

RESEARCH ARTICLE

RIBOZYMES

Processive RNA polymerization and promoter recognition in an RNA World

Razvan Cojocaru and Peter J. Unrau*

Early life is thought to have required the self-replication of RNA by RNA replicases. However, how such replicases evolved and subsequently enabled gene expression remains largely unexplored. We engineered and selected a holopolymerase ribozyme that uses a sigma factor-like specificity primer to first recognize an RNA promoter sequence and then, in a second step, rearrange to a processive elongation form. Using its own sequence, the polymerase can also program itself to polymerize from certain RNA promoters and not others. This selective promoter-based polymerization could allow an RNA replicase ribozyme to define “self” from “nonself,” an important development for the avoidance of replicative parasites. Moreover, the clamp-like mechanism of this polymerase could eventually enable strand invasion, a critical requirement for replication in the early evolution of life.

The RNA World Hypothesis posits that at the dawn of evolution, RNA played a key role in the establishment of life (1). Central to this hypothesis is the existence of an RNA replicase ribozyme capable of copying its own genome using a supply of prebiotically synthesized nucleotide monomers and RNA polymers (2, 3). Ever since the class I ligase ribozyme was isolated from a high-diversity RNA pool (4), there has been a sustained effort to produce highly processive polymerase ribozymes (5–12). Because the affinity of these polymerases for their RNA templates is weak, with Michaelis constant (K_M) values in the millimolar range (13), the most successful strategies to date have colocalized polymerase ribozymes with their substrates using concentration-enhancing micelles (7) or by anchoring either the RNA template (8, 9) or the RNA primer to be extended (10, 11) to the polymerase ribozyme. These strategies create a high local concentration of primer template with respect to the polymerase but fail to create a truly processive polymerase by virtue of the tethering strategies used to enhance polymerization.

Here, we report a natural linkage between the emergence of processivity and promoter selectivity in an RNA polymerase ribozyme. We hypothesized that an RNA polymerase ribozyme could be partially hybridized to a sigma factor-like specificity primer. This “open” clamp form (Fig. 1, A and D) would be able to search for a single-stranded RNA (ssRNA) promoter. Strand invasion would then allow template sequences containing a promoter to strip the specificity primer away from the primer-binding site (PBS) of the polymerase

(Fig. 1B), triggering a structural rearrangement to a processive, “closed” clamp form (Fig. 1, C and E). Such a mechanism is analogous to that used by extant bacterial DNA-dependent RNA polymerases (DdRPs), which have evolved to recognize promoters through a two-step process involving sigma factor-dependent promoter recognition and nucleoside triphosphate (NTP)-dependent structural rearrangement to a final processive elongation form (Fig. 1, A to C, bottom panels) (14–17). All extant DdRPs, including the bacteriophage polymerases (18–21), use a variation of this two-step process. Thus, we selected a ribozyme with a similar mechanism to explore the potential connection between promoter recognition and processivity.

Selection of a promoter-specific RNA polymerase ribozyme

To investigate this hypothesis, we started with the two-domain RNA polymerase ribozyme B6.61 (6), which consists of a catalytic ligase core and a secondary accessory domain that confers NTP extension ability through its AJ3/4 element (11). We engineered three changes into this parental ribozyme by appending a PBS to its 5' end, synthesizing a high-diversity pool containing 10^{13} sequence variants by inserting random sequence libraries at three distinct sites, and removing sequence from the B6.61 accessory domain known to be redundant (11) (fig. S1 and table S1).

Three selection schemes, a negative selection, a clamping selection, and a processivity selection (fig. S2), were alternated for 30 rounds to select for functional ribozymes (fig. S3 and table S2). The negative selection removed pool molecules that could hybridize to a linear, randomly generated selection template (T1) immobilized onto streptavidin magnetic beads (fig. S2A). The clamping selection (fig. S2B) first formed P1:Pool^{OPEN} molecules by incu-

bating them with the P1 specificity primer (table S3), which is fully complementary to the 22-nucleotide (nt) promoter found within the T1 template. To retain pool molecules that could make the transition to the Pool^{CLOSED} state, P1:Pool^{OPEN} molecules were added to circularized T1 (cT1) immobilized on streptavidin beads. Correctly clamped closed pool molecules retained on cT1 were then recovered by adding fresh specificity primer to again reform the open state and release correctly clamped pool molecules from the circular template. This process of transitioning from open to closed and back to open was performed either once or twice during clamping rounds of selection for increased selective pressure.

The processivity selection scheme incorporated polymerization activity into the clamping selection (fig. S2C). Here, activated P1:Pool^{OPEN} complexes were added to free circular template and incubated with adenosine triphosphate (ATP), guanosine triphosphate (GTP), and uridine triphosphate (UTP) (4 mM each), together with 1 mM biotin-11-cytidine-5'-triphosphate (CTP^B). Templates encoding for CTP incorporation at the first, third, or 10th extension position were then used to select for polymerization through the incorporation of CTP^B (tables S2 and S3). After this incubation, the pool-primer-template mixture was bound to streptavidin beads and washed. As before, captured pool ribozymes were recovered by adding specificity primer, reforming the open state, and allowing the recovery of clamping polymerase (CP) ribozymes with significant polymerization activity. After every round of selection, recovered pool RNA was reverse transcribed, polymerase chain reaction (PCR) amplified, and transcribed for the next round of in vitro selection (see the supplemental materials).

After 16 rounds of selection, the pool exhibited the anticipated clamping activity but only minor polymerization ability relative to the B6.61 progenitor. The selection pool was therefore mixed with three new subpools in which the existing clamping domain pool diversity was preserved but the ligase and accessory core of the polymerase were modified to contain the following: (i) nine high-frequency mutations found throughout the cloned pool, (ii) 11 point mutations and a deletion found previously by other groups (8, 10), and (iii) the union of subpools 1 and 2 mutations. To further increase diversity, the combined pools were subjected to mutagenic PCR. After 23 rounds of selection, the pool was found to add CTP^B to the 3' terminus of pool molecules. This was suppressed by extending the 3' terminus of the pool by a single A residue (fig. S4 and table S1C).

