A PUBLICATION OF THE RNA SOCIETY

Selection of an improved RNA polymerase ribozyme with superior extension and fidelity

Hani S. Zaher and Peter J. Unrau

References	This article cites 39 articles, 14 of which can be accessed free at: **http://rnajournal.cshlp.org/content/13/7/1017.full.html#ref-list-1** Article cited in: **http://rnajournal.cshlp.org/content/13/7/1017.full.html#related-urls**
Email alerting service	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here**

Selection of an improved RNA polymerase ribozyme with superior extension and fidelity

HANI S. ZAHER and PETER J. UNRAU

Department of Molecular Biology and Biochemistry, Simon Fraser University, BC Canada, V5A 1S6

ABSTRACT

Our current understanding of biology suggests that early life relied predominantly on RNA for catalysis and replication. Here, we report the isolation of an RNA polymerase ribozyme called B6.61 that exhibits superior extension and fidelity relative to its progenitor, the Round-18 polymerase. The B6.61 polymerase was selected from a mutagenized pool containing $\sim 9 \times 10^{14}$ sequence variants through the use of a novel large-scale in vitro compartmentalization system. B6.61 polymerized all tested primer–template (PT) complexes faster than the Round-18 variant. For one PT, B6.61 exhibited dramatically faster elongation past one full helical turn and incorporated at least 20 nucleotides of sequence, setting a new extension record for an RNA polymerase ribozyme. The increased efficiency of the B6.61 construct was related to improvements in fidelity, with the new variant incorporating less incorrect wobble base pairs than its parent. This new polymerase demonstrates the feasibility of evolving an artificial RNA replicase ribozyme in the foreseeable future.

Keywords: ribozyme; polymerase; replication; in vitro compartmentalization; RNA world

INTRODUCTION

The RNA world hypothesis suggests that modern metabolism evolved from a system where RNA served the genomic role now filled by DNA and the catalytic role presently performed by protein (Gilbert 1986). The plausibility of this model has been strengthened by the discovery that RNA can catalyze a diverse range of chemical reactions (Wilson and Szostak 1999; Joyce 2002; Fiammengo and Jaschke 2005) and that functional RNAs play critical roles in modern metabolism, most notably the ribosome, RNaseP, and the riboswitches (Guerrier-Takada et al. 1983; Ban et al. 2000; Winkler and Breaker 2005). Although RNA machines like the ribosome are believed to be relics of an RNA-based past, the current dominance of protein catalysts prevents us from learning how RNA-based life might have replicated. Nevertheless, one can imagine constructing a minimal autoreplicative system using RNA catalysts confined within small model compartments, in analogy to a modern cell (Bartel and Unrau 1999).

Arguably the greatest limitation of the RNA world hypothesis is the lack of an RNA replicase ribozyme, which has been coined the holy grail in the ribozyme community (Szostak et al. 2001; McGinness and Joyce 2003). Guided by the catalytic mechanisms utilized by protein polymerases, attempts have focused on isolating ribozymes that are capable of polymerizing a primer in a template-dependent manner as a key first step in the creation of a replicase ribozyme (Been and Cech 1988; Doudna and Szostak 1989; Bartel et al. 1991; Chowrira et al. 1993; Doudna et al. 1993; McGinness and Joyce 2002). No approach has been more successful than the efforts that resulted in the 189-nucleotide (nt)-long Round-18 polymerase, which was selected from a pool built by appending a 76-random-nucleotide accessory sequence to the class I ligase (Bartel and Szostak 1993; Ekland and Bartel 1995, 1996; Ekland et al. 1995; Johnston et al. 2001). This polymerase can extend an RNA primer-template (PT) complex by up to 14 nt in a template-directed manner after 24 h of incubation. Nonetheless, the ability of the Round-18 ribozyme to polymerize more than one RNA helical turn is limited; more typically the Round-18 polymerase adds only a few nucleotides to a given PT complex. This poor polymerization ability has been attributed to weak and highly variable PT recognition (Lawrence and Bartel 2003), which in turn is related in a

---

complex fashion to the fidelity of the enzyme. Although the fidelity is high if the polymerase successfully extends a primer by as much as 11 nt (Johnston et al. 2001), this measurement does not take into account aborted extension products that terminate due to a misincorporation of a nucleotide. Therefore, the efficiency of this polymerase must be greatly improved in order to create an RNA system with true autoreplicative potential. Achieving this goal will require a complex combination of improved PT recognition, processivity, fidelity, and NTP utilization. As these polymerase properties are interdependent, it will be increasingly difficult to optimize these properties simultaneously using standard selection methods. Indeed, recent attempts to isolate superior polymerase ribozymes from the same pool that was used to isolate the Round-18 polymerase using conventional in vitro selection techniques have not yet proven fruitful (Lawrence and Bartel 2005).

A key problem associated with conventional selection is that it does not directly select for true *trans*-acting enzymes. Instead, ribozymes are selected based on self-induced chemical modifications that result from a *cis*-dependent reaction. Thus, the selected sequences are not true catalysts, as they do not perform catalysis in *trans*. This problem has been surmounted for some ribozymes by breaking the *cis* construct into two modules, one being the catalyst and the other acting as a substrate (Fiammengo and Jaschke 2005). The Round-18 ribozyme was isolated precisely this way, with the PT tethered to the ribozyme, and as a result, a substantial loss of activity was observed when the PT was broken away from the polymerase (Johnston et al. 2001).

Darwinian evolution has solved the problem of correlating genotype to phenotype by utilizing cellular membranes to ensure that the benefits of an advantageous gene are not lost to the external environment. Although cells have been used for the selection of nucleic acids (Ferber and Maher 1998; Soukup and Maher 1998; Zimmerman and Maher 2002; Buskirk et al. 2003; Wadhwa et al. 2004), they offer little or no control over selection conditions, and the diversity of the starting library is very limited. Recently, in vitro compartmentalization (IVC) has been used as an alternative for conventional in vitro selection in the direct isolation of *trans*-ribozymes performing multiple-turnover chemistries (Agresti et al. 2005; Levy et al. 2005). The methodology, which was first described for protein selections (Tawfik and Griffiths 1998; Miller et al. 2006), allows the coupling of genotype (DNA rather than RNA) and phenotype by compartmentalizing both within water-in-oil droplets but with nucleic acid diversities only slightly higher that those possible with cellular systems, and 10^3–10^4 times smaller than is possible with conventional in vitro selection where a typical sequence diversity of 10^{15} can be sampled (Agresti et al. 2005; Fiammengo and Jaschke 2005; Levy et al. 2005).

In this study we have developed a large-scale IVC system to select for an improved RNA polymerase ribozyme from a mutagenized library based on the Round-18 ribozyme and with a diversity of $\sim 9 \times 10^{14}$ sequences. After six rounds of selection, where ribozyme polymerases were required to act in a completely *trans*-fashion, we isolated a sequence that had extension properties superior to the Round-18 ribozyme in every assay tested.

RESULTS AND DISCUSSION

Pool construction

The Round-18 polymerase, called WT hereafter, is the only ribozyme that has been shown to exhibit general RNA-templated primer extension ability. In the presence of a particular PT complex, the WT can polymerize a full turn of an RNA helix at 200 mM Mg^{2+} and pH 8.5 after 24 h of incubation (Johnston et al. 2001). However, its polymerization activity is extremely sensitive to the PT sequence, with the addition of a single nucleotide to either the primer or the template sequence often having a dramatic and unpredictable effect on activity (Lawrence and Bartel 2003). With the aim of improving ribozyme processivity and fidelity, we decided to select for improved variants of the WT from a high diversity doped pool (Fig. 1A) using a large-scale compartmentalized selection scheme.

