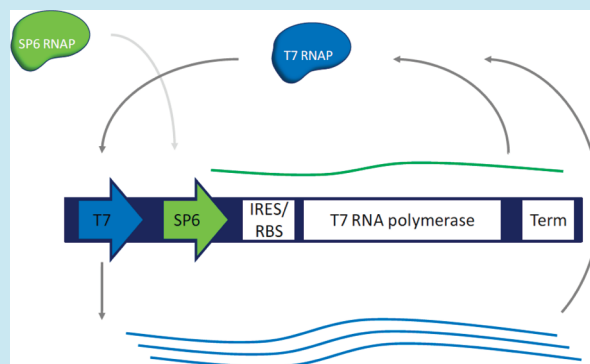# An *in vitro* Autogene

Eric A. Davidson, Adam J. Meyer, Jared W. Ellefson, Matthew Levy, and Andrew D. Ellington*

Department of Chemistry and Biochemistry, Institute for Cell and Molecular Biology, Center for Systems and Synthetic Biology, University of Texas at Austin, Austin, Texas 78712, United States

S *Supporting Information*

ABSTRACT: Recent technological advances have allowed development of increasingly complex systems for *in vitro* evolution. Here, we describe an *in vitro* autogene composed of a self-amplifying T7 RNA polymerase system. Functional autogene templates in cell-free lysate produce T7 RNA polymerase, which amplifies the autogene genetic information through a positive feedback architecture. Compartmentalization of individual templates within a water-in-oil emulsion links genotype and phenotype, allowing evolution.

KEYWORDS: *T7 RNA polymerase, in vitro selection, in vitro compartmentalization, autogene, self-amplifying genetic system*

One of the first demonstrations of *in vitro* evolution was also one of the most powerful: Sol Spiegelman and his co-workers showed that bacteriophage Qβ replicase, in the absence of a living cell, could replicate its substrate, the Qβ RNA phage,[1] and in the process yield both better replicating templates (mini-monsters[2]) and also RNA molecules with altered functionalities (resistance to ethidium bromide[3]). Since that time, there have been a number of additional demonstrations of the *in vitro* replication and evolution of nucleic acids in the presence of exogenous enzymes and added templates, including the development of *in vitro* selection or SELEX[4,5] and the predator:prey system known as CATCH.[6] However, in each of these instances it was only the nucleic acid, not the protein, components of the system that evolved.

The development of robust, cell-free transcription and translation (TNT) systems has enabled engineering and evolution of cell-free genetic systems with both protein and nucleic acid components. With optimized reaction conditions and genetic instructions, cell-free lysates can be used to implement complex genetic regulatory circuits or generate patterns.[7,8] In one example, polymerases, repressors, and reporter genes were connected through regulatory DNA sequences to create systems with increasing complexity.[7] These results highlight a barrier to engineering cell-free lysates: even when supplemented with energy regeneration systems, a three-gene cascade reduced reporter protein expression by approximately 10^3-fold. Improved design principles or improved cell-free systems may overcome this limitation. Indeed, better understanding of cell-free metabolism is now leading to advances in cell-free translation systems.[9]

Functional polymerases can potentially be selected by directly linking protein function (polymerization) to the gene templates encoding that function, in particular through *in vitro* compartmentalization by emulsification.[10] In a particularly noteworthy demonstration, Holliger and co-workers have emulsified *E. coli* in order to evolve variants of Taq DNA polymerase with novel properties, such as increased substrate specificity.[11] However, in this instance the production of the enzyme is still coupled to a living cell, albeit emulsified.

The completely *in vitro* replication and evolution of a nucleic acid template that encodes the means for its own replication has never before been demonstrated. In order to approach this problem, we started from a T7 RNA polymerase autogene, an expression cassette that encodes a polymerase capable of acting on its own promoter and that thus forms a positive feedback loop between RNA and protein expression.[12] We were able to adapt the autogene to function in a cell-free lysate, emulsify this system, and show that we could get increasing amounts of polymerase and RNA production over time. By carrying out multiple rounds of emulsification, we were further able to show the change in allelic frequency of the polymerase gene over time and hence to confirm that the system is in fact capable of protein evolution on its own, in the absence of any living cell.

■ RESULTS AND DISCUSSION

***In Vitro* Autogene Schema.** We attempted to reproduce the T7 RNA polymerase autogene by a selection for fecundity that in part involved amplification within an emulsion and in part involved amplification outside of the emulsion. The T7 RNA polymerase autogene was modified such that it could be "kick started" by from a second promoter. Amplification would

then proceed by the *in vitro* translation of the initial T7 RNA polymerase mRNA and continued transcription from an embedded T7 RNA polymerase promoter (Figure 1).
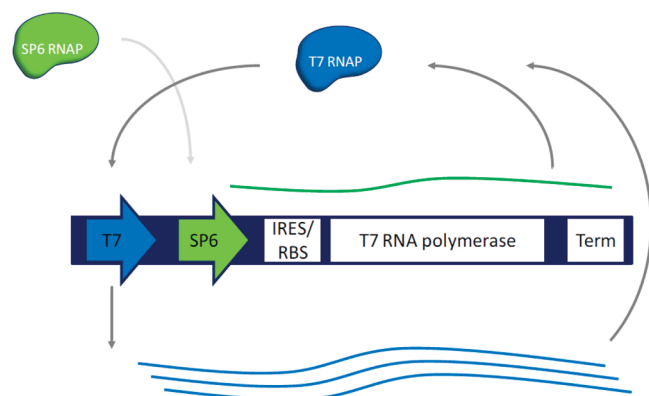

Figure 1. SP6 initiated T7 autogene. In this implementation, the SP6 RNA polymerase is present to initiate transcription and translation of T7 RNA polymerase. The T7 RNA polymerase produced feeds back to the T7 promoter.