After 25 rounds of selection, a substantial decrease in pool diversity occurred, with the final five rounds of selection being dominated by five major ribozyme polymerase families

Department of Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, British Columbia, Canada V5A 1S6.
*Corresponding author. Email: punrau@sfu.ca

(Fig. 2A). This loss in diversity was directly correlated with the emergence of significant polymerization on the cT1 template (Fig. 2B). One ribozyme from family 1, CP, was characterized further. The CP ribozyme contained seven point mutations and deletions in the ligase core and

17 point mutations or insertions in the accessory domain relative to the progenitor B6.61 ribozyme. Of the 24 mutations found, 14 were activity-enhancing mutations that have been found previously (8, 10); of these 14 mutations, 11 were deliberately designed into the selection

pool at round 16 and three evolved independently (Fig. 1D and table S3).

Clamping domain characterization

Removal of the newly selected 3' clamping domain abolished polymerization activity (fig. S5

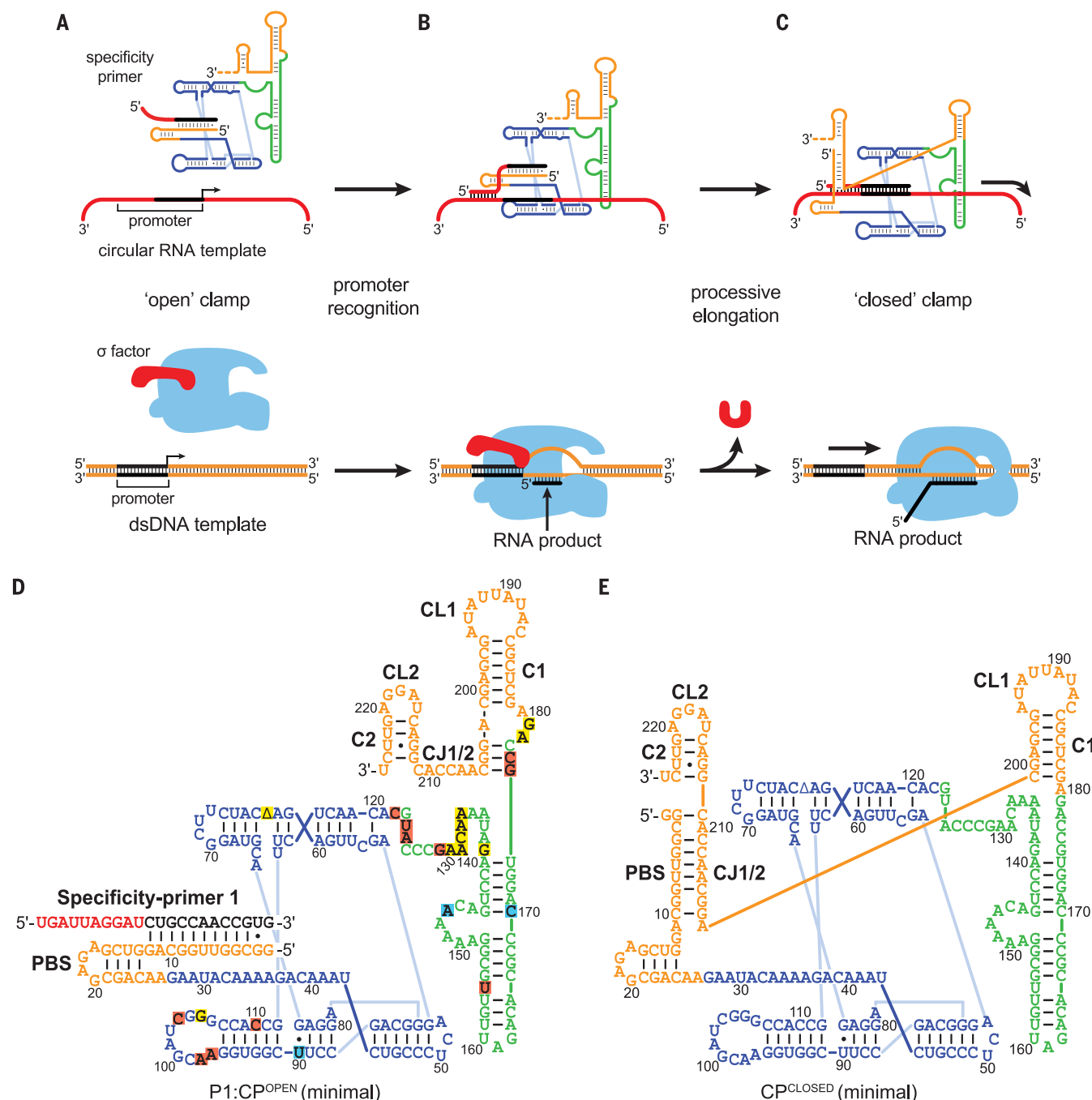


Fig. 1. Clamping RNA-dependent RNA polymerase (CP RdRP) ribozyme and DNA-dependent RNA polymerase (DdRP) transcriptional initiation processes. (A) RNA specificity primer-activated “open” form P1:CP^{OPEN} (top) and DdRP holo-enzyme (bottom). (B) The specificity primer localizes CP to an ssRNA promoter (top), whereas a sigma factor localizes the DdRP to a DNA promoter (bottom). (C) In both cases, a clamped “closed” state forms,

enhancing polymerization. (**D** and **E**) Secondary structure of the minimal P1:CP^{OPEN} form (D) and the CP^{CLOSED} form (E). Colored lines indicate the ligase core (blue), accessory domain (green), and minimal clamping domain (orange). Up mutations designed into the selection are shown in red boxes, rediscovered up mutations in teal boxes, and newly discovered mutations in yellow boxes.

and table S4, construct 5), whereas transplanting the clamping domain from the family 1 CP onto a lower-activity family 4 ribozyme (12% activity of CP) enhanced its activity to CP's level (fig. S6), implicating this new domain in

processive polymerization. Truncation analysis (table S4, constructs 1 to 18) and secondary structure prediction of the 3' clamping domain revealed a minimal 45-nt domain composed of two stem-loop structures, C1-CL1 and C2-CL2,

separated by a junction sequence, CJ1/2, shown in the predicted P1:CP^{OPEN} form (Fig. 1D). In the CP^{CLOSED} form, the C1 helix shortens by up to three base pairs (bp), allowing the CJ1/2 junction to form a 7-bp noncontinuous helix with the 5' PBS sequence as the specificity primer transfers to the template sequence (Fig. 1E). The closed form of the minimal clamping domain is highly structured, naturally precluding base-pair interactions with template sequence.