The ligase core of the ribozyme has been previously shown to play an integral role in phosphodiester bond formation and has been found to contain highly conserved sequence elements (Johnston et al. 2001). As a result, two sequences corresponding to the ligase core having no mutagenesis (L_W) or a mutagenesis level of 3% (L_M) were synthesized and converted into double-stranded DNA using large-scale PCR (Zaher and Unrau 2005). A split-pool synthesis was utilized for both constructs so as to incorporate random nucleotides of the form $(NN)_{0-5}$ at the 5'-end of the ribozyme (Fig. 1A; Zaher and Unrau 2005). This random insertion provided sequence diversity in a region of the polymerase that had remained fixed ever since the initial characterization of the class I ligase (Ekland and Bartel 1995), largely due to limitations of the conventional selection scheme. In contrast to the ligase core, much less is known about the role of the accessory domain (the auxiliary domain appended to the ligase core during the polymerase selection that confers polymerization ability to the ligase core). Therefore, two accessory double-stranded DNA sequences having mutagenesis levels of either 0% (A_W) or 10% (A_M) were synthesized (Zaher and Unrau 2005). The ligase core sequences were designed to contain a BtgI restriction site upstream of a T7 promoter and an EarI site at their 3'-end. The accessory pool elements contained an EarI site at their 5'-end, allowing them to be ligated to the ligase core pool elements with only 3 nt of fixed sequence (Fig. 1A) after restriction digestion of both with EarI. Pool fragments were ligated to produce a full-length pool using 0.75 nmol of $L_M A_M$, 0.375 nmol of $L_W A_M$, and
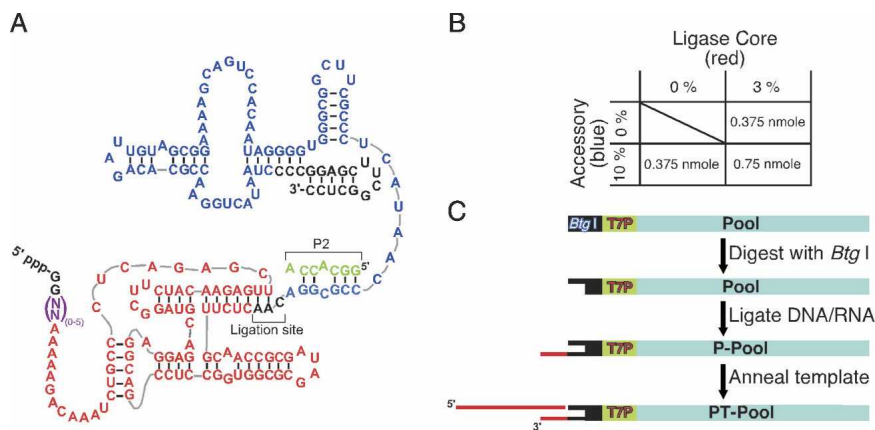
FIGURE 1. Pool design and construction. (*A*) Secondary structure of the Round 18 polymerase, which was used to construct the starting pool. The ligase core is shown in red, while the accessory domain is depicted in blue, and the P2 oligonucleotide, found to be important for the WT polymerase ribozyme, is drawn in green. Black residues indicate fixed sequence that was required for transcriptional initiation, PCR amplification, or ligation of the two domains after EarI digestion. A random insert of variable length was also introduced, $(NN)_{0-5}$, 3′ to the pair of initiating guanosine residues (magenta). (*B*) The starting pool was made from three ligation reactions using the molar amounts shown. (*C*) Pool processing for the selection (T7P indicates T7 promoter). The DNA pool was digested with BtgI, leaving a 4-nt overhang that was used to attach the single-stranded selection primer using T4 DNA ligase (P-pool). This DNA–RNA hybrid (RNA in red) was then annealed to an RNA template in preparation for the selection process (PT-pool).

0.375 nmol of $L_M A_W$ (Fig. 1B). The final pool was then digested with BtgI and ligated to the D11P12 DNA/RNA primer element (Table 1) using T4 DNA ligase. The pool was then annealed to a twofold excess of the selection template T21 (Table 1) to produce the initial PT-DNA pool (Fig. 1C).

Large-scale emulsion

IVC has been previously used to isolate RNA ligase and Diels-Alderase ribozymes from libraries having diversities of $\sim 10^9$ and 10^{11}, respectively, using emulsion volumes of 0.5–1 mL (Agresti et al. 2005; Levy et al. 2005). We increased the emulsion volume to 3000 mL, and by having on average 30 genomes per compartment (weighted by volume distribution), the total pool diversity was expanded to $\sim 9 \times 10^{14}$. The large-scale emulsion was made in 50 mL falcon tube batches, using glass beads and a vortexer (Fig. 2A). This procedure resulted in compartments having radii typically from 0.5 to 1.5 μm (volumes of 5×10^{-16} to 1.5×10^{-14} L), with a weighted average volume of 5×10^{-15} L as judged by microscopy (Fig. 2B) and sedimentation experiments (data not shown). Under these conditions, T7 RNA polymerase (T7 RNAP) was functional with an activity threefold lower than that of a nonemulsified reaction (Fig. 2C). The free Mg^{2+} concentration during the incubation was optimized to balance the requirements of both T7 RNAP, which is inhibited by high magnesium levels, and ribozyme activity, which in this case is optimal

at high magnesium concentrations. We found that when the free Mg^{2+} concentration (excess over the total NTPs concentration of 14 mM) was set at 36 mM, emulsified transcription yielded ~ 200 RNA transcripts per DNA sequence (with the DNA substrate below the enzymes ~ 16 nM affinity for the T7 promoter [Zaher and Unrau 2004], 3 h incubation, 10 U/μL T7 RNAP). Under the same conditions, the WT polymerase was able to extend the untethered selection PT (D11P12:T21) complex by up to 6 nt. Furthermore, when this primer was ligated to the WT DNA sequence (PT-WT), it was extended by 4 nt when incubated with either gel-purified WT ribozyme or T7 RNAP (Fig. 3A). When the same primer was ligated to the initial DNA pool, extension was observed in the presence of gel-purified WT ribozyme but not when T7 RNAP was added, demonstrating that functional ribozyme polymerases resulting from transcription were required to mediate the extension of the PT-DNA complex. Overall these conditions imply that if a single functional DNA sequence was in a compartment, the transcribed ribozymes resulting from this sequence (~ 200) would be approximately sevenfold in excess of all PT-DNA substrates found in that particular compartment (~ 30).