Implementation and Optimization To Achieve Self-Amplification. We initially chose to execute this circuit using a commercially available *in vitro* transcription and translation reaction based on an *E. coli* lysate (*E. coli* S30 Extract System for Linear Templates, Promega). The *E. coli*-based system was expected to readily produce a bacteriophage polymerase, and

the choice of a commercial extract helped ensure reproducibility between reactions.

The autogene circuit was initially constructed so that gene expression would be kick-started from a tetA promoter and endogenous *E. coli* RNA polymerase. The production of T7 RNAP was assayed by inclusion of a pT7-GFP reporter. Protein production was found to peak after 20 min of incubation at 30 °C. Unfortunately, mRNA could not be detected from either the initiating *E. coli* promoter or the T7 RNA polymerase promoter, as determined by real time RT-PCR (Figure 2A). Given that protein production was observed, this was likely the result of nuclease degradation of the nascent mRNA. Exploration of a variety of other commercially available *E. coli* lysates, and alternative reaction conditions such as using biotinylated templates, different polymerases, and different means of preparing the template, did not substantially enhance mRNA production or recovery. In the end, it was determined that *E. coli* lysates in general were unsuitable for implementation of the autogene schema.

There were other reports that T7 RNA polymerase could be successfully produced in eukaryotic lysates.[13] We next attempted implementation in the TNT SP6 High-Yield Wheat Germ Protein Expression System (Promega). For this and subsequent formulations, the circuit was changed to allow initiation from the SP6 RNA polymerase promoter via SP6 RNAP. Gratifyingly, the wheat germ lysate was found to yield appreciable levels of not only T7 RNA polymerase protein but also mRNA from both promoters, again as determined by real-time PCR. However, the amount of the mRNA made by
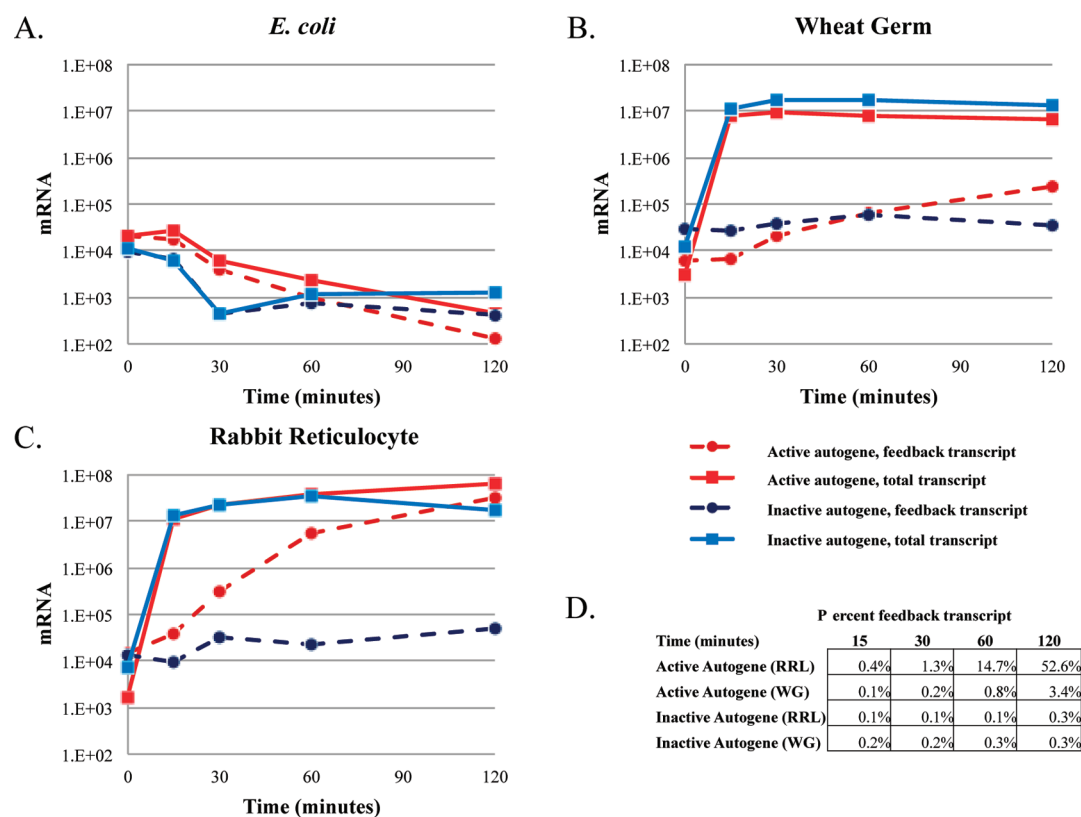

Figure 2. Autogene transcriptional dynamics in lysates. Total transcript (solid) and feedback transcript (dashed) are measured by real time PCR. The active autogene (red) is compared to an inactive version (blue) containing a truncated T7 RNA polymerase gene. Autogene dynamics are investigated in *E. coli* lysate (A), wheat germ lysate (WG; B), and rabbit reticulocyte lysate (RRL; C). Feedback transcription from panels B and C are represented as the percent of total transcription (D).

feedback on the T7 promoter was found to be only 3% as much as the amount from the initiating SP6 promoter (Figure 2B,D). Thus, although mRNA stability is less of a problem in the wheat germ lysate than in the *E. coli* lysate, feedback and amplification were still inadequate.