Removing the C1-CL1 stem-loop structure destroyed polymerization activity, whereas replacing either the stem sequence or the loop sequence with a GCAA tetraloop had a minimal effect on activity (table S4, constructs 19 to 25). Hybridization of a DNA oligonucleotide to this region also suppressed activity (table S4, DNA 1). Thus, the C1-CL1 stem-loop structure plays an important mechanistic role in forming the active form of the clamped polymerase. The C2-CL2 stem-loop structure was less critical because removing or mutating it had only an intermediate effect on activity.

The CJ1/2 region, which hybridizes to the PBS in the closed state (Fig. 1, D and E), showed the highest effect on activity when blocked or mutated (table S4, constructs 26 to 51 and DNAs 2 to 7). Introducing a G5:C208U wobble mutation in the PBS:CJ1/2 helix resulted in a 19% increase in extension compared with wild-type on cT1 (construct 36). Weakening this stem further affected activity, with G8:C205U and U6:A207G lowering extension to 43 and 50%, respectively, whereas combining the two mutations lowered activity to 11% (constructs 34 to 35 and 38). Changing the CJ1/2 sequence by seven nucleotides from ...A GGC AAC CAC G... to ...A CGG CCA AAA G... (underlined residues are predicted to hybridize to PBS; fig. S7A) was predicted to preserve the net hybridization between the PBS and the CJ1/2 in the clamped helix, and indeed had 53% activity, indicating that base-pair formation rather than sequence in this region is essential for forming a correctly closed clamp (construct 26). Conversely, strengthening hybridization in the PBS:CJ1/2 stem through A11:A202U or C3:A210G mutations lowered extension to 34 or 47%, respectively (constructs 29 and 30), while strengthening by 3 bp with A11:A202U, G9:G204C, and C3:A210G mutations dropped extension to only 2% (construct 31). Construct 31 prevented P1 hybridization and formation of the P1:CP^{OPEN} complex (fig. S7B, C). Because both stabilization and destabilization of the clamping helix can lower polymerization activity, a thermodynamic balance between the open, primer-bound form and the closed form of the polymerase is required for correct promoter-dependent polymerase function.

The clamping domain confers long-range extension and promoter selectivity

In addition to the cT1 template, we created a second template called cT2. This template,

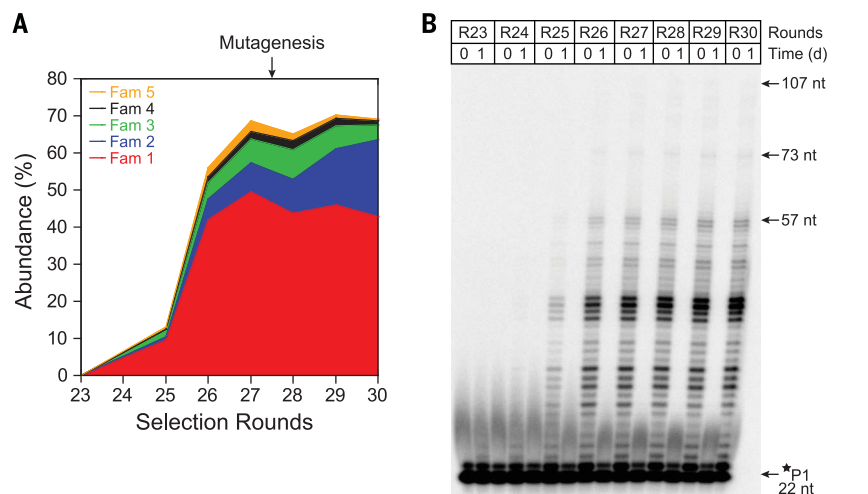


Fig. 2. Pool diversity and the emergence of polymerization. (A) Pool diversity from rounds 23 to 30. Families contain sequences with a pairwise distance $d \leq 2$. Arrow indicates 2% mutagenesis of the R28 DNA pool. (B) Extension activity of P1:Pool^{OPEN} on cT1 by selection rounds. Reaction conditions: P1 specificity primer (0.1 μ M) was mixed with ribozyme pools (0.12 μ M) in 100 mM MgCl₂, 100 mM KCl, and 100 mM Tris-HCl at pH 8.5 for 20 min at room temperature. Reactions were started by the addition of 4 mM concentrations of each NTP and 0.14 μ M cT1 template. Reactions were stopped by heating at 95°C for 5 min after adding equivolume 80% formamide, 200 mM EDTA, 0.025% xylene cyanol, 0.025% bromophenol blue, and 10-fold excess of an RNA oligonucleotide complementary to cT1 before loading on 10% PAGE.