Selection of the superior B6.61 polymerase

The primary selective step involved the isolation of ribozyme polymerase-encoding genomes possessing correctly extended PT complexes. After incubating the emulsion, the aqueous fraction was recovered, the nucleic acid was denatured, and an excess of a biotinylated capture probe, having a sequence similar to the template used in the PT complex, was added (Table 2). Upon renaturation, this

TABLE 1. PT sequences used in extension assays

Oligo name	Sequence
D11P12	5′d(CACGAGCCACG)CUGCCAACCGUG3′
D4P12	5′d(CACG)CUGCCAACCGUG3′
P7	5′CUGCCAA3′
T21	3′GACGGUUGGCACGCUUCGCAG5′
T31	3′GACGGUUGGCACGCUUCGCAGCCCCCCCCGG5′
T26	3′GACGGUUGGCACGCCAUCGGACUAUG5′

D, P, and T denote deoxynucleotide, primer, and template, respectively.
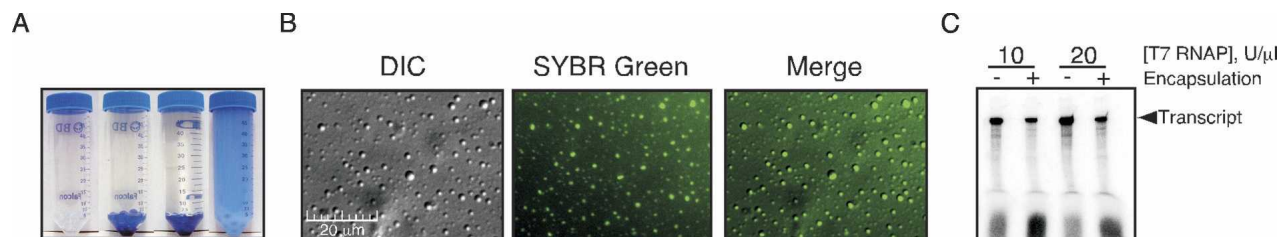
FIGURE 2. Large-scale emulsion manufacture. (*A*) Photographs of the steps used to make 50 mL of emulsion. First (*left* to *right*), DEPC-treated glass beads are added into a 50 mL Falcon tube; second, the aqueous phase (bromophenol blue is included for visualization) is added; third, the oil-phase is poured on top of the aqueous phase before the whole mixture is vortexed to make the water-in-oil droplets. (*B*) Microscope images of emulsion containing RNA (180-nt-long transcripts at 5 μM) that was stained with SYBR green. Fluorescence resulting from 450 to 490 nm excitation colocalized with the DIC image. The droplets are 1–3 μm in diameter. (*C*) A denaturing polyacrylamide gel showing that emulsified transcription takes place with a yield one-third that of the normal one.

material was applied to streptavidin magnetic beads and washed so as to retain only those DNA pool sequences whose RNA primers had been correctly extended by a polymerase ribozyme. These retained DNA genotypes were then eluted and amplified by PCR. This new pool was subsequently digested using BtgI and ligated to the selection DNA/RNA primer (Table 1) before undergoing another Round of selection. A purification factor of 100 could easily be obtained by this hybridization approach (Fig. 4).

A total of seven rounds of selection labeled as branch A (Table 2) were performed and resulted in the isolation of parasitic DNA sequences containing internal and presumably enzymatically incorporated multiple repeats of the sequence 5′-*TTTGGACGCTTCGCACGGTTG*CAGD (cap-

ture probe sequence in italics D: A, G, or T) that allowed the bottom DNA strand to be captured during the selective step. An extensive search found two variants (A7.15 and A7.29) with catalytic activity comparable to the WT together with 10 WT sequences, suggesting that superior catalysts might be found by further selection. A second selection step was therefore added to increase stringency and was used for Rounds 4–6 of a B branch of selection. The new Rounds of selection were performed by adding 0.5 mM ^{4S}UTP to the transcription mix (Fig. 5; Table 2). This nucleotide, when incorporated into the WT polymerase by transcription, did not inhibit activity and was also found to be a substrate for the WT ribozyme (Fig. 3B). PT-pool sequences containing ^{4S}UMP in the PT element could be
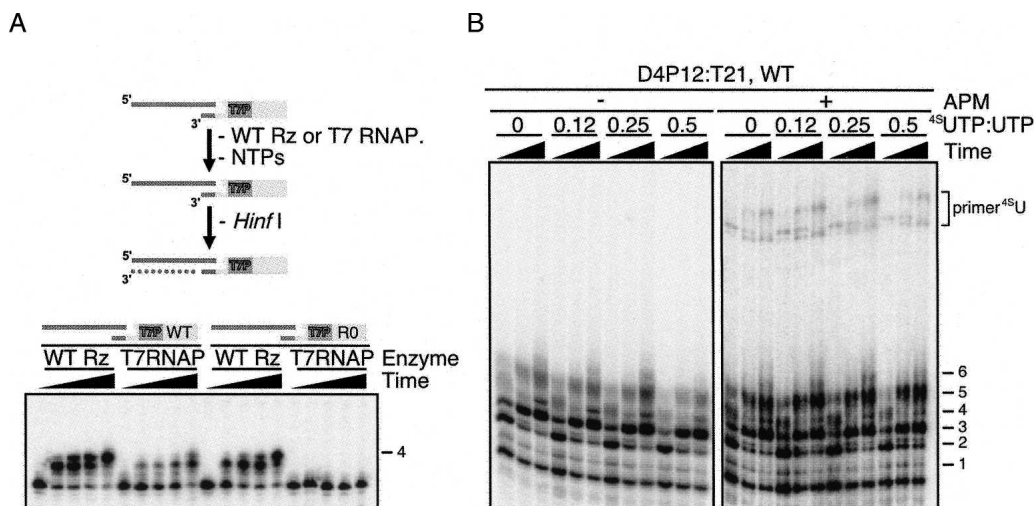
FIGURE 3. The WT polymerase is active in the selection context. (*A*) The PT-WT or the PT-R0 genome constructs were incubated with gel-purified WT (2 μM final concentration) ribozyme polymerase or T7 RNAP, and incubated in the selection buffer. Time points were taken at 0, 1, 2, 4, and 20 h, and the DNA genomes were subsequently digested with HinfI to improve gel resolution. Both constructs were extended when the WT ribozyme was added, but only the PT-WT was extended in the presence of T7 RNAP, indicating that transcripts produced by T7 RNAP were catalytically active. (*B*) The WT ribozyme was transcribed in the presence of varying ^{4S}UTP concentrations (relative to 2 mM UTP which was always present) and gel purified. The four resulting ribozymes were then assayed for their ability to extend PT(D4P12:T21) in the selection buffer supplemented with 0.5 mM ^{4S}UTP. Time points were taken at 0.5 h, 4.5 h, and 24 h and resolved on denaturing PAGE plus or minus APM. Ribozymes transcribed in the presence of 1 mM ^{4S}UTP and 2 mM UTP were almost as active as ribozymes that were transcribed in the absence of ^{4S}UTP. All ribozymes were capable of incorporating ^{4S}UMP as indicated by the mercury-dependent gel shift shown in the *right* panel (marked as primer ^{4S}U in the figure).

