGGUGGCGCGAUAGCGCCAACGUUCUCAACAGGCGCCCAAUACU CCCGCUUCGGCGGGUGGGGAUAACACCUGACGAAAAGGCGAUGUUAGACACGCCAA GGUCAUAAUCCCCGGAGCUUCGGCUCC
	C30	IVT, RNE	GGACAACCAAAAGACAAAUCUGCCCUCAGAGCUUGAGAACAUCUUCGGAUGUAGAG GAGGCAGCCUCCGGUGGCGCAAUAGCGCCAACGUUCUCAACAGAUACCCAAUACUC CCGCUCCGGCGGGUGGGGAUAACACCUGACGAAAAGGCGCUGUUAGACACGCCAAG GUCAUAAUCCCCGGAGCUUCGGCUCCGCGGCCGCAAAAAAAAAAGGCUUACC
	C8	IVT, RNE	GGACAACCAAAAAGACAAAUCUGCCCUCAGAGCUUGAGAACAUCUUCGGAUGCAGA GGAGGCAGCCUCCGGUGGCGCGAGAGCGCCAACGUUCUCAACAGACGCACAAUACU CCCGCUUCGGCGGGUGGGGAUAACACCUGACGAAAAGGCGAUGUUAGACACGCCAA GGUCAUAAUCCCCGGAGCUUCGGCUCCGCGGCCGCAAAAAAAAAAGGCUUACC
	Stem2 (part of wt, R18, C30, C8)	Dharmacon, GP	GGCACC–dideoxyC
	W	IVT, RNE	GGACAACCAAAAAGACAAAUCUGCCCUCAGAGCUUGAGAACAUCUUCGGAUGUAGA GGAGGCAGCCUUCGGUGGCGCAAUAGCGCCAACGUUCUCAACAGAUACCCAAUACU CCCGCUCCGGCGGGUGGGGAUAACACCUGACGAAAAGGCGCUGUUAGACACGCCCA GGUCAUAAUCCCCGGAGCUUCGGCUCC
	Z Z^{-RTT} lacks bold sequence	IVT, RNE	GGACAACCAAAAAGACAAAUCUGCCCUCAGAGCUUGAGAACAUCUUCGGAUGCAGA GGAGGCAGCCUUCGGUGGCGCGAUAGCGCCAACGUUCUCAACAGACACCCAAUACU CCCGCUUCGGCGGGUGGGGAUAACACCUGACGAAAAGGCGAUGUUAGACACGCCCA GGUCAUAAUCCCCGGAGCUUCGGCUCC**GCGGCCGCAAAAAAAAAAGGCUUACC**
	Y	IVT, RNE	GGACAACCAAAAAGACAAAUCUGCCCUCAGAGCUUGAGAACAUCUUCGGAUGCAGA GGAGGCAGCCUUCGGUGGCGCGAGAGCGCCAACGUUCUCAACAGACGCACAAUACU CCCGCUUCGGCGGGUGGGGAUAACACCUGACGAAAAGGCGAUGUUAGACACGCCAA GGUCAUAAUCCCCGGAGCUUCGGCUCC
	tC19	IVT, GP	GUCAUUGAAAAAAAAAGACAAAUCUGCCCUCAGAGCUUGAGAACAUCUUCGGAUGC AGAGGAGGCAGCCUCCGGUGGCGCGAUAGCGCCAACGUUCUCAACAGACGCCCAAU ACUCCCGCUUCGGCGGGUGGGGAUAACACCUGACGAAAAGGCGAUGUUAGACACGC CAAGGUCAUAAUCCCCGGAGCUUCGGCUCC
	tC19Y	IVT, GP	GUCAUUGAAAAAAAAAGACAAAUCUGCCCUCAGAGCUUGAGAACAUCUUCGGAUGC AGAGGAGGCAGCCUUCGGUGGCGCGAGAGCGCCAACGUUCUCAACAGACGCACAAU ACUCCCGCUUCGGCGGGUGGGGAUAACACCUGACGAAAAGGCGAUGUUAGACACGC CAAGGUCAUAAUCCCCGGAGCUUCGGCUCC
	tC9Y	IVT, GP	GUCAUUGAAAAAAAAAAAAGACAAAUCUGCCCUCAGAGCUUGAGAACAUCUUCGG AUGCAGAGGAGGCAGCCUUCGGUGGCGCGAGAGCGCCAACGUUCUCAACAGACGCA CAAUACUCCCGCUUCGGCGGGUGGGGAUAACACCUGACGAAAAGGCGAUGUUAGAC ACGCCAAGGUCAUAAUCCCCGGAGCUUCGGCUCC