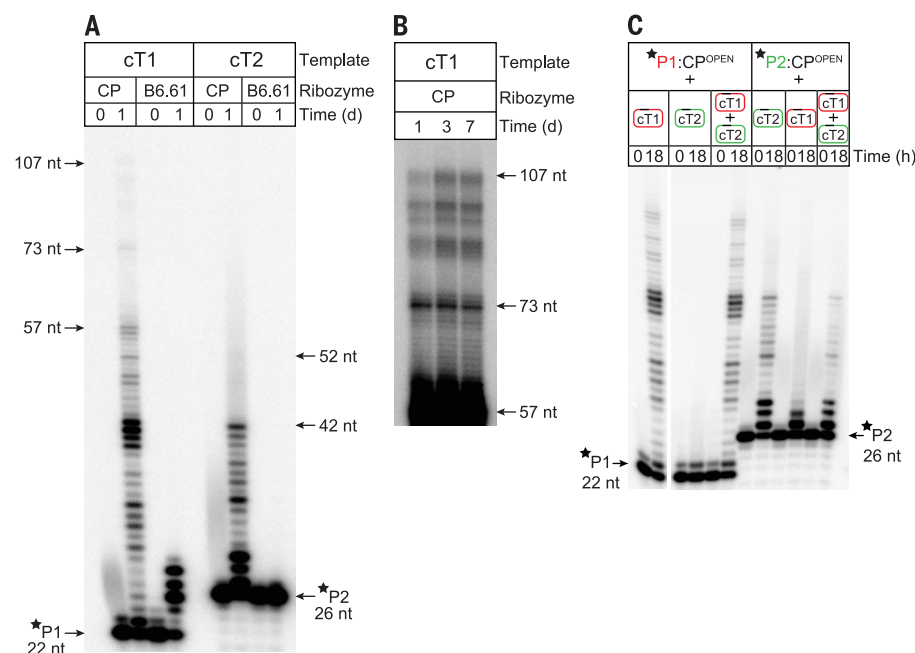


Fig. 3. Promoter-specific polymerization by CP on random sequence templates. (A) P1:CP^{OPEN} and P2:CP^{OPEN} extension relative to its progenitor, B6.61, on cT1 and cT2 promoter templates. (B) Contrast-enhanced long-range extension for a 7-day time course (full gel of both templates in fig. S6). (C) Promoter-mediated template selectivity by CP. Shown are P1:CP^{OPEN} or P2:CP^{OPEN} extensions with cT1, cT2, or both templates simultaneously.

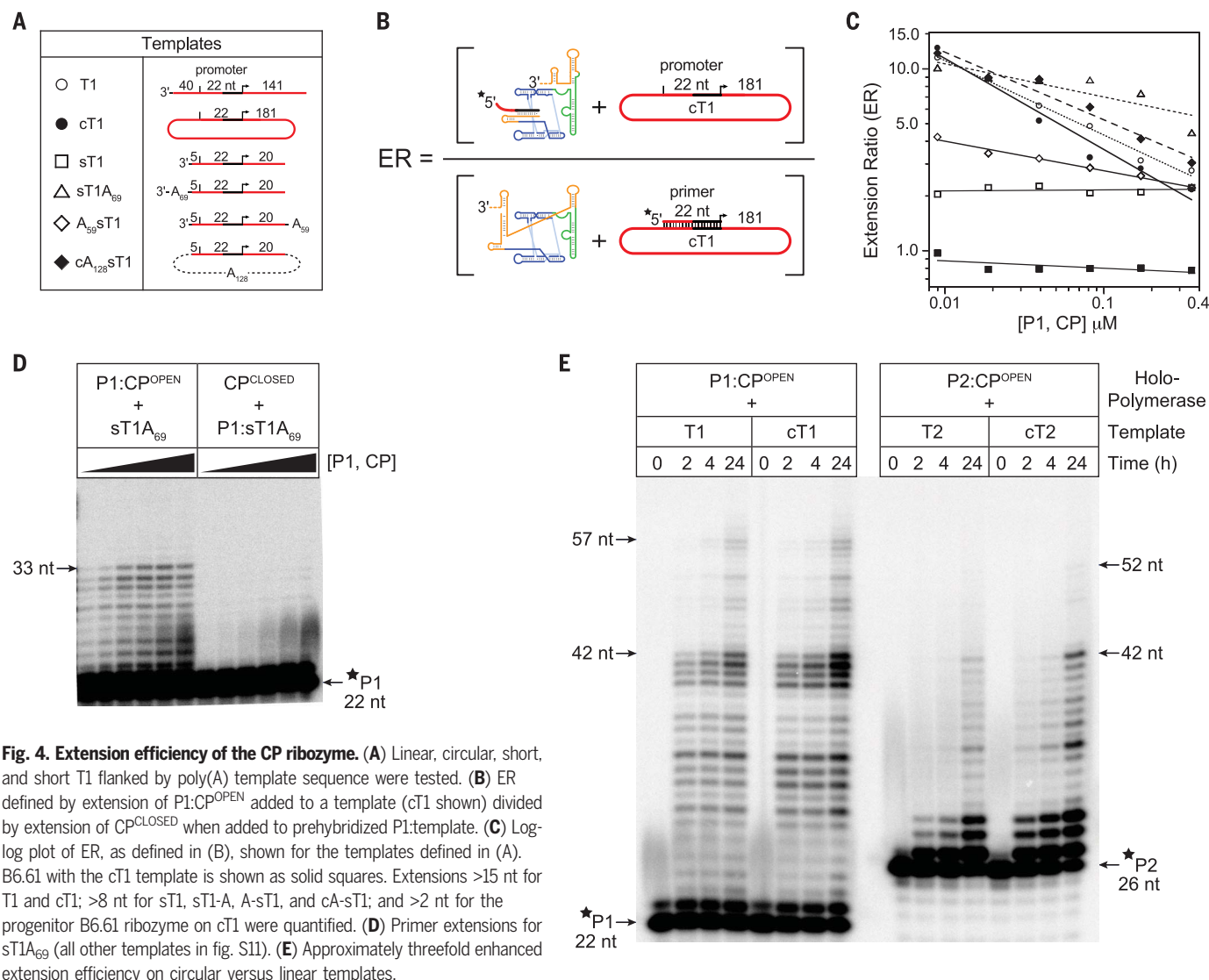


Fig. 4. Extension efficiency of the CP ribozyme. (A) Linear, circular, short, and short T1 flanked by poly(A) template sequence were tested. (B) ER defined by extension of P1:CP^{OPEN} added to a template (cT1 shown) divided by extension of CP^{CLOSED} when added to prehybridized P1:template. (C) Log-log plot of ER, as defined in (B), shown for the templates defined in (A). B6.61 with the cT1 template is shown as solid squares. Extensions >15 nt for T1 and cT1; >8 nt for sT1, sT1-A, A-sT1, and cA-sT1; and >2 nt for the progenitor B6.61 ribozyme on cT1 were quantified. (D) Primer extensions for sT1A₆₉ (all other templates in fig. S11). (E) Approximately threefold enhanced extension efficiency on circular versus linear templates.