TABLE 2. Selection conditions

Round	Capture Probe	^{4S}UTP	P2	[Genome] nM[a]	Emulsion volume	Inc. time, h
AI	d(TTB$_T$GGACGCTTCGCACGGTTG)	−	+	10	3000 mL	O/N
AII	d(TTB$_T$GGACGCTTCGCACGGTTG)	−	+	10	300 mL	O/N
AIII	d(TTB$_T$GGACGCTTCGCACGGTTG)	−	+	10	10 mL	O/N
AIV	d(TTB$_T$GGACGCTTCGCACG)	−	+	10	1 mL	O/N
AV	d(TTB$_T$GGACGCTTCGCACG)	−	+	5	1 mL	O/N
AVI	d(TTB$_T$GGACGCTTCGCAC)	−	+	5	1 mL	O/N
AVII	d(TTB$_T$GGACGCTTCGCAC)	−	+	5	1 mL	O/N
BIV	d(TTB$_T$GGACGCTTCGCACG)	+	−	10	1 mL	O/N
BV	d(TTB$_T$GGACGCTTCGCACG)	+	−	5	1 mL	O/N
BVI	d(TTB$_T$GGACGCTTCGCAC)	+	−	5	1 mL	4

Selection PT:
 (D11P12) 5′-d(**CACG**AGCCACG)CUGCCAACCGUG-3′
 (T21)   3′-GACGGUUGGCACGCUUCGCAG − 5′

B$_T$ indicates a biotinylated thymidine residue.
[a]10 nM corresponds to a weighted average, by volume, of 10–30 DNA genomes per compartment.

selected via a thiol-N-acryloyl-aminophenyl-mercuric acid (APM)-dependent gel shift. In addition, the P2 RNA was omitted from this branch of selection as it was found to inhibit extension of the PT(D4P12:T21) by the WT polymerase (data not shown).

The effect of adding an APM-purification step was apparent after only two rounds of selection, and by Round 6 of the B branch selection, extension activity had increased markedly (Fig. 6A). Screening 60 isolates from this round identified a total of 10 active and 50 inactive isolates with one clone, called B6.61, having polymerization activity significantly superior to that of the WT polymerase. B6.61 extended the PT(D4P12:T21) by 6 nt or more three- to fourfold faster than either the WT ribozyme or the best isolates from the branch A selection (Fig. 6B). The B6.61 sequence differed from the WT sequence and had the following sequence at its 5′ terminus: 5′-GG**AAUAC**AAAAG... (most simply interpreted as a 4-nt insertion immediately after G_2 and a A_7C mutation) (Fig. 6C). This insertion, presumably derived from the $(NN)_{0-5}$ region designed into the starting pool, was identical to that found in the A7.15 variant and was similar to that of the 6-nt insertion found in the A7.29 isolate (5′-GG**CAUACC**AAAAAG...). In addition, all three variants acquired an $A_{170}G$ mutation in the accessory domain, suggesting a convergent functional requirement for all three sequences. Indeed, the overall similarity in sequence between B6.61 and the WT polymerase suggests that the WT/B6.61 lineage is highly optimized.

B6.61 incorporates > 20 nt of sequence

The faster kinetics exhibited by B6.61 were also apparent when the shorter primer P7 was annealed to the T21 template (Table 1). Although both the WT and the B6.61 polymerases extended the PT(P7:T21) by 14 nt, B6.61

polymerized extension of > 9 nt about threefold faster than the WT (Fig. 7). More interesting was the finding that B6.61 extended PT(P7:T31) beyond the 14-nt limit, reaching a total extension of > 20 nt (Fig. 7, T31 is based on T21 and contains 10 additional residues; Table 1), while the WT ribozyme failed to extend more than ∼ 11 nt. The rate of RNA polymerization beyond 10 nt using
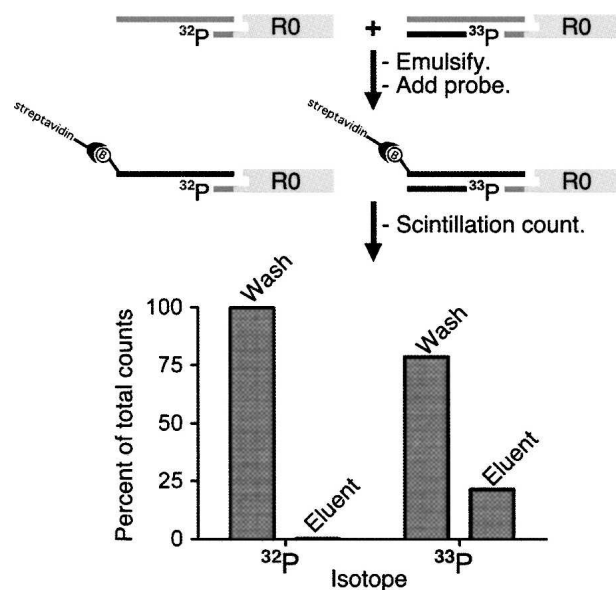
FIGURE 4. The hybridization approach is efficient in isolating extended DNA constructs. A ^{33}P-end labeled D11P12 primer was extended by dNTPs in the presence of T21 using superscript reverse transcriptase (Invitrogen), gel purified, and ligated to the DNA pool. This material was then mixed with an equal amount of DNA that had been ligated to a ^{32}P-end-labeled unextended D11P12 primer. A mock round of selection recovered $21.4 \pm 0.7\%$ of the ^{33}P-labeled material (mock extended), while only $0.21 \pm 0.03\%$ of the ^{32}P-labeled construct (unextended) was retained, as determined by spectrum gated scintillation counting.
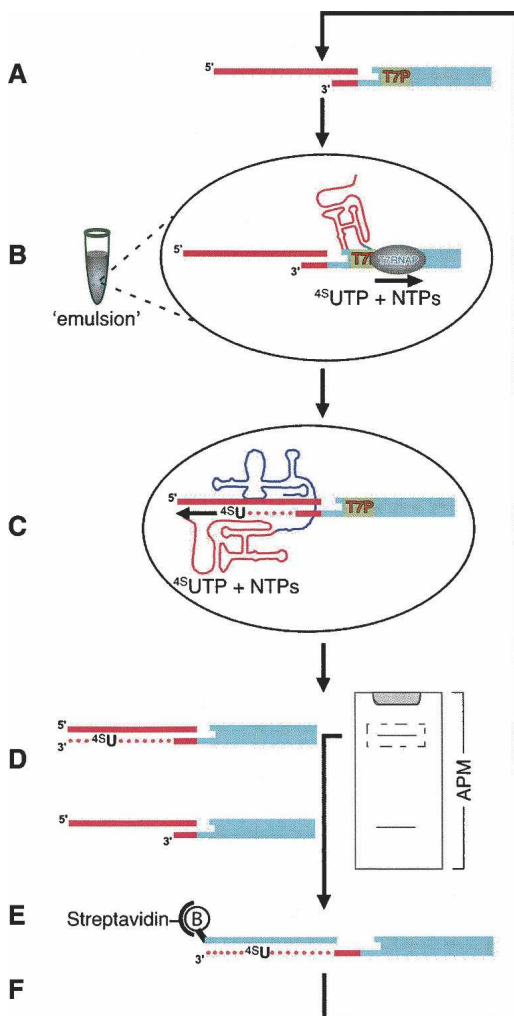
FIGURE 5. Selection scheme used in the branch B selection. (*A*) DNA constructs, which were ligated to an RNA primer, are annealed to the selection template. (*B*) This PT-pool is then compartmentalized within water-in-oil droplets together with NTPs, ^{4S}UTP, and T7 RNAP. Incubation results in T7 RNAP-mediated transcription of the RNA phenotype. (*C*) Active polymerase ribozymes recognize the PT complex that was attached to their parent DNA genotype and extend it in a template-directed manner while also incorporating ^{4S}UMP. (*D*) The emulsion is broken open and the recovered DNA genomes separated on an APM gel to enrich in those modified with ^{4S}UMP. (*E*) The recovered material is then hybridized to a biotinylated probe having a sequence complementary to the extended primer, and applied to streptavidin magnetic beads. (*F*) The purified pool is eluted from the beads, amplified, digested with BtgI, and ligated to the selection DNA/RNA primer ready for the next round of selection.