	tC19Z	IVT, GP	GUCAUUGAAAAAAAAAGACAAAUCUGCCCUCAGAGCUUGAGAACAUCUUCGGAUGC AGAGGAGGCAGCCUUCGGUGGCGCGAUAGCGCCAACGUUCUCAACAGACACCCAAU ACUCCCGCUUCGGCGGGUGGGGAUAACACCUGACGAAAAGGCGAUGUUAGACACGC CCAGGUCAUAAUCCCCGGAGCUUCGGCUCC
CBT DNA minicircles[11]	DNAcirc+7 [template I]	Sigma	GTTACCTTTCAATGAATCCACGCTTCGCACGGTTGGTGTAACGACTTTTCGGATTT CTAGGATCTCCAAGTATGTTCTAAGTC
	DNAcirc+5 [template II]	Sigma	GTTACTTTTGCCTCCCTTCGCACGGTTCTTTGTAACGATCTTTCTGGATTTCTAGG ATCTCGTCCCTATAGTGAGTCGGTTCTAGATC

	DNAcirc0 [template I]	Sigma	GTTACTTTTCAATGAATCCACGCTTCGCATGTAACGACTTTTCGGATTTCTAGGATCTCCAAGTATGTTCTAAGTC
	DNAcirc0 [template II]	Sigma	GTTACTTCTATCTCCCTTCGCATTCTTTGTAACGATCTTTCTGGATTTCTAGGATCTCGTCCCTATAGTGAGTCGGTTCTAGATC
RPA Probes[11]	P2	Sigma	Biotin–TCCCTTCGCACGGTT
	P3	Sigma	Biotin–TGAATCCACGCTTCGCACGG
RNA primers	A	Dharmacon	Fluorescein–CUGCCAACCG
	TAMRA-A	IDT	TAMRA–CUGCCAACCG
	BioU10-A	IDT	Biotin–UUUUUUUUUUCUGCCAACCG
	BioU10-Aext	Dharmacon	Biotin–UUUUUUUUUUCUGCCAACCGUGCGAAGGGAG
	BioFITC-A	IDT	Fluorescein–(C6–NH$_2$–dT)–CUGCCAACCG
	BioFITCU10-A	Dharmacon	Fluorescein–(C6–NH$_2$–dT)–UUUUUUUUUCUGCCAACCG
	BioloopA10	IDT	Biotin-Biotin–GACUCUUCGGAGUCCUGCCAACCG
	B	Dharmacon	Fluorescein–GAAUCAAGGG
	C	Dharmacon	Fluorescein–GAUAGGUAG
Ribozyme templates	I	Dharmacon	CAAUGAAUCCACGCUUCGCACGGUUGGCAGAACA
	II	Dharmacon	UUCUAUCUCCCUUCGCACGGUUGGCAG
	I-s	Dharmacon	GACGCUUCGCACGGUUGGCAG
	IV	Sigma, Fill-in	Forward: GATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTATA Reverse: TTTTTTTTTTGAATCAAGGGCCGAGGTCCAATCTTCATTGTCTATAGTGAGTCGTATTAATTTC
		IVT, RNE	Transcript: GCUUAAACAGAUUGGACCUCGGCCCUUGAUUCAAAAAAAAAA
	VI	Sigma, Fill-in	Forward: GATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTATA Reverse: TTTTTTTTTTGATAGGTAGCTACGCCGTGGGTTTCATTGTCTATAGTGAGTCGTATTAATTTC
		IVT, RNE	Transcript: GUCUUUAGAACCCACGGCGUAGCUACCUAUCAAAAAAAAAA
	I-n (n = 1-10)	Sigma, Fill-in	Forward: GATCGATCTCGCCCGCGAAATTAATACGACTCACTATA Reverse: TTTTTTTTTTCTGCCAACCG(TGCGAAGCGTG)$_n$TCATTGACTATAGTGAGTCGTATTAATTTC
		IVT, GP	Transcript: GUCAAUGA(CACGCUUCGCA)$_n$CGGUUGGCAGAAAAAAAAAA