also generated from random sequence, is distinct from cT1 and contains a new promoter region complementary to a 26-nt P2 specificity primer (table S3). P2 shares 10 nt in common with P1 at its 3' end, enabling its hybridization to the PBS just as P1 does (Fig. 1D). When incubated with cT1 for 24 hours, P1:CP^{OPEN} extends P1 by 85 nt, whereas P2:CP^{OPEN} extends P2 on cT2 by 26 nt. By contrast, the progenitor B6.61 ribozyme extended only 5 nt on cT1 and had no observable extension on cT2 (Fig. 3A). The long-range extension on cT1 saturated at ~85 nt after 3 days of incubation, whereas cT2 extended by up to ~40 nt (Fig. 3B and fig. S8). The holocomplexes formed from either P1 or P2 were template specific, with P1:CP^{OPEN} being specific for the cT1 promoter and P2:CP^{OPEN} being specific for the cT2 promoter, even when mixtures of cT1 and cT2 were presented simultaneously to either holopolymerase (Fig. 3C).

Relative to the B6.61 progenitor, the CP ribozyme also requires less Mg²⁺ to become fully

functional, with the emergence of polymerization occurring at 50 mM Mg²⁺ and saturating at 75 to 100 mM. By contrast, B6.61 polymerization showed no such saturation, with polymerization extension doubling from 75 to 100 mM and then tripling from 100 to 200 mM Mg²⁺ (fig. S9).

The clamping domain confers polymerization efficiency

A correctly clamped CP ribozyme should ideally stay localized to the primer-template complex that triggered the formation of its closed elongation form, whereas a less-processive ribozyme might dissociate. However, just as in extant biology, the CP ribozyme should only initiate elongation when its open holopolymerase form is presented to an ssRNA promoter template and should not polymerize efficiently when its closed form is presented to a primer-promoter-template complex. We found this to be the case, with P1:CP^{OPEN} extension on cT1 being

significantly better than CP^{CLOSED} extension on P1:cT1 (fig. S10).

To quantify clamp-driven extension efficiency, we defined and measured an extension ratio (ER) for a range of promoter templates (Fig. 4A). The numerator was defined as the percentage extension past a particular RNA product size when the open holopolymerase ribozyme was added to a promoter template (P1:CP^{OPEN} + template). The denominator was defined as the extension when the promoter template was first hybridized to the P1 specificity primer, and then added to the closed polymerase ribozyme (CP^{CLOSED} + P1:template; Fig. 4B).

We measured ER by maintaining a one-to-one stoichiometric ratio of the P1 primer and CP ribozyme, which was then titrated over two orders of magnitude of concentration on fixed 1-μM templates (Fig. 4C and fig. S11). As expected, the B6.61 progenitor ribozyme had a concentration-independent ER value of ~1,