PT(P7:T31) was at least 75-fold faster than the WT, accounting for a total of 7.5% of the primer after 24 h of incubation while the WT had only extended ∼0.1%. The relative rate enhancement for longer extension products was even higher but could not be quantified accurately.

The importance of the mutations in B6.61 for this increased polymerization activity was not immediately apparent. The effect of magnesium and pH on the ability of B6.61 and the WT to extend PT(D4P12:T21) was examined. While both ribozymes were strongly affected by the concentrations of magnesium and hydroxide, the ratio of the extension rates between the two ribozymes was quite similar, and the ratio was independent of ion concentration—with the B6.61 ribozyme always exhibiting superior extension kinetics. Magnesium was titrated from 12.5 to 200 mM at a pH of 8.5 while pH dependence was measured from 5.7 to 9.5 at 50 and 200 mM magnesium. We had expected that if the phosphodiester bond forming catalytic core of the polymerase—the class I ligase—had been significantly affected by the mutations found in B6.61, the relative rates of the two ribozymes then might have differed as a function of pH or magnesium concentration. While it is certainly still possible that the mutations found in B6.61 directly affect chemistry, the superior polymerization ability of B6.61 most likely resides with some other aspect of the polymerization process. This is consistent with the location of the B6.61 5′ mutations; based on a tertiary structure model of the class I ligase, they are in a relatively unstructured region (Bergman et al. 2004).

Titrating the PT(P7:T31) concentration indicated that neither the WT nor B6.61 ribozyme bound the PT with high affinity. Only when the PT complex concentration was as high as 15 μM did the relative extension efficiency of B6.61 improve relative to the WT. As B6.61 demonstrated superior polymerization even when a large excess of polymerase was used relative to the PT substrate, this modest improvement in PT binding affinity could not easily explain the superior extension properties of the B6.61 polymerase.

B6.61 is more accurate

A clue to B6.61's superiority was revealed by observing that B6.61 had improved fidelity. Whereas B6.61 was always faster, changing primers while using the T21 and T31 templates affected the overall polymerization in a complex way. Unexpectedly, the PT(DP12:T31) complex resulted in the very fast incorporation of nucleotides that were not simply encoded by the template sequence (Fig. 8A, marked by asterisks). Since T31 shares nucleotides 10–31 (5′ to 3′) with T21, the same extension pattern up to the 9-nt addition point on either template had been expected. Nevertheless, B6.61 synthesized template-encoded nucleotides in the 4- to 8-nt range more rapidly than the WT and appeared to have produced correct extensions further up the gel. This strongly suggested that B6.61 has superior fidelity to the WT. This hypothesis was confirmed by a second extension assay. Polymerization of a completely new PT(P7:T26) by the WT polymerase produced a misincorporation after the first extension product (Fig. 8B, asterisk in the -PT WT lanes). This error was found to be most likely a uridine rather than the cytidine encoded by the template (Fig. 9). B6.61 suppressed this error (Fig. 8B),
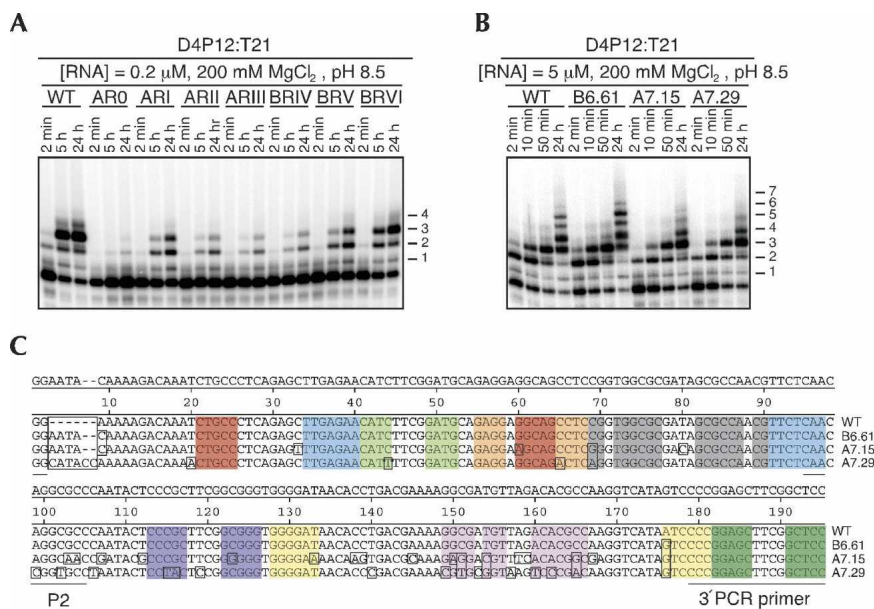
FIGURE 6. Progress and outcome of the selection. (*A*) Polymerization activity of gel-purified selection rounds under the conditions indicated. Aliquots were taken at the indicated time intervals and extension products separated on 20% PAGE. Length of extension products are shown on the *right*. After six rounds of the branch B selection, extension activity was enriched by a factor of 60. (*B*) Extension of PT(D4P12:T21) complex by variants from branches A and B of the selection. B6.61 extended >6 nt of the PT complex threefold faster than the WT, while A7.15 and A7.29 were slightly slower. (*C*) Sequence alignment of the selected isolates relative to the WT sequence. Elements highlighted in the same color are proposed to form a secondary structure with each other based on the previous secondary structure determination of the class I ligase and WT polymerase. Residues that are different from the WT sequence are boxed. The initiating guanosine residues are underlined, while the underlined CAA stretch indicates the fixed sequence that was required to ligate the two modules together to make the starting pool.

element interacts with the polymerase, and it has been assumed that the accessory domain sequence stabilizes the P2–polymerase interaction. Our data suggest that the P2 element regulates the fidelity of the WT enzyme but at the expense of overall extension rate (Fig. 8B). Indeed, when P2 RNA was added to the reaction, extension by either polymerase decreased dramatically (data not shown). The mutations found in B6.61 appear to bypass the detrimental aspects of the P2 RNA element and provide a new way for the polymerase to extend a PT quickly and with improved fidelity.

B6.61 exhibited better polymerization kinetics than the WT for all PT complexes tested, and with certain PT complexes, the overall rate of extension beyond a helical turn was at least 75 times faster. A major component of this superiority may be attributed to the ability of B6.61 to correctly incorporate nucleotides encoded by the template sequence (Fig. 8). The improved fidelity demonstrated by B6.61 is important as it is achieved by a mechanism that also improves elongation efficiency. Both features are critical attributes of a true polymerase.

A polymerase must hold both the PT complex and the incoming nucleotide in such a way that only Watson–Crick-paired nucleotides are incorporated onto the elongating strand. This discrimination can only be improved by stabilizing correct A:U and G:C pairs or by destabilizing inappropriate pairings. As the Watson–Crick face of a potential nucleotide:template base pair is by necessity hidden from the ribozyme, it is possible that B6.61 has evolved a simple molecular caliper able to stabilize correct Watson–Crick pairs at the site of nucleotide incorporation. As this interaction would be transient—occurring only when PT and nucleotide are bound—it could increase both polymerization efficiency and fidelity without dramatically changing PT affinity or the ribozyme's pH or metal-ion dependence, consistent with the observed data. Molecular calipers of precisely this sort are used by protein polymerases. These enzymes, by making hydrogen bonds to the purine N3 and pyrimidine O2, measure nearly invariant properties of Watson–Crick pairs that would be violated by, for example, a G:U wobble pair (Kunkel and Bebenek 2000).

indicating that the mutations found in B6.61 provide a mechanism to improve the polymerase's ability to distinguish a Watson–Crick pair from that of a wobble.