	I-n (n = 11-19)	IDT Ultramers, Fill-in	Forward: GATCGATCTCGCCCGCGAAATTAATACGACTCACTATA Reverse: TTTTTTTTTTCTGCCAACCG(TGCGAAGCGTG)$_n$TCATTGACTATAGTGAGTCGTATTAATTTC
		IVT, GP	Transcript: GUCAAUGA(CACGCUUCGCA)$_n$CGGUUGGCAGAAAAAAAAAA
Competing oligonucle-otides	CompI (for I, I-s)	Dharmacon	CUGCCAACCGUGCGAAGCGUGGAUUCAUUG
	CompII (for II)	Dharmacon	GCUGCCAACCGUGCGAAGGGAGAUAGAA
	CompIV (for IV)	Sigma, Fill-in	Forward: GATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTATA Reverse: GCTTAAACAGATTGGACCTCGGCCCTTGATTCCTATAGTGAGTCGTATTAATTTC
		IVT, RNE	Transcript: GGAAUCAAGGGCCGAGGUCCAAUCUGUUUAAGC
	CompVI (for VI)	Sigma, Fill-in	Forward: GATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTATA Reverse: GTCTTTAGAACCCACGGCGTAGCTACCTATCCTATAGTGAGTCGTATTAATTTC
		IVT, RNE	Transcript: GGAUAGGUAGCUACGCCGUGGGUUCUAAAGAC
	CompI-6 (for I-1-6)	Sigma, Fill-in	Forward: GATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTATA Reverse: GTCAATGACACGCTTCGCACACGCTTCGCACACGCTTCGCACACGCTTCGCACACGCTTCGCACACGCTTCGCACGGTTGGCAGCCTATAGTGAGTCGTATTAATTTC
		IVT, RNE	Transcript: GGCUGCCAACCGUGCGAAGCGUGUGCGAAGCGUGUGCGAAGCGUGUGCGAAGCGUGUGCGAAGCGUGUGCGAAGCGUGUCAUUGAC
	CompI-11 (for I-7-11)	IDT Ultramer, Fill-in	Forward: GATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTATA Reverse: GTCAATGACACGCTTCGCACACGCTTCGCACACGCTTCGCACACGCTTCGCACACGCTTCGCACACGCTTCGCACACGCTTCGCACACGCTTCGCACACGCTTCGCACACGCTTCGCACGGTTGGCAGCCTATAGTGAGTCGTATTAATTTC
		IVT, RNE	Transcript: GGCUGCCAACCGUGCGAAGCGUGUGCGAAGCGUGUGCGAAGCGUGUGCGAAGCGUGUGCGAAGCGUGUGCGAAGCGUGUGCGAAGCGUGUGCGAAGCGUGUGCGAAGCGUGUGCGAAGCGUGUGCGAAGCGUGUCAUUGAC
	CompI-19 (for I-12-19)	IDT Ultramer, Fill-in	Forward: GATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTATA Reverse: GTCAATGACACGCTTCGCACACGCTTCGCACACGCTTCGCACACGCTTCGCACACGCTTCGCACACGCTTCGCACACGCTTCGCACACGCTTCGCACACGCTTCGCACACGCTTCGCACACGCTTCGCACACGCTTCGCACACGCTTCGCACACGCTTCGCACACGCTTCGCACACGCTTCGCACACGCTTCGCACACGCAAGGCACGGTTGGCAGCCTATAGTGAGTCGTATTAATTTC