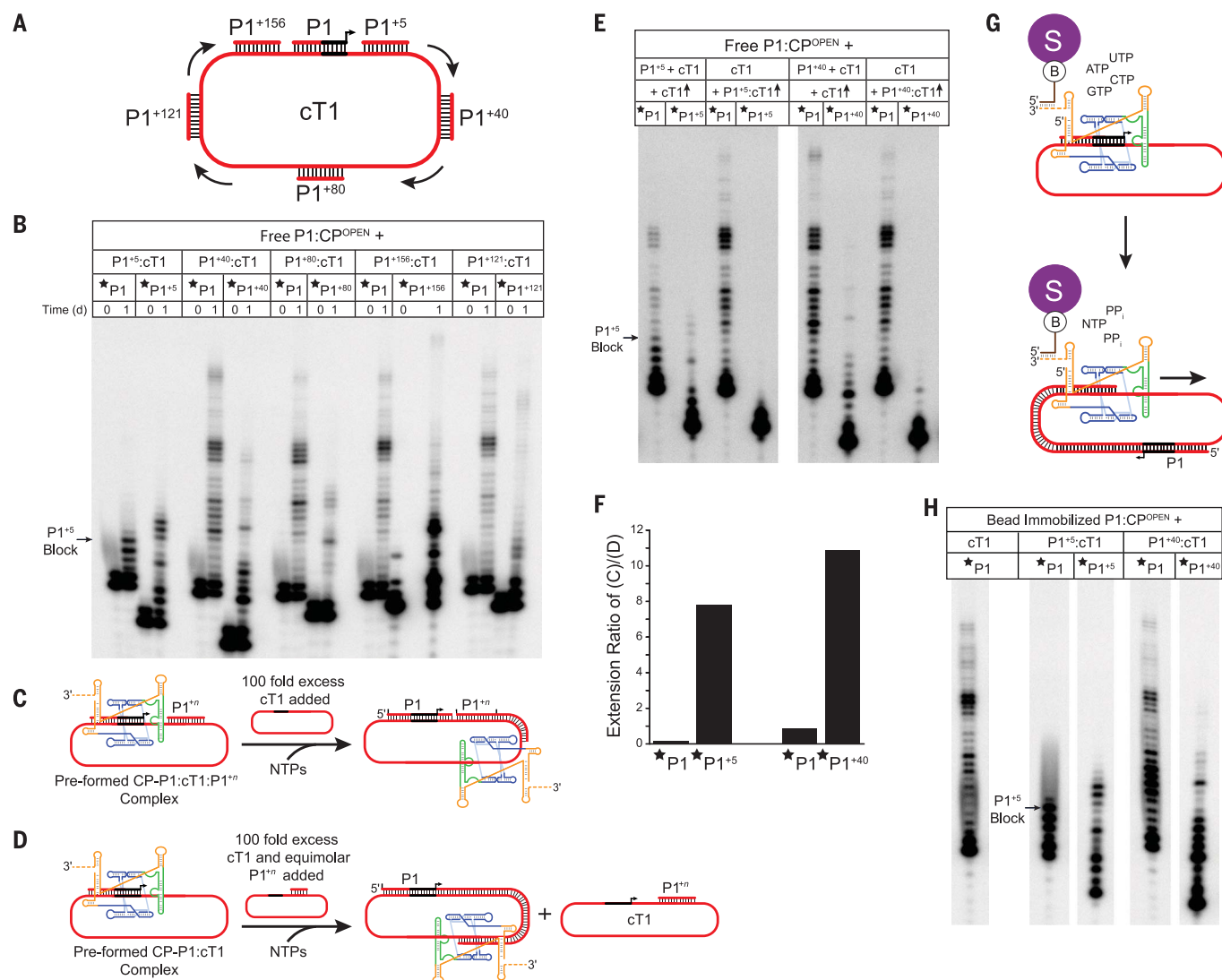


Fig. 5. CP processively extends multiple primers on the same promoter template.

(A) Distal primer locations relative to the P1:cT1 initiation site. Arrows indicate direction of polymerization. (B) Extension assay of P1:CP^{OPEN} and P1ⁿ ($n = 5, 40, 80, 121$, and 156) primers added to cT1. Either P1 or the P1ⁿ primers were radiolabeled as indicated (star). (C) To prepare primers in a correlated fashion, a CP^{CLOSED}-P1:cT1:P1ⁿ complex ($n = 5$ or 40) was preformed by mixing P1:CP^{OPEN} ($0.1 \mu\text{M}$), P1ⁿ ($0.2 \mu\text{M}$), and cT1 ($0.1 \mu\text{M}$), and then rapidly diluted 10-fold into cT1 (final concentration: $1 \mu\text{M}$, CT \uparrow). (D) To prepare

uncorrelated mixtures, preformed CP^{CLOSED}-P1:cT1 complex was diluted into high cT1 ($1 \mu\text{M}$) containing prehybridized P1ⁿ ($0.02 \mu\text{M}$). (E) Correlated (C) and uncorrelated (D) 4-hour extensions with P1⁵ (left) and P1⁴⁰ (right). (F) Ratio of quantified correlated and uncorrelated extensions (>5 nt for P1 and >2 nt for P1ⁿ). (G) CP^{CLOSED}-P1:cT1 immobilized on streptavidin beads were washed for 1 hour and extended with NTPs. (H) Four-hour extensions of cT1 or cT1 prehybridized with P1ⁿ ($n = 5$ or 40 , respectively) by bead immobilized P1:CP^{OPEN}.

consistent with this polymerase lacking a clamping domain (Fig. 4C, solid squares). However, CP showed an ~ 12 -fold higher ER at low primer-polymerase concentrations with both linear and circular T1 (Fig. 4C, solid and empty circles). Truncating the linear T1 to the much shorter sT1 construct resulted in a low ER value that was completely independent of polymerase concentration (Fig. 4C, open squares). This lack of polymerization efficiency could be rescued by adding oligo(A) template sequence either to the 5' terminus (A₅₅sT1, open diamonds) or the 3' terminus (sT1A₆₉, open triangles) of sT1, with the 3' rescue being more pronounced

(Fig. 4, C and D). Similarly, circularizing the sT1 construct with oligo (A) fully rescued polymerization efficiency (cA₁₂₈sT1; Fig. 4C, solid diamonds).

Template extension efficiency was about threefold superior on circular cT1 and cT2 templates relative to their linear counterparts, where efficiency was defined as: $E = \text{Extension}(\text{P1:CP}^{\text{OPEN}} + \text{cTemplate}) / \text{Extension}(\text{P1:CP}^{\text{OPEN}} + \text{Template})$ (Fig. 4E). Further highlighting the importance of correct clamping for efficient polymerization, we eliminated the 5' primer-binding region within the PBS, preventing both the P1:CP^{OPEN} and the CP^{CLOSED} states

from forming. The removal of this region reduced the extension ratio to ~ 1 , as expected (fig. S12). These data are consistent with the promoter triggering correct clamping of the polymerase, forming a processive elongation complex able to extend a range of templates provided they have sufficiently long sequence flanking the specificity-primer:promoter duplex.

The clamped complex is stable and allows extension at multiple primed sites

Extant DNA polymerases use multimeric clamp proteins to facilitate high polymerization rates and processivity (22). Without the clamping

domain of the CP ribozyme, a close relative of the B6.61 polymerase ribozyme has been reported to have negligible processivity, where on average only ~50% of nucleotide extensions results in a second nucleotide being added by the same polymerase on the same template (13). To explore the mechanism of processivity further, we immobilized the P1:CP^{OPEN} complex to streptavidin magnetic beads by hybridizing the ribozyme's 3' terminus to a biotinylated DNA oligonucleotide (fig. S13). The P1:CP^{OPEN} complex was then incubated with a range of templates, and the off-rate of either the radiolabeled P1 or template was measured by scintillation counting. The reverse complement of P1 and the sT1 template nearly quantitatively stripped P1 off the immobilized P1:CP^{OPEN} complex (fig. S13, C to E). All other templates were retained together with P1 on the immobilized CP ribozyme (fig. S13, C to E). Addition of cT1 in two-fold excess retained 64% of P1, whereas the

addition of sT1 retained only 8% after 1 hour of stringent washing (fig. S13D). Likewise, after 4 hours of washing, 41 to 53% of the cA₁₂₈sT1, sT1A₆₉, and A₅₉sT1 templates were retained, whereas only 1% of sT1 remained bound (fig. S13E). Because these templates differ from sT1 only by the addition of oligo(A) residues, the CP clamp must intrinsically operate in a sequence-independent fashion to retain the immobilized complex.