We finally switched to the TNT SP6 Quick Coupled Transcription/Translation System (Promega), which is based on rabbit reticulocyte lysate (RRL). Bulk TNT reactions with the autogene circuit again yielded protein and detectable transcripts from both promoters. However, in this instance feedback on the T7 RNA polymerase promoter accounted for about 50% of total transcription (Figure 2C,D). This was a significant amount of RNA amplification, and further experiments therefore focused on this reaction mix and template configuration.

Adaptation to Emulsion. While a variety of functional proteins have been produced in emulsions,[10,14−16] there have also been numerous reports of low yields and activities, in part due to interactions between proteins and the hydrophobic oil phase.[17] Ghadessy and Holliger showed that ensconcing the RRL reaction in a mineral oil/4% Abil EM90 emulsion retained as much as 40% of the protein production seen in solution. We therefore chose this formulation for the autogene-RRL system. However, in initial attempts, T7-based transcription was only slightly above background (not shown).

In order to improve expression, several alterations to the template were made, including biotinylation of the 5′ ends, inclusion of an EMCV IRES, increasing the spacing between the T7 promoter and the 5′ end of the template, and incorporating a stabilizing F880Y mutation.[18] With this new construct utilized in bulk RRL, transcription from the T7 promoter accounted for nearly all the total transcript. In emulsion, approximately 40% of the transcription came from the T7 promoter (Figure 3).
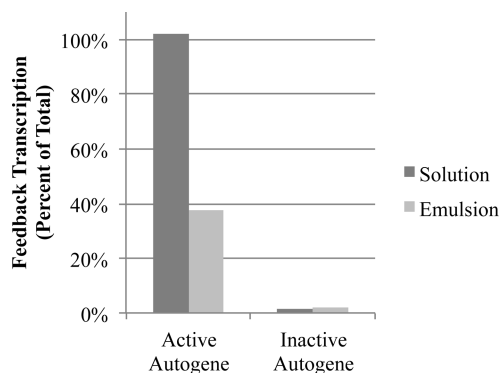

Figure 3. Comparison of solution and emulsion activity of a T7 autogene in rabbit reticulocyte lysate. Transcript levels determine by real time PCR using primer/probe sets that are (feedback) or are not (total) specific to feedback transcripts. Data is an average of two independent experiments.

Selection. Mock selections were carried out with the improved autogene. Two separate emulsion TNT reactions were prepared. The first contained a wild-type version of the SP6-initiated autogene. The second contained a truncated form of the polymerase (engineered to have 2 consecutive, premature stop codons). For each reaction, total RNA was recovered, and a cycle course RT-PCR was performed using primers specific to transcripts originating from the T7 promoter. The active autogene amplified roughly 9 cycles

earlier than the inactive autogene, implying at least a several-hundred-fold selective advantage (Figure 4).
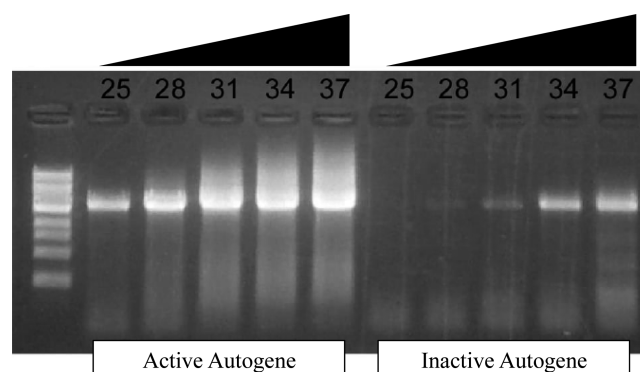

Figure 4. Selective amplification of mRNA. Separate emulsion TNT reactions with either a wild-type autogene or a truncated version were prepared. RNA was recovered and subjected to RT-PCR for varying number of cycles. The active autogene amplified approximately 9 cycles earlier than the inactive version.

The active and inactive templates were then directly competed with one another. A two-member pool of 10% active templates and 90% truncated templates was mixed in RRL and emulsified. Total RNA was recovered, subjected to RT-PCR, and cloned. Of 10 clones, 8 were identified as being from the active autogene (Figure 5). The apparent 8-fold
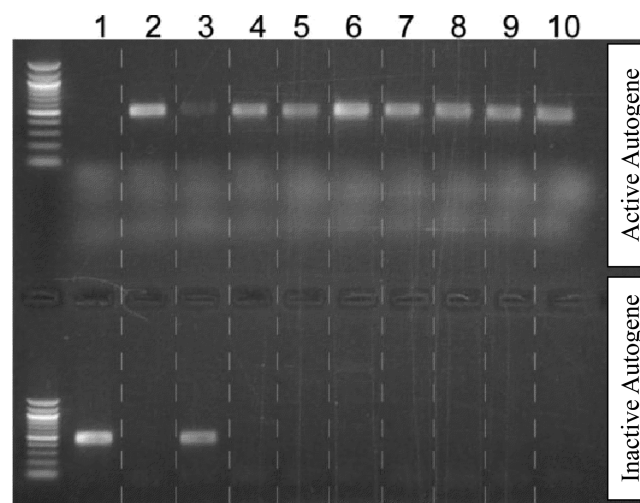

Figure 5. Enrichment of active autogene sequences. Clones from one round of selection from a two-member pool were identified by PCR. The starting pool was 10% active and 90% inactive. Eight of the 10 clones after the selection were from the active autogene. Identification was confirmed by sequencing.

enrichment was smaller than expected but should still be suitable to obtain active clones over multiple rounds of selection and amplification.