		IVT, DNAse-treated, RNE	Transcript: GGCUGCCAACCGUGCCUUGCGUGUGCGAAGCGUGUGCGAAGCGUGUGCGAAGCGUGUGCGAAGCGUGUGCGAAGCGUGUGCGAAGCGUGUGCGAAGCGUGUGCGAAGCGUGUGCGAAGCGUGUGCGAAGCGUGUGCGAAGCGUGUGCGAAGCGUGUGCGAAGCGUGUGCGAAGCGUGUGCGAAGCGUGUGCGAAGCGUGUGCGAAGCGUGUCAUUGAC
	CompA50 (for A50)	Sigma, Fill-in	Forward: GATCGATCTCGCCCGCGAAATTAATACGACTCACTATA Reverse: GTCAATGACACGCTTCGCACACGCTTCGCACACCTCCACCTCGATCGCTTCCCGCCCCCCCCCCCGGTTGGCAGCCTATAGTGAGTCGTATTAATTTC
		IVT, RNE	Transcript: GGCUGCCAACCGGGGGGGGGGCGGGAAGCGAUCGAGGUGGAGGUGUGC

			GAAGCGUGUGCGAAGCGUGUCAUUGAC
	CompB50 (for B50)	Sigma, Fill-in	Forward: GATCGATCTCGCCCGCGAAATTAATACGACTCACTATA Reverse: GTCAATGACACGCTTCGCACACGCTTCCCACATCTCCTCTTCGATCGCC CCACACCCCCCCCTCGGTTGGCAGCCTATAGTGAGTCGTATTAATTTC
		IVT, RNE	Transcript: GGCUGCCAACCGAGGGGGGGGUGUGGGGCGAUCGAAGAGGAGAUGUGG GAAGCGUGUGCGAAGCGUGUCAUUGAC
	CompC50 (for C50)	Sigma, Fill-in	Forward: GATCGATCTCGCCCGCGAAATTAATACGACTCACTATA Reverse: GTCAATGACACGCTTCGCACACGCTTCATAGCACACTACACCACCCGAC TCCCCGCTTCCCCCGGTTGGCAGCCTATAGTGAGTCGTATTAATTTC
		IVT, RNE	Transcript: GGCUGCCAACCGGGGGAAGCGGGGAGUCGGGUGGUGUAGUGUGCUAUG AAGCGUGUGCGAAGCGUGUCAUUGAC
	CompD50 (for D50)	Sigma, Fill-in	Forward: GATCGATCTCGCCCGCGAAATTAATACGACTCACTATA Reverse: GTCAATGACACGCTTCGCACACGCTTCCCGCTCCTCCACCCACTCCTTC CTTACCATCCCGCGGTTGGCAGCCTATAGTGAGTCGTATTAATTTC
		IVT, RNE	Transcript: GGCUGCCAACCGCGGGAUGGUAAGGAAGGAGUGGGUGGAGGAGCGGGA AGCGUGUGCGAAGCGUGUCAUUGAC
	CompE50 (for E50)	Sigma, Fill-in	Forward: GATCGATCTCGCCCGCGAAATTAATACGACTCACTATA Reverse: GTCAATGACACGCTTCGCACACGCTTCCCCCCCCCCCCCTCTTCGCCTCC ACTCTTCCCCTTTCCCGGTTGGCAGCCTATAGTGAGTCGTATTAATTTC
		IVT, RNE	Transcript: GGCUGCCAACCGGGAAAGGGGAAGAGUGGAGGCGAAGAGGGGGGGGGG GGAAGCGUGUGCGAAGCGUGUCAUUGAC
Template selection procedure[11]	T7C19	Sigma	GATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTATAGTCAATGACACGCT TCGCACACGCTTC
	TempLib	Sigma	CCGCCAACCGnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnGAAGCGTGTG CGAAGCGTG
	Template library	IVT	GUCAAUGACACGCUUCGCACACGCUUCnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn nnnnnnnCGGUUGGCGG
	TempRec	Sigma	CCTTATTAGGGTTTACCATTCGCACACGCTTC