Mechanistically, the formation of the clamped state allows the polymerase to reach and extend primers found a substantial distance away from the promoter-binding site. A set of primers (P1ⁿ; $n = 5, 40, 80, 121$, and 156, where n indicates the distance of the 5' termini of each primer hybridized downstream from the P1 specificity primer promoter start site on cT1) could all be significantly extended by the polymerase (Fig. 5, A and B, and fig. S14). As expected, the P1 specificity primer could only be extended by 5 and 40 nt before

being blocked by the P1⁺⁵ and P1⁺⁴⁰ primers, respectively (Fig. 5 and fig. S14).

When P1:CP^{OPEN} was mixed with cT1 pre-hybridized to the P1⁺⁵ or P1⁺⁴⁰ primers and the resultant complex diluted into 100-fold excess cT1 (Fig. 5E), simultaneous extension of P1 and the P1⁺⁵ or P1⁺⁴⁰ occurred to a much greater extent than if the P1 and P1ⁿ primers were found on distinct templates (Fig. 5, D and E, and fig. S15). The extension ratios for P1⁺⁵ and P1⁺⁴⁰ were eightfold and 11-fold higher when correctly clamped (Fig. 5F), similar to the ratios observed for P1 extension (Fig. 4C). Further, immobilizing P1:CP^{OPEN} onto streptavidin beads (Fig. 5G) resulted in similar extension of templates with P1ⁿ primers, with 75 to 90% of the correctly clamped complexes remaining on the beads after 4 hours of polymerization (Fig. 5H and fig. S16). The three-component complex (CP^{CLOSED}-P1:cT1) is therefore a stable and processive polymerase-primer:template complex.

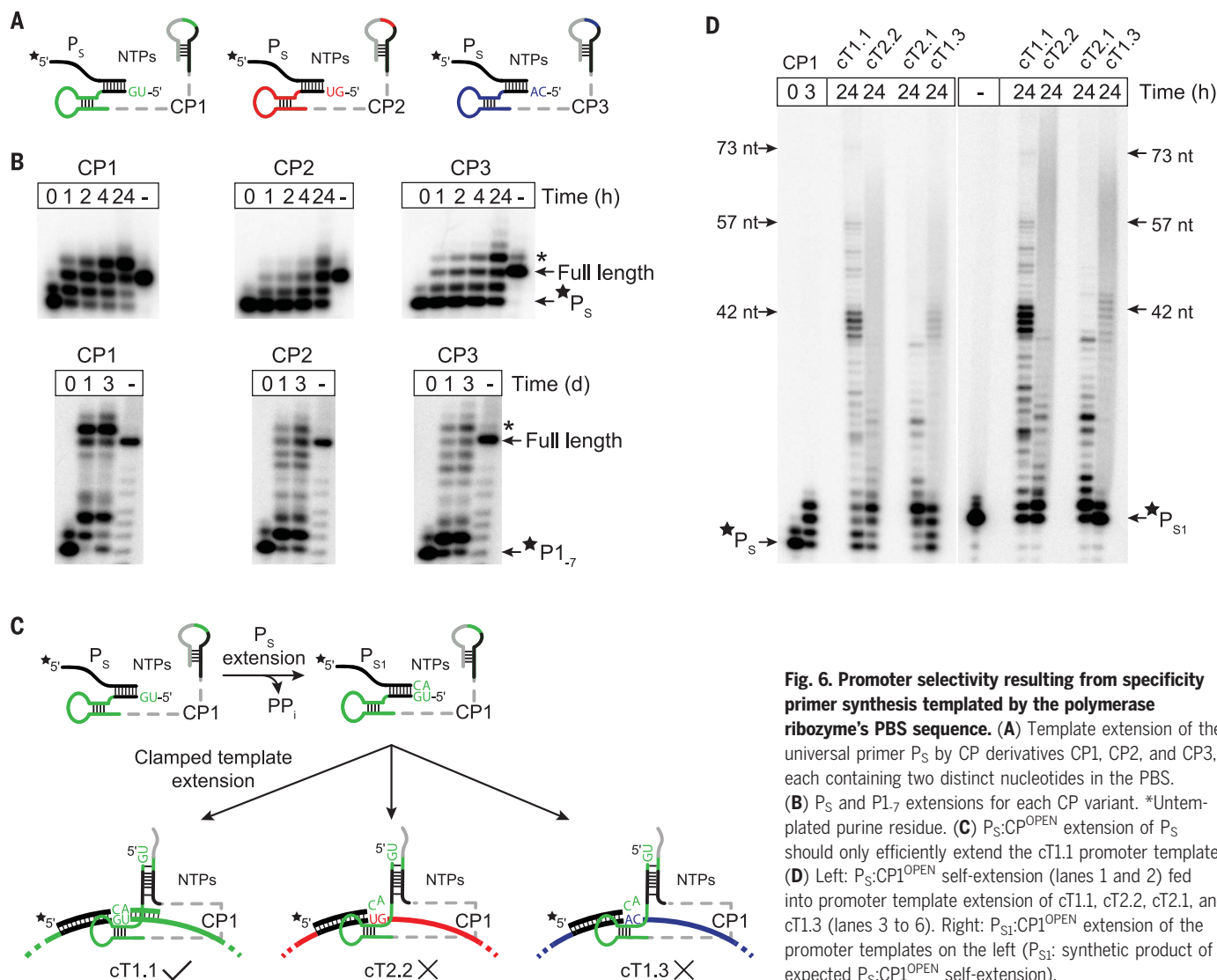


Fig. 6. Promoter selectivity resulting from specificity primer synthesis templated by the polymerase ribozyme's PBS sequence. (A) Template extension of the universal primer P_S by CP derivatives CP1, CP2, and CP3, each containing two distinct nucleotides in the PBS. (B) P_S and $P1_{7.7}$ extensions for each CP variant. *Untemplated purine residue. (C) P_S :CP^{OPEN} extension of P_S should only efficiently extend the cT1.1 promoter template. (D) Left: P_S :CP1^{OPEN} self-extension (lanes 1 and 2) fed into promoter template extension of cT1.1, cT2.2, and cT1.3 (lanes 3 to 6). Right: P_S :CP1^{OPEN} extension of the promoter templates on the left (P_{S1} : synthetic product of expected P_S :CP1^{OPEN} self-extension).