Collectively the data presented in Figures 7 and 8 suggest that the mutations in B6.61 confer both greater fidelity and improved elongation rate to the polymerase ribozyme. Intriguingly, the beneficial effect of the B6.61 mutation could be suppressed by the addition of the P2 RNA. In the presence of P2 RNA, B6.61 behaves very much like the WT ribozyme; with both enzymes exhibiting a nearly identical pattern of misincorporation and extension (Fig. 8B). This suggests a model whereby the new sequence found in B6.61 interacts with a region of the polymerase that previously interacted with the P2 RNA element.

The P2 element was originally added to the WT polymerase selection to allow the formation of a 6 base-pair helix called P2 thought to be critical for the class I ligase (Bergman et al. 2004). At the end of the selection, the WT polymerase was found to contain a U to C mutation in the P2 helical arm that would be expected to destabilize the P2 helix (Fig. 1A). Nevertheless, this mutation had a dramatic effect on polymerization, and when changed back to a U, a significant loss in activity was observed (Johnston et al. 2001). Therefore, it has been mysterious how the P2

All modern protein DNA polymerases, in addition to being complex multisubunit machineries, depend on other enhancing factors to reach their full processive potential.
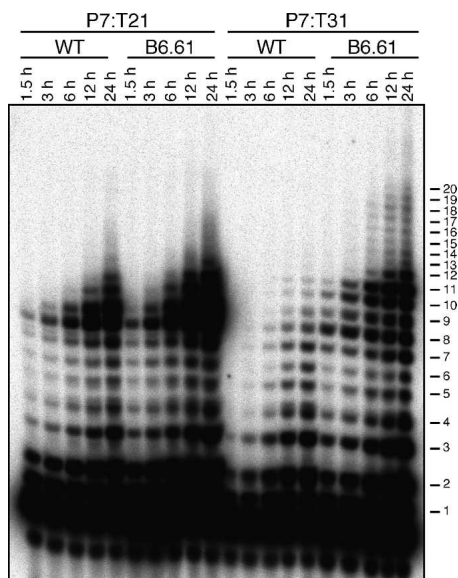
FIGURE 7. B6.61 polymerizes a PT complex by at least 20 nt. The PT complexes were incubated in the presence of WT or B6.61 ribozymes using optimum conditions. Reactions were stopped at the specified time intervals and resolved using 20% PAGE. Although both polymerases extended PT(P7:T21) by 14 nt (the end of the template), B6.61 was threefold faster in synthesizing > 9 nt extension products (*left* panel). B6.61 was able to extend PT(P7:T31) by at least 20 nt in contrast to the 11 nt of extension detectable by the WT ribozyme (*right* panel).

For instance, the core DNA polymerase III of *Escherichia coli* polymerizes DNA with a rate of 20 nt/sec and a processivity only of tens of nucleotides. By interacting with the β-clamp protein, its rate of polymerization is improved by a factor of ∼40, with a processivity exceeding 50,000 nt (Benkovic et al. 2001). B6.61 has a respectable polymerization rate with incorporation rates reaching 0.6 nt/min (Fig. 6), still far less than the 100 min^{-1} k_{cat} observed for phosphodiester bond formation by the class I ligase core (Bergman et al. 2000). As a result, it is reasonable to assume that selection for a B6.61 RNA cofactor could dramatically improve the polymerization rate of the B6.61 complex. Furthermore, as the polymerase currently exhibits a low affinity for the PT complex, it is conceivable that the association of B6.61 with this RNA cofactor may dramatically improve the processivity of this ribozyme polymerase. The IVC technique described here can be easily modified to isolate such a helper RNA that would be extremely difficult or impossible to achieve using conventional techniques. Round 6 from the branch B selection contained ∼ 80% inactive genomes, which implies that the catalytic power of active ribozyme polymerases such as B6.61 is sufficient to carry forward a mixed population of DNA genomes through the selection process. If these inactive genomes could be replaced with a high diversity pool and placed under selective pressure in the presence of B6.61 so as to improve its polymerization capability, it might be possi-

ble to isolate a truly processive multisubunit polymerase ribozyme.

MATERIALS AND METHODS

Oligonucleotides

Deoxy-oligonucleotides were synthesized on an ABI 392 DNA/RNA synthesizer using Expedite chemistry. Ribo-oligonucleotides were purchased from Dharmacon and deprotected according to the company's protocol. The selection primer having the sequence 5′-d(CACGAGCCACG)CUGCCACCGUG(D11P12) (Table 1) was made by ligating a DNA and an RNA oligonucleotides together using T4 DNA ligase and a splint having the sequence 5′-d(CAGCGTGGCTCGTG).

Pool construction

The ligase core of the pool was synthesized from two oligonucleotides having the following sequence: 5′-CAACTAG**CTCTTC**A GTTgagaacgttggcgctatcgcgccaccggaggctgcctctgcatccgaagatgttctcaag ctctgagggcagatttgtctttt (NN)$_{0–5}$CC*TATAGTGAGTCGTAT*. The lowercase residues were mutated at a frequency of 0 (L_W) or 3% (L_M). The sequence corresponds to the reverse complement of the ligase core and contains a T7 promoter sequence in italics and an EarI site in bold. A split and pool strategy was used to synthesize the (NN)$_{0–5}$ region (Fig. 1A). The two sequences were PCR amplified with the top strand PCR primer 5′-CTCGCACCT G**CCGTGG**GAGACGTTCATCCTTTA*ATACGACTCACTATAGG* (introducing the BtgI cleavage sequence in bold, which leaves the underlined overhang after digestion) and the bottom strand PCR primer 5′-CAACTAG**CTCTTC**A<u>GTT</u> (EarI recognition site in bold and leaving the final double-stranded DNA with the underlined sequence overhanging after digestion). Two oligonucleotides corresponding to the accessory domain sequence 5′-GACTGCA **CTCTTC**C<u>AAC</u>aggcgcccaatactcccgcttcggcgggtggggataacacctgacgaa aaggcgatgttagacacgccaaggtcataatCCCGGAGCTTCGGCTCC (EarI recognition site in bold, with corresponding overhang underlined) were also synthesized; again one strand was not mutagenized (A_W), while the second had the lowercase residues mutated at a frequency of 10% (A_M). These oligonucleotides were made double stranded by PCR amplification with the top strand primer 5′-GACTGCA**CTCTTC**C<u>AAC</u> and the bottom strand PCR primer 5′-GGAGCCGAAGCACCGGG and were digested by EarI (bold residues) to produce double-stranded DNA with a 3-nt overhang (underlined residues in primer). The EarI-digested ligase core and accessory domain pool fragments were mixed pairwise and ligated together to make three pool subpopulations as shown in Figure 1B. The final pool was digested with BtgI and ligated to the D11P12 primer (Table 1) using T4 DNA ligase producing the P-pool. This pool was then annealed to a twofold excess of the selection template T21 (Table 1) to produce a DNA pool attached to the PT (called the PT-pool) (Fig. 1C).