	PriRec	Sigma	ACACGACGCTCTTCCGATCTCACGGGTTTTTTTTTTC
	U10D	Sigma	TTTTTTTTTTCTGCCAACCG
Template selection	Template structure	N/A	GUCAAUGACACGCUUCGCACACGCUUC–selected sequence– CGGUUGGCAGAAAAAAAAA
Round 4 output template sequences	95 selected sequences	N/A	--GCACCACCAUUCUUACACUCGCGCGCUUCAUCCC ---------CCCCCAUUCGCCCCUGUCCUCCCCCCC --CCGACCUCCCCCCCUACACCCCCACUCUCCACCCC --CCACCCUUCCCCCCCCGCUCCUUCCUUCCCCCCUC ---CUCCCUCCCCCCCGCAUUCCUUCAUCGCCUCCCC ---CUAGAUACCCCCCCCUUCUCUCCACUCCACCUCC ---CCACACACUCGCCCUCGCCUUCGCACCGUUUCC --CCGCUCCUCCACCCACUCCUUCCUUACCAUCCCG = D50 ---CCUCUCUCCCCCCCUCUCUUACCCCCCCACGCCC UCUCCCGCACCCACUCUUCCCUUCCUCGCUCACACC ----CCGUGCCCUCUCCUCGCUUCCCUUCGCUCGCC -CCUCCCUCUCCCCCCCUUCCCUUCCCUUCCACGCCC CACUCCCUCGACGCCUCUCCCUUCCCUUCUCUACCC --CACUCACUCGACGCCCGGCCCCACCUCUCUACUC --CACCCACCCGCCGCCUCUCCCCACUCCUCUACCC ----CUUCCACACCUCCUCCCUCCUCACACCCACCC CCCCAUACCUCUUCGCCUUCGUCCACUCCCCUCCCC ----CACCCUCGCUCCUUCGUACCAUCGCCAUCCAC -UACCCUCCACAUCCUCCGCACCCACUUCCAUCGCU UCACCCUCCACACUCUCUGCACCCGCCGCCAUCGCC --CUCGCCCGCUCGCACACCCACCCAUCCACCCCUU ----CGCCCCCACACACAUAUUACCCACCCUCCCGC -CGCACUCCACACACACAUUACCCUUCCCUUCCCCC CGCACUCCCACACACACAUUACCCCUCCCUCCUCGC CGCACUCCCACACACACAUUACCCCUCCCUUCUCGC CGCACUCACACACACAUAACACCCUUCCCUUCUCGC CGCACUCACACACACAUACCACCCCUCCCUCCUCGC -----GCACCCACAAGCAUCCCCCCCCCCUCCUCGC ----CGCACUCACGCGCACUACCCACCCCCCCUCGC -CCACACACCCACCCUCACCCUUCGCACCACCCUCC -CCGCACACUCGCCCUCCCCCUUCGCAUCAUCCCCC --CGCACGCUCACCCCCGCCCCUCGCAUCAUCCCUC -CCGCACACUCACCCCCGCCCCUCGCACCAUCCUCC -CCGCACACUCGCCCCCACUUCUCGCAUCAUCCUCC --CUCCCACACCUCGUCCCACCUCCCACACCACUCC --CUCCACACCUCACCCACACCUCGCACACCACCCC ----CUCACCUCCACCCACUCCUCACACACCACCCC -UUCCACACCUUACCCCUCUCCUCCCACACCGCCCC ----CUACCCACUCUCCCCUCACUCCUCCCUUCCCCC ---------------CCUCUCCUCCUCCCUUCACCG -----CAACCCACCGCACGCACACUUCCCUCCCCUU -GUCCCUUCCCACUCGCUCCGUACCUCCCACUCCUA ---UCCCUCCUACCCUCCACCCUCCACCCCCUCCCG ----------------UCCCUCUCCGCCCUCACACU GACUCCAUCGCACCUCCCUCUCCUCGCCCCGUUCGU

```
GACUCCACUCUCCCCACUCCGCCUCGCUCUCAUCCC
-------------------------ACACUCCACCC
--UACUCUACUCCUCCCUCCUACACACCCUCCUCCC