However, initial attempts to perform successive rounds of selection were largely unsuccessful. RNA was recovered and DNA templates regenerated over 2 rounds, but in the third round recovery RT-PCR was unsuccessful. Sequencing the products recovered from round 2 recovery products indicated that the system is highly mutagenic. Of 13 unique clones sequenced, 6 were full length and predicted to be potentially

functional (Supplementary Table S1). The 3 kb ORF contained an average of 12 point mutations (9 non-synonymous; n = 6.) This works out to ca. 2 mutations per kb per round. The mutations accumulated were distributed throughout the ORF but were generally predicted to be located on the periphery of the polymerase, away from the active site. It should be noted that the EMCV IRES contained twice as many mutations per kb as the ORF. This suggests that the IRES was either less fit to begin with or less subject to purifying selection.

The remaining 7 clones had large deletions. These deletions were remarkably similar to each other. From the 5′ end, the deletions began early in the EMCV IRES and spanned most of ORF. These deletions likely arose during reverse transcription, with the most plausible scenario being that reverse transcriptase was stalled within the highly structured 3′ end of the ORF (residues 2700−3300 in Figure 6) and the truncated cDNA
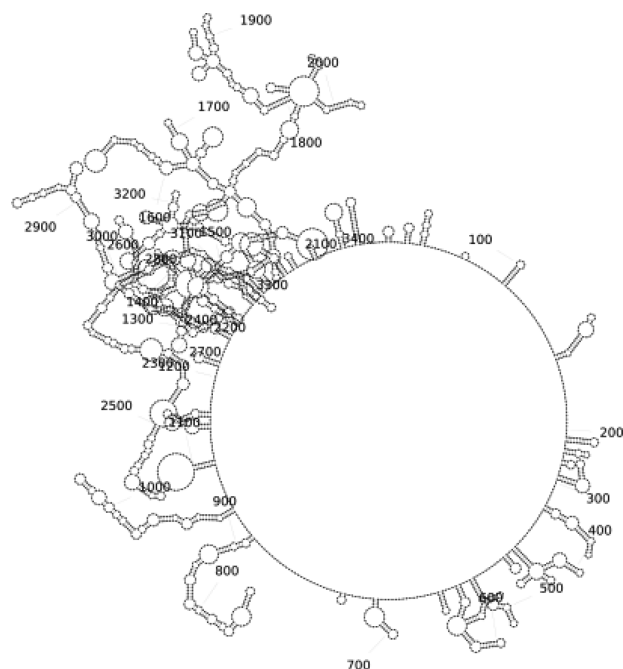

Figure 6. Wild-type T7 RNA polymerase. NUPACK generated structure of the wild-type T7 RNA polymerase mRNA at 55 °C.

reinitiated within the less structured and C-repeat rich IRES (residues 100−150 in Figure 6). The accumulation of these deletion variants is in large measure why recovery failed at round 3.

The takeover of this selection by deletion mutants has precedent in many selections based on replication. Parasitic templates can take advantage of replication size or efficiency to propagate disproportionately to their function. Many laboratory molecular replication processes have generated parasites, most prominently isothermal nucleic acid amplifications such as self-sustained sequence replication (3SR) and Qβ RNA-dependent RNA polymerase amplification. Many classic Qβ experiments resulted in the evolution of progressively minimized replicating RNA, when the Qβ genome was used as a starting template.[2,19] 3SR is modeled after retroviral replication strategies involving two simultaneous reactions: transcription of a DNA template to RNA and reverse transcription of RNA to DNA.[20,21] When 3SR is utilized for the amplification step of in vitro selection of functional nucleic acids, RNA species that replicate efficiently dominate the population, regardless of additional functional

selection.[22] One solution to replication parasites is to increase the complexity required for replication. CATCH is a cooperative replication variant of 3SR. The CATCH amplification scheme requires more than two primer binding events. Ultimately, this did not avoid parasitism but did produce a more complex "parasite".[6]

Similarly, iterative cycles of self-replication of the in vitro autogene generated molecular parasites. Compartmentalization is intended to prevent resource sharing among individuals within the population; however, partitioning of individuals cannot be strictly enforced. It is likely that a combination of imperfect partitioning and the amplification efficiencies of the deletions allowed them to outcompete the full-length functional product. As with CATCH, it is possible that adding an additional phase or gene to the replication cycle could delay or avoid parasite accumulation. We first sought to minimize the most obvious mechanisms for generating deletions.