Emulsion procedure

The emulsion protocol, which was originally described by Tawfik and Griffiths (1998) for protein selections, was modified to allow for the generation of liters of emulsion in 50 mL batches. A 2.5 mL

A

B
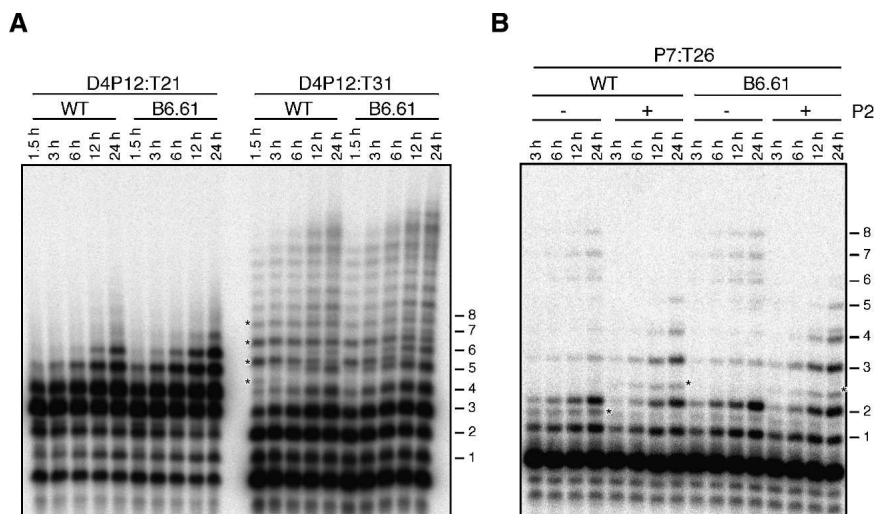
FIGURE 8. The fidelity of B6.61 is superior. (*A*) The extension pattern of PT(D4P12:T31) by both polymerases was compared with the pattern observed for PT(D4P12:T21). Reactions were conducted using optimum conditions, and aliquots were taken at the indicated time points. Expected extension products are numbered, while incorrect products are marked by asterisks. Note that B6.61 produces approximately threefold more of the expected products relative to the WT. Products beyond the 8-nt addition were not labeled as there was no simple standard to determine whether they were correct additions. (*B*) Effect of P2 on ribozyme fidelity. Ribozyme-catalyzed extension of PT(P7:T26) was conducted in the absence or presence of the P2 oligonucleotide using optimum conditions. In the absence of the P2 element, the WT polymerase adds an incorrect product after the first addition (marked by an asterisk), while B6.61 produces three- to fourfold less of the same anomalous product. In the presence of P2, the accumulation of this product by the WT ribozyme was inhibited, but the addition of an incorrect nucleotide after the second addition was enhanced by both polymerases.

T7 RNAP transcription reaction containing a 10 nM PT-pool in the selection buffer (40 mM Tris-HCl at pH 7.9, 2.5 mM spermidine, 50 mM MgCl$_2$, 0.01% Triton X-100, 10 mM DTT, 8 mM GTP, 2 mM ATP, 2 mM CTP, and 2 mM UTP, 7.5 μM P2 RNA if added) (see Table 2) was prepared in a Falcon tube (Becton Dickinson) and incubated on ice for at least 10 min before the addition of T7 RNAP at 10 U/μL. This was followed immediately by the addition of 47.5 mL of the oil surfactant mixture composed of 0.5% Tween 80 (Sigma) and 4.5% Span 80 (Sigma) in heavy mineral oil (Paraffin oil white, Anachemia) at a temperature of 4°C. The mixture was emulsified by adding DEPC-treated 6-mm glass beads so as to completely fill the Falcon tube, taping the tube to a mini vortexer (VWR), and vortexing at full speed for 5 min. The order of addition of the glass beads was not essential, and in some instances where large numbers of tubes were utilized (i.e., Fig. 2A), the prechilled beads were added first to facilitate batch processing. For later rounds, where the volume of the emulsion was reduced to 1 mL, the emulsion was prepared in an Eppendorf tube using a dispersing tool (IKA WORKS). The resulting emulsion from both procedures was incubated for 3 h at 37°C to drive the T7 RNAP-dependent transcription followed by heat inactivation of T7 RNAP for 15 min at 65°C. The resulting emulsion was then incubated at room temperature either overnight or as indicated. The aqueous phase was then recovered by centrifuging the emulsion at 12,000*g* for 20 min, which resulted in a white pellet and clear oil supernatant. A twofold excess of 30 mM EDTA was added to the pellet, which was then resuspended by vigorous vortexing. Excess oil was removed by extraction with

an equivalent volume of diethyl ether. The recovered aqueous phase was ethanol precipitated, phenol/chloroform extracted, and subjected to a second ethanol precipitation step before resuspension in water.

Selection

A total of 1.5 nmol (9 × 10^{14} different sequences) of pool were used in the first round of selection and required 3 L of emulsion. The emulsion was incubated for ∼18 h at room temperature after heat inactivating the T7 RNAP. The recovered nucleic acid was denatured by adding KOH to a final concentration of 50 mM. A 10-fold excess of a biotinylated DNA capture probe (Table 2), able to hybridize to the extended PT complexes, was then added and the mixture promptly neutralized with Tris-HCl. The pool was applied to streptavidin magnetic beads (NEB) and washed three times with SSC 0.5× buffer. Hybridized nucleic acid was eluted by applying 50 mM KOH to the beads. Recovered material was neutralized with Tris-HCl and ethanol precipitated. This DNA was then PCR amplified, BtgI digested, and ligated to the selection DNA/RNA primer before starting the subsequent round of selection. In branch B of the selection, 0.5 mM ^{4S}UTP was added to the selection buffer and P-pool elements that were successfully extended by ^{4S}UMP were isolated based on their slower mobility in APM polyacrylamide gels (Igloi 1988).
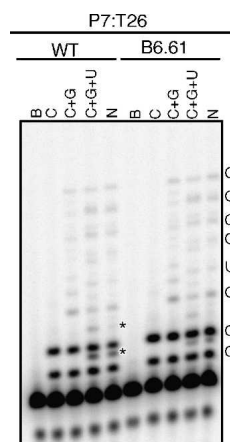
FIGURE 9. The WT polymerase utilizes wobble base pairing during extension more often than B6.61. Polymerization of PT(P7:T26) by both ribozymes was conducted in the absence of NTPs (B); CTP (C); CTP and GTP (C+G); CTP, GTP, and UTP (C+G+U); and all four NTPs (N). Only in the presence of UTP is the misincorporation observed, where the WT ribozyme incorporates UMP across a guanosine residue approximately threefold more than the B6.61 ribozyme.

Polymerization assay

5′-Radiolabeled primer (0.1 μM final concentration) was annealed to a fivefold excess of template and added to either the selection buffer or optimal buffer (200 mM MgCl$_2$, 100 mM Tris-HCl at pH 8.5, and 4 mM of each NTP). Gel-purified ribozymes (5 μM final concentration or otherwise indicated) were added to the reaction and incubated at 22°C. Reactions were stopped by adding a ninefold excess of a stop mix (80% formamide, 40 mM EDTA, 0.025% Xylene Cyanol, and 0.025% bromophenol blue) in the presence of ∼ 20-fold excess of a competing oligonucleotide able to hybridize to the template strand. Products were resolved using 20% sequencing PAGE. In reactions where the P2 oligonucleotide was included, it was added at 7.5 μM. The relative extension rate of a polymerase was calculated by first quantifying extension past a particular nucleotide and dividing this value by the equivalent number found for the Round-18 polymerase.