----GCCAUCCCCCGCAUCACACACACACCCCUCCC
CGCACCCCCCCGCAGCUCACCCGCUUUUCCCACCCCC
-CCGACCUCAUCCUUACACCCCCACUCUCCACCCUC
---GACGUCCACGGACCCCUCGCUCUAUCCCCCCUC
---CCGAACCUCAUCUCUACCCCACCCUCCACCCUC
--CCUGCACUUCUUCGCCUUCGCUCAUUCUCUCCCC
-CCCCGCACUUCUUCGCCUUCGCUCACUUCCCCCCC
GCACACCUCCACCUCGAUCGCUUCCCGCCCCCCCCC = A50
CCACAUCUCCUCUUCGAUCGCCCCACACCCCCCCCU = B50
-CUUCCACCCUCCCCCUUCUCCCUCUCCGCCCCACU
------CUCCCUCCCCCUCUCUCCCUCCGCCCCACU
--GCACAUCUCCACUUCGAUCGCCCCACUCCCCCCU
---ACUCUCACACAUUCGCACGCUAUUCUCCCCUCU
-CACAUCUCCCGAGCUCACCCGCUCUCCUUACACCU
-AAACCCCUUCCUCGCUCCACACCCUCCCCCUCUUC
----CCCCUCUCCCCCCCCUCUCUCUCCUUCUCUAC
---CACCGUCCCUCUUCCCACUCCUCCCCCACCAUC
---UACCCGCCGCACUCGCCCCCCCCCCCUCACAUC
-CUACCCUCCACACCCUCCACACCCGCUCCAUCGCU
-GUACCACACCCACCACCCAGACUCCUUGCCUCCCU
GCAACACACCACACCACCCAGUCCCCUCGCAUCCCU
ACACCACACCACACCACCCAGACUCCUCGCUUCCUC
GUAACACACUACACUACCCAGACUCCUCGCUUCCCC
GUACCACACUACACCACCCAGAUUCCCCGCCUGCCC
-------CACACCCUCCCCCGCCCUUCACUCUCCCC
---CUCGCACACCCGCUCGCAUCCUUAGCCCUCACG
---CUCCCCCAACCCACCUCCACUCCCAAUCCCUCA
-AUAGCACACUACACCACCCGACUCCCCGCUUCCCC = C50
---GUAGCACACCACACCACCCCCCCCCGCCUCUCC
-------------------------CUCCCCCUCGC
-----CACAUUCUCCCACACACCCUUUCCCUCCCGC
---CACACUCACACACACACCACCCCCCCUUCUCGC
--GCACUAACACACGCACACUACCCUUCCCUCUCGC
-CGCACCCACGCACACACCCUCCCUCCCCCUCCUCGC
---CGCACUCCCACACACAGACCCUCCCCCCCUCGC
-------CCCACACACACACACUACCCUUCCCUCCU
CACAUUACGCACACCCUCGCACUCCACUCCACUCUA
---UUCCCUCCUCCACCCACUCCCUCCUCCGCCCUA
---UUCCCCCUUCCACCCACUCCCUCCUCCACCCUA
---UUCCCCCUUCCACCCACUCCCUCCUCCACCCUA
--UCCCGCACCUCUUCGCCCUCGCUCAUCCCUUCCC
---CCCCGUAUCCUUCGCCUUCGCUCAUUCCCUCUC
CCCCCCCCCCCUCUUCGCCUCCACUCUUCCCCUUUCC = E50
----UACCACCCCCCCCCACACUCCCACUCAUCCAC
-------------------------CACUCAUCCAC
----CACCCCUACCGUCCACGCUCGUCCUCCCACCC
```

Supplementary references

33 Vaidya, N. *et al.* Spontaneous network formation among cooperative RNA replicators. *Nature* **491**, 72-7 (2012).

34 Zhao, H. *et al.* Molecular evolution by staggered extension process (StEP) in vitro recombination. *Nat. Biotechnol.* **16**, 258-261 (1998).