In light of the above data, steps were taken to reduce the occurrence of mutations and deletions. UV light-aided gel purification was replaced by blue-light aided purification, the PCR steps were carried out with a higher fidelity enzyme, and the IRES and ORF were replaced with sequences that were designed to form much less secondary structure (dubbed T7RSS; Figure 7). With these new features, the scheme was
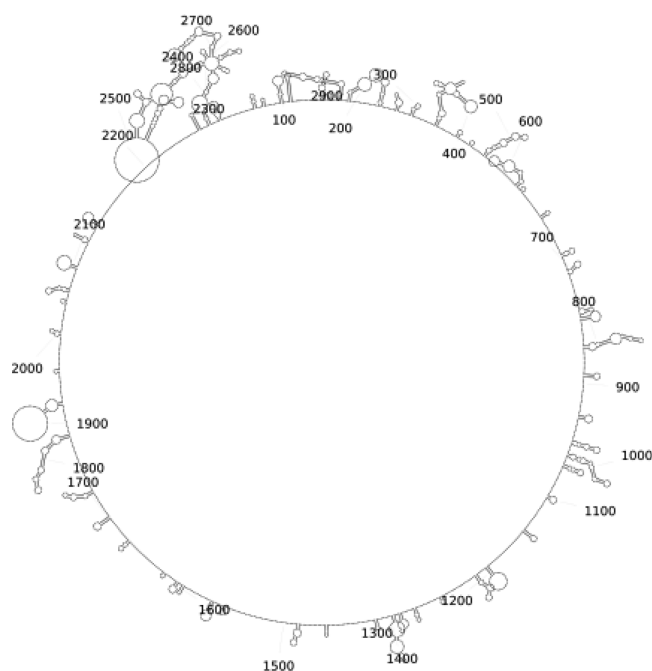

Figure 7. T7 RNA polymerase reduced secondary structure (RSS.) NUPACK generated structure of the destructured T7 RNA polymerase mRNA at 55 °C.

carried out robustly for 4 rounds. Sequencing from each of these optimized rounds did not reveal any of the large scale deletions found in previous experiments. There were also fewer mutations per gene per round (Supplementary Tables S2−S5). However, the number of point mutations, especially synonymous mutations, increased with each round. After one round, there were 0.2 synonymous mutations per kb (20% of total mutations; similar to the predicted 21.7% synonymous mutations by random mutagenesis of the T7-RSS ORF). This increased to 1.3 synonymous mutations per kb (40% of total mutations) by round 4 (Table 1). This bias toward

Table 1. Sequence Results from Successive Rounds of Selection for Wild-Type Function

	round			
	1	2	3	4
point mutations (per kb)	0.8	1.7	2.3	3.4
synonymous mutations (per kb)	0.2	0.6	0.8	1.4
synonymous mutations (% of total)	20	33	35	40
K_a/K_s	1.0	0.5	0.5	0.4
predicted viable sequences (%)	70	67	70	83

synonymous mutations is suggestive of purifying selection.[23] Of the 37 sequences analyzed, only 10 have mutations that would obviously affect function (a stop codon, frame shift, or mutation of a residue in the active site or involved in promoter recognition).

In order to demonstrate the activity of the heavily mutated (~11 amino acid substitutions per gene) variants that remained in the round 4 population, we assayed 8 polymerases for their ability to make RNA (Figure 8). One of the eight clones (R4.3)
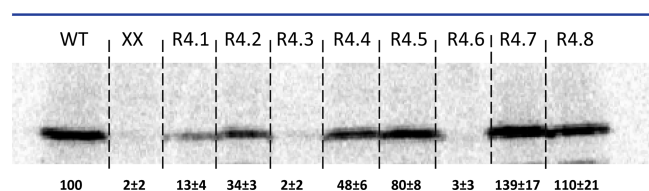

Figure 8. Assay of activity of 8 clones from round 4 in rabbit reticulocyte lysate. Six of the 8 clones show activity comparable to that of wild-type. Numbers shown are averaged from three independent measurements. R4.3 has a frame shift and nonsense mutation. R4.6 has 8 point mutations, but no obviously fatal mutation.

contained a frame-shift and premature stop codon and thus did not produce a full-length polymerase (or demonstrate polymerase activity). For 6 of the remaining 7 clones (all except R4.6), activities roughly comparable to that of the wild-type T7 RNA polymerase were observed. It is not immediately clear which of the 8 non-synonymous mutations in R4.6 ablated activity.

It can be argued that both mutation and selection are roughly constant for each round. The first round was seeded with the wild-type gene, and thus most or all mutations observed after this round likely occurred during the transcription and RT-PCR steps. On the basis of the observation that 30% of the variants that emerged from the first round were inactive, it might be suspected that this is the mutational load for each additional round. The fact that by the conclusion of round 4 we still see roughly 30% inactive variants strongly indicates that at each round inactive polymerases are purged, but that the high mutation rate results in a re-establishment of an extremely mutant population.

The rescue of function by a reduction in the mutation rate suggests that the loss of function is a form of "lethal mutagenesis". Lethal mutagenesis (often erroneously equated with Eigen's "error catastrophe"[24]) is the deterministic extinction of a population from a sustained, high mutation rate — too many mutations occur in each generation for the population to maintain functional genomes.[25] The high mutation rate leads to the accumulation of deleterious mutations, fitness decline, and eventual reduction in viable offspring number to the point that parents are no longer replacing themselves, and the functional genotype in essence goes extinct.