REFERENCES

Agresti, J.J., Kelly, B.T., Jaschke, A., and Griffiths, A.D. 2005. Selection of ribozymes that catalyse multiple-turnover Diels-Alder cyclo-additions by using in vitro compartmentalization. *Proc. Natl. Acad. Sci.* **102:** 16170–16175.

Ban, N., Nissen, P., Hansen, J., Moore, P.B., and Steitz, T.A. 2000. The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* **289:** 905–920.

Bartel, D.P. and Szostak, J.W. 1993. Isolation of new ribozymes from a large pool of random sequences. *Science* **261:** 1411–1418.

Bartel, D.P. and Unrau, P.J. 1999. Constructing an RNA world. *Trends Cell Biol.* **9:** M9–M13.

Bartel, D.P., Doudna, J.A., Usman, N., and Szostak, J.W. 1991. Template-directed primer extension catalyzed by the *Tetrahymena* ribozyme. *Mol. Cell. Biol.* **11:** 3390–3394.

Been, M.D. and Cech, T.R. 1988. RNA as an RNA polymerase: Net elongation of an RNA primer catalyzed by the *Tetrahymena* ribozyme. *Science* **239:** 1412–1416.

Benkovic, S.J., Valentine, A.M., and Salinas, F. 2001. Replisome-mediated DNA replication. *Annu. Rev. Biochem.* **70:** 181–208.

Bergman, N.H., Johnston, W.K., and Bartel, D.P. 2000. Kinetic framework for ligation by an efficient RNA ligase ribozyme. *Biochemistry* **39:** 3115–3123.

Bergman, N.H., Lau, N.C., Lehnert, V., Westhof, E., and Bartel, D.P. 2004. The three-dimensional architecture of the class I ligase ribozyme. *RNA* **10:** 176–184.

Buskirk, A.R., Kehayova, P.D., Landrigan, A., and Liu, D.R. 2003. In vivo evolution of an RNA-based transcriptional activator. *Chem. Biol.* **10:** 533–540.

Chowrira, B.M., Berzal-Herranz, A., and Burke, J.M. 1993. Novel RNA polymerization reaction catalyzed by a group I ribozyme. *EMBO J.* **12:** 3599–3605.

Doudna, J.A. and Szostak, J.W. 1989. RNA-catalysed synthesis of complementary-strand RNA. *Nature* **339:** 519–522.

Doudna, J.A., Usman, N., and Szostak, J.W. 1993. Ribozyme-catalyzed primer extension by trinucleotides: A model for the RNA-catalyzed replication of RNA. *Biochemistry* **32:** 2111–2115.

Ekland, E.H. and Bartel, D.P. 1995. The secondary structure and sequence optimization of an RNA ligase ribozyme. *Nucleic Acids Res.* **23:** 3231–3238.

Ekland, E.H. and Bartel, D.P. 1996. RNA-catalysed RNA polymerization using nucleoside triphosphates. *Nature* **382:** 373–376.

Ekland, E.H., Szostak, J.W., and Bartel, D.P. 1995. Structurally complex and highly active RNA ligases derived from random RNA sequences. *Science* **269:** 364–370.

Ferber, M.J. and Maher 3rd, L.J. 1998. Combinatorial selection of a small RNA that induces amplification of IncFII plasmids in *Escherichia coli*. *J. Mol. Biol.* **279:** 565–576.

Fiammengo, R. and Jaschke, A. 2005. Nucleic acid enzymes. *Curr. Opin. Biotechnol.* **16:** 614–621.

Gilbert, W. 1986. The RNA world. *Nature* **319:** 618.

Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., and Altman, S. 1983. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **35:** 849–857.

Igloi, G.L. 1988. Interaction of tRNAs and of phosphorothioate-substituted nucleic acids with an organomercurial. Probing the chemical environment of thiolated residues by affinity electrophoresis. *Biochemistry* **27:** 3842–3849.

Johnston, W.K., Unrau, P.J., Lawrence, M.S., Glasner, M.E., and Bartel, D.P. 2001. RNA-catalyzed RNA polymerization: Accurate and general RNA-templated primer extension. *Science* **292:** 1319–1325.

Joyce, G.F. 2002. The antiquity of RNA-based evolution. *Nature* **418:** 214–221.

Kunkel, T.A. and Bebenek, K. 2000. DNA replication fidelity. *Annu. Rev. Biochem.* **69:** 497–529.

Lawrence, M.S. and Bartel, D.P. 2003. Processivity of ribozyme-catalyzed RNA polymerization. *Biochemistry* **42:** 8748–8755.

Lawrence, M.S. and Bartel, D.P. 2005. New ligase-derived RNA polymerase ribozymes. *RNA* **11:** 1173–1180.

Levy, M., Griswold, K.E., and Ellington, A.D. 2005. Direct selection of *trans*-acting ligase ribozymes by in vitro compartmentalization. *RNA* **11:** 1555–1562.

McGinness, K.E. and Joyce, G.F. 2002. RNA-catalyzed RNA ligation on an external RNA template. *Chem. Biol.* **9:** 297–307.

McGinness, K.E. and Joyce, G.F. 2003. In search of an RNA replicase ribozyme. *Chem. Biol.* **10:** 5–14.

Miller, O.J., Bernath, K., Agresti, J.J., Amitai, G., Kelly, B.T., Mastrobattista, E., Taly, V., Magdassi, S., Tawfik, D.S., and Griffiths, A.D. 2006. Directed evolution by in vitro compartmentalization. *Nat. Methods* **3:** 561–570.

Soukup, G.A. and Maher 3rd, J.J. 1998. Selection and characterization of RNAs that relieve transcriptional interference in *Escherichia coli*. *Nucleic Acids Res.* **26:** 2715–2722.

Szostak, J.W., Bartel, D.P., and Luisi, P.L. 2001. Synthesizing life. *Nature* **409:** 387–390.

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

Wadhwa, R., Yaguchi, T., Kaur, K., Suyama, E., Kawasaki, H., Taira, K., and Kaul, S.C. 2004. Use of a randomized hybrid ribozyme library for identification of genes involved in muscle differentiation. *J. Biol. Chem.* **279:** 51622–51629.

Wilson, D.S. and Szostak, J.W. 1999. In vitro selection of functional nucleic acids. *Annu. Rev. Biochem.* **68:** 611–647.

Winkler, W.C. and Breaker, R.R. 2005. Regulation of bacterial gene expression by riboswitches. *Annu. Rev. Microbiol.* **59:** 487–517.

Zaher, H.S. and Unrau, P.J. 2004. T7 RNA polymerase mediates fast promoter-independent extension of unstable nucleic acid complexes. *Biochemistry* **43:** 7873–7880.

Zaher, H.S. and Unrau, P.J. 2005. Nucleic acid library construction using synthetic DNA constructs. *Methods Mol. Biol.* **288:** 359–378.

Zimmerman, J.M. and Maher 3rd, L.J. 2002. In vivo selection of spectinomycin-binding RNAs. *Nucleic Acids Res.* **30:** 5425–5435.