It should be noted that it is possible for a gene population to remain functional under a highly mutagenic regime. For example, the TEM1 lactamase has been shown to retain functionality for at least 18 rounds of mutation and selection (at a rate of 2 mutations per gene per round[23,26]). Thus, the persistence or extinction of a drifting population is trade-off between the mutation of active to inactive variants and the fecundity of active variants. The extinction threshold has been derived in simple cases to be

$$f \exp(-U_d) < 1$$

where U_d is the genomic deleterious mutation rate, and f is the "fecundity" of the mutation-free genome. In our *in vitro* system, fecundity is the average number of emulsion bubbles carrying one or more RNA molecules derived from a single bubble with a wild-type RNAP molecule in the previous cycle. U_d is the deleterious mutation rate, the average number of deleterious mutations per (functional) RNA molecule in cycle 2 from a parent molecule in cycle 1. While we do not know f, especially because of the likely heterogeneity of the emulsion reaction bubbles, it can be estimated. On the basis of our results, it appears that U_d can be approximated as 0.3, which in turn implies that on average f must have been less than 1.4, that is, each functional polymerase carried less than 1.4 copies of itself into a subsequent generation. While this value is obviously low, it is consistent with the large technical difficulties encountered in obtaining polymerase function in the emulsion-based transcription and translation system. Interestingly, the fact that the system is going extinct does not prevent us from identifying novel, functional polymerases, as we have shown. This is especially true given that deep (NextGen) sequencing can now be used to sort through entire artificial phylogenies of polymerases, looking for conserved functional mutations against a large background of deleterious mutations.

This analysis does provide insight into how to better preserve function and avoid extinction into the future. One method would obviously be to lower the mutation rate, although given that most RNA polymerases are inherently error-prone this may be difficult. The other solution would of course be to increase fecundity, the number of mRNAs made per emulsion bubble. As we continue to improve the system, it should prove possible to not only select functional polymerases but also to promote their continued evolution.

■ METHODS

Template Preparation. Plasmids were generated in which the SP6 promoter was placed upstream of either the wild-type T7 RNA polymerase or a truncated version with two premature stop codons. The T7 RNA polymerase promoter was not included in the plasmid, due to toxic effects on the host. PCR was then carried out with primers designed to extend the template to include the T7 RNA polymerase promoter, a T7 terminator, and 5′ biotins. PCR was carried out with either Platinum Taq High Fidelity (Invitrogen), Accuprime PFX (Invitrogen), or PFU Ultra II (Agilent). Intermediate PCR products were gel purified using Wizard Gel and PCR Clean-Up System (Promega). The final PCR product was further purified by phenol/chloroform extraction. In short, 100 μL of PCR reaction was mixed with 100 μL 25:24:1 phenol/chloroform/isoamylalcohol pH 7.9 (Ambion). After centrifugation, the aqueous phase was removed and mixed with 100 μL of chloroform. After centrifugation, the aqueous phase was removed and mixed with 10 μL of 3 M sodium acetate and 300

μL of 100% ethanol. The solution was stored at −20 °C for 30 min and then centrifuged at 13,000 × g at 4 °C for 30 min. The supernatant was removed and the pellet was washed with 70% ethanol before being air-dried and resuspended in 50 μL of water. This yield of PCR product was quantified on an agarose gel using DNA quantitation standards.

In Vitro Transcription and Translation Reactions in Solution. Several different *in vitro* transcription and translation reactions were performed, each according to manufacturer's specifications. Specifically, the perfomance of *E. coli* S30 Extract System for Linear Template (Promega), TNT SP6 High-Yield Wheat Germ Protein Couple Expression System (Promega), and TNT SP6 Quick Coupled Transcritpion Translation System (Promega) were assessed. Reactions contained 100 ng (∼3 × 10^9 molecules) of full-length autogene template in 50 μL reaction. Reactions were assembled on ice and incubated at 30 °C for up to 2 h.

In Vitro Transcription and Translation Reactions in Emulsion. Individual tubes (Sarstedt, 95 × 16.8 mm polypropylene) were set up to contain 96 μL of mineral oil and 4 μL of Abil EM90 (Evonik) and placed on ice. A 50 μL *in vitro* transcription and translation reaction was prepared as described above. All the reagents were kept on ice to prevent premature initiation of transcription and translation, and the template was added immediately before emulsification. In general, 0.34 ng (∼1 × 10^8 molecules) to 34 ng (∼1 × 10^10 molecules) was added to a given emulsified reaction. A stir bar (Spinplus 9.5 × 9.5 mm Teflon) was added to the tube containing the oil/surfactant mixture. The tube was moved into a beaker containing ice water on top of a magnetic stir plate (Corning) and stirred on the "high" setting for 1 min (1150 rpm). While the oil/surfactant mixture was stirred, the 50 μL *in vitro* transcription and translation reaction was added drop-by-drop over 1 min and then stirred for an additional 4 min. The fully emulsified reaction was incubated at 30 °C for an additional 2 h.

The emulsion reaction was stopped by placing the tube on ice for 10 min and adding Stop Solution (50 mM Tris, pH 7.0, 50 mM EDTA) sufficient to bring the aqueous volume to 100 μL total. The emulsion was broken by vortexing the reaction with 1x volume chloroform. The reaction was centrifuged at 13,000 × g to separate the aqueous phase from the organic phase. A waxy plug composed of the emulsifying reagents and proteins forms between the aqueous layer and the organic phases at this step. The aqueous phase was removed for further purification.

mRNA Purification. The entire transcription-translation reaction (for bulk solution reactions) or the aqueous phase (after breaking an emulsion reaction as above) was added to 400 μL of Trizol reagent (Invitrogen). The tube was shaken and incubated for 2 min at room temperature, 150 μL of chloroform was added, and the tube was again shaken and incubated for 2 min to allow the phases to separate. The reaction was centrifuged at 13,000 × g at room temperature for 2 min. The aqueous phase was removed and precipitated by addition of 1 μg of glycogen and 0.7x volume of isopropyl alcohol. After incubation at room temperature for 15 min the reaction was again centrifuged at 4 °C for 30 min. The supernatant was removed, and the pellet was washed with 70% ethanol before being air-dried.

The pellet was resuspended in 43 μL of water and 0.1x volume 10x DNase buffer, and then 2 U Turbo DNase (Ambion) and 40 U RNasin Plus RNase inhibitor (Promega)

were added, bringing the final volume to 50 μL. The DNA digestion reaction was incubated at 37 °C for 30 min, and then 0.1x volume DNase Inactivation Reagent was added. The solution was mixed gently for 2 min, and the DNase Inactivation Reagent was removed by centrifugation at 13,000 × g for 2 min at 4 °C. The aqueous supernatant was removed and could be stored at −80 °C or carried into further amplification reactions.

mRNA Quantitation. The purified RNA was converted to cDNA by reverse transcription with SuperScript III reverse transcriptase (Invitrogen) according to manufacturer's specifications. Some 5 μL (∼10%) of the total RNA preparation was reverse-transcribed in a 20 μL reaction using 2 pmol of a reverse primer that annealed to the 3′UTR of the autogene. The reaction was incubated at 55 °C for 1 h, and then at 70 °C for 15 min.

The relative amount of mRNA recovered was evaluated by real time PCR with a 7300 Real Time PCR System (Applied Biosystems). Reaction conditions were 0.5x volumes 2x FastStart Universal Probe Master with Rox (Roche), 0.9 μM forward and reverse primer (each), and 0.25 μM Taqman probe in a 25 μL reaction. The reverse transcription reaction was diluted from 20 μL to 100 μL with water, and then 10 μL (10%) was added as input to a real-time PCR reaction. The thermal cycling parameters were (1) 2 min at 50 °C; (2) 10 min at 95 °C; and (3) 40 cycles of 15 s at 95 °C, then 1 min at 60 °C. Reactions were set up in triplicate. Each autogene sample was analyzed by real-time PCR with at least two separate real-time primer-probe sets in separate reactions. One real-time primer-probe set, referred to as the "total transcription" primer-probe set, specifically amplified a sequence contained wholly within the T7 RNA polymerase gene. This amplicon will be present in both transcripts produced from the SP6 RNA polymerase promoter and the upstream T7 RNA polymerase promoter. The second real-time primer-probe set, referred to as the "feedback transcription" primer-probe set, specifically amplified a sequence present only in transcripts initiated from the T7 RNA polymerase promoter.

Template Regeneration. Reverse transcription and PCR amplification of full length mRNAs were performed in a single-step reaction using Superscript III one step RT-PCR System with Platinum Taq High Fidelity (Invitrogen). The RT-PCR product was gel purified and used for regeneration of the autogene template. This protocol is the same as described in Template Preparation. Briefly, there were two regenerative PCR steps. The first PCR product was gel-purified, and the second product was phenol/chloroform-extracted and ethanol-precipitated. The PCR steps add back the T7 promoter, 3′UTR, and 5′ biotins.

Cloning and Sequencing. Gel-purified RT-PCR products were amplified using primers that added appropriate restriction sites for cloning into a pASK sequencing vector. Some 5−10 clones from each round were sequenced using multiple primers that spanned the full length of the gene. Sequences were assembled and analyzed using BioEdit (Ibis Biosciences) and Geneious (Biomatters). K_a/K_s was determined as the observed ratio of non-synonymous mutations to synonymous mutations normalized to the ratio expected from random chance.

T7 RNA Polymerase Activity Assay. A SP6 promoter and EMCV IRES were appended to the genes for T7 RNA polymerase variants isolated from round 4 via overlap PCR. DNA templates were gel-purified, and 3 × 10^9 templates were mixed with RRL, 200 pM ^35S methionine (1175 Ci per mmol;

Perkins Elmer), and a plasmid containing the T7 promoter, EMCV IRES, and GFP gene. The reaction mixtures were incubated at 30 °C for 2 h, followed by addition of an equal volume of Stop Solution (50 mM Tris, pH 7.0, 50 mM EDTA), 10 mM DTT, and 4x LDS loading dye (Invitrogen). Samples were developed on a 4−12% NuPAGE gel (Invitrogen) with 1x MOPS, transferred to a 0.45 μM nitrocellulose membrane (Invitrogen), and exposed to a storage phosphor screen (Molecular Dynamics) before imaging on a STORM 840 Phospoimager (GE Healthcare). Autoradiographs were analyzed and quantified using ImageQuant (GE Healthcare). A band at 100 kDa indicated proper expression of the T7 RNA polymerase variants, while a band at 27 kDa indicated expression of GFP and thus T7 RNA polymerase activity.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Sequences for T7 RNA polymerases recovered through selection. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
*E-mail: andy.ellington@mail.utexas.edu.

Notes
The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Haruna, I., and Spiegelman, S. (1965) Autocatalytic synthesis of a viral RNA in vitro. *Science 150*, 884−886.

(2) Mills, D., Peterson, R., and Spiegelman, S. (1967) An extracellular Darwinian experiment with a self-duplicating nucleic acid molecule. *Proc. Natl. Acad. Sci. U.S.A. 58*, 217−224.

(3) Saffhill, R., Schneider-Bernloehr, H., Orgel, L. E., and Spiegelman, S. (1970) In vitro selection of bacteriophage Q-beta ribonucleic acid variants resistant to ethidium bromide. *J. Mol. Biol. 51*, 531−539.

(4) Ellington, A. D., and Szostak, J. W. (1990) In vitro selection of RNA molecules that bind specific ligands. *Nature 346*, 818−822.

(5) Tuerk, C., and Gold, L. (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science 249*, 505−510.

(6) Ellinger, T., Ehricht, R., and McCaskill, J. S. (1998) In vitro evolution of molecular cooperation in CATCH, a cooperatively coupled amplification system. *Chem. Biol. 5*, 729−741.

(7) Noireaux, V., Bar-Ziv, R., and Libchaber, A. (2003) Principles of cell-free genetic circuit assembly. *Proc. Natl. Acad. Sci. U.S.A. 100*, 12672−12677.

(8) Isalan, M., Lemerle, C., and Serrano, L. (2005) Engineering gene networks to emulate Drosophila embryonic pattern formation. *PLoS Biol. 3*, e64.

(9) Jewett, M. C., Calhoun, K. A., Voloshin, A., Wuu, J. J., and Swartz, J. R. (2008) An integrated cell-free metabolic platform for protein production and synthetic biology. *Mol. Syst. Biol. 4*, 220.

(10) Tawfik, D. S., and Griffiths, A. D. (1998) Man-made cell-like compartments for molecular evolution. *Nat. Biotechnol. 16*, 652−656.

(11) Ghadessy, F. J., Ong, J. L., and Holliger, P. (2001) Directed evolution of polymerase function by compartmentalized self-replication. *Proc. Natl. Acad. Sci. U.S.A. 98*, 4552−4557.

(12) Dubendorff, J. W., and Studier, F. W. (1991) Creation of a T7 autogene. Cloning and expression of the gene for bacteriophage T7 RNA polymerase under control of its cognate promoter. *J. Mol. Biol. 219*, 61−68.

(13) Finn, J., Lee, a C. H., MacLachlan, I., and Cullis, P. (2004) An enhanced autogene-based dual-promoter cytoplasmic expression system yields increased gene expression. *Gene Ther. 11*, 276−283.

(14) Griffiths, A. D., and Tawfik, D. S. (2003) Directed evolution of an extremely fast phosphotriesterase by in vitro compartmentalization. *EMBO J. 22*, 24−35.

(15) Levy, M., and Ellington, A. D. (2008) Directed evolution of streptavidin variants using in vitro compartmentalization. *Chemi. Biol. 15*, 979−989.

(16) Doi, N., Kumadaki, S., Oishi, Y., Matsumura, N., and Yanagawa, H. (2004) In vitro selection of restriction endonucleases by in vitro compartmentalization. *Nucleic Acids Res. 32*, e95.

(17) Ghadessy, F. J., and Holliger, P. (2004) A novel emulsion mixture for in vitro compartmentalization of transcription and translation in the rabbit reticulocyte system. *Protein Eng., Des. Sel. 17*, 201−204.

(18) Akio Sugiyama, T.; Yoshiaki Nishiya, T.; Bunsei Kawakami, F.-K. RNA polymerase mutants with increased themostability. U.S. Patent US 7,507,567 B2, March 24, 2009.

(19) Kacian, D. L., Mills, D. R., Kramer, F. R., and Spiegelman, S. (1972) A replicating RNA molecule suitable for a detailed analysis of extracellular evolution and replication. *Proc. Natl. Acad. Sci. U.S.A. 69*, 3038−3042.

(20) Guatelli, J. C., Whitfield, K. M., Kwoh, D. Y., Barringer, K. J., Richman, D. D., and Gingeras, T. R. (1990) Isothermal, in vitro amplification of nucleic acids by a multienzyme reaction modeled after retroviral replication. *Proc. Natl. Acad. Sci. U.S.A. 87*, 1874−1878.

(21) Compton, J. (1991) Nucleic acid sequence-based amplification. *Nature 350*, 91−92.

(22) Breaker, R. R., and Joyce, G. F. (1994) Emergence of a replicating species from an in vitro RNA evolution reaction. *Proc. Natl. Acad. Sci. U.S.A. 91*, 6093−6097.

(23) Bershtein, S., Segal, M., Bekerman, R., Tokuriki, N., and Tawfik, D. S. (2006) Robustness-epistasis link shapes the fitness landscape of a randomly drifting protein. *Nature 444*, 929−392.

(24) Eigen, M. (1971) Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften 58*, 465−523.

(25) Bull, J. J., Sanjuán, R., and Wilke, C. O. (2007) Theory of lethal mutagenesis for viruses. *J. Virol. 81*, 2930−2939.

(26) Bershtein, S., Goldin, K., and Tawfik, D. S. (2008) Intense neutral drifts yield robust and evolvable consensus proteins. *J. Mol. Biol. 379*, 1029−44.