In aggregate, these data indicate that the CP ribozyme uses a sequence-independent topological clamp of the form shown in Fig. 1; however, we cannot rule out the possibility of a “gripping” type clamp model known to form by fusing DNA-binding domains to *Taq* polymerase (23). In either case, an organized transition exists from the open holopolymerase to the closed-clamp form, which allows the active site of the polymerase to extend a broad set of RNA templates often by multiple helical turns. Simultaneously, but only when correctly triggered by a specificity primer, the clamp confers the ability to find and extend primers found at widely spaced locations within a single RNA template.

Programmable promoter recognition by the holopolymerase

The CP ribozyme can use its PBS sequence as a template to extend shortened specificity primers. To find the minimal primer that the CP^{OPEN} complex can extend, we partially hydrolyzed 5′ end-labeled P1. We found that a 7-nt 3′ truncation of P1 (P1₇) was the shortest primer that could be efficiently extended by the CP ribozyme (fig. S17), consistent with the proposed hybridization mechanism of P1 to the CP ribozyme’s PBS domain (Fig. 1D). Three new CP variants were created, each having a specific 5′-terminal dinucleotide sequence appended to the CP ribozyme after first removing its 5′ G. The variants were activated with a shortened universal primer, P_S (fig. S17B), designed to allow two nucleotides of self-templated extension (Fig. 6A and table S3). P_S was extended nearly quantitatively by the CP1 variant (5′ UG template), whereas the CP2 (5′ GU template) and the CP3 (5′ CA template) variants showed lower extension after 24 hours of incubation (Fig. 6B). This extension behavior was quite robust, with all three polymerase constructs also extending the P1₇-truncated primer (Fig. 6B). All polymerase variants added at least one untemplated purine residue (Fig. 6B and fig. S18), similar to many protein repair polymerases such as *Taq*, which add primarily an untemplated A to blunt-ended duplex DNA (24).

Such holopolymerase-dependent extension of the universal primer sequence allows some promoter template RNAs but not others to be copied based on the sequence of the polymerase itself (Fig. 6C). We replaced the P1 and P2 promoters in the cT1 and cT2 constructs with three new promoter sequences corresponding to primer sequences synthesized by the CP1, CP2, and CP3 ribozymes. Forming the P_S:CP1^{OPEN} holopolymerase, the P_S primer was extended by incubation for 3 hours (Fig. 6D). The holopolymerase ribozymes containing the extended P_S primer (P_S-ext:CP1^{OPEN}) were then incubated with the newly constructed promoter templates and the primer extension was

measured. Polymerization by CP1 was ~12-fold better on the cT1 template with the CP1 promoter (cT1.1) compared with the same template having a CP3 promoter (cT1.3; Fig. 6D). Likewise, extension on the cT2 template having a CP1 promoter (cT2.1) was about fourfold superior to extension on the same template with a CP2 promoter (cT2.2; Fig. 6D). Other combinations of CP derivatives and their promoters were less significantly regulated, but these permutations demonstrate that self-templated primer synthesis by the polymerase itself can have a marked effect on selective polymerization ability.

Discussion

The ability of a polymerase to recognize a promoter presents a fundamental evolutionary tension: Molecular recognition of a promoter is a static process, whereas processive polymerization is a dynamic one. Through in vitro evolution, we have found an RNA polymerase that can search for a promoter by first forming a functional open holopolymerase complex and then in a second step rearrange into a processive elongation form. The correct assembly of this CP^{CLOSED} complex results in a more than one order of magnitude increase in extension, with extension on randomly generated templates being directly comparable to the best RNA polymerase ribozymes isolated to date, which on highly repetitive tethered templates are able to synthesize 75 to 203 nt of sequence (8–10).

RNA replication results in long stretches of duplex RNA. Thus, just as in modern biology, the ability of an RNA polymerase ribozyme to invade duplex RNA would be of fundamental importance in early evolution. Although the CP polymerase is incapable of strand invasion (Fig. 5), its processivity and correlated primer extension ability indicates that it entrains templates through a “sticky” topological clamp when correctly clamped (Fig. 1). The CP can synthesize duplexes from ~50 to 107 bp in size, a linear extent ranging from 175 to 360 Å, which is threefold to sixfold larger than its class I ligase catalytic core (25). The polymerase must therefore move while not disengaging from the template because polymerization can occur while washing the immobilized processive complex. Precedent for such a sticky clamp exists in modern RNA biology. The ribosome creates a topological clamp by assembling the large and small subunits around an mRNA. This clamp is stable (26) but allows robust movement of mRNA 3 nt at a time during translation. Similarly, coupling the force generated from NTP incorporation with the embryonic CP clamp could lead to the development of a polymerase ribozyme with ratchet-like and strand invasion capabilities (27).

The CP ribozyme can also synthesize part of its own specificity primer, providing evidence

that a replicase in an RNA World could have avoided replicative parasites by a strategy akin to the genomic tag hypothesis of Weiner and Maizels (28). Compartmentalization has long been recognized as a key element in the solution to this problem (29), but early evolution may have undergone a period during which replicating systems existed without cellularization. In such a situation, a replicase able to synthesize all or part of its own specificity primer could, through mutations to its own sequence, have rapidly evolved a sense of self to avoid replicative parasites early in evolution.

Although many outstanding challenges remain to producing a self-evolving system in the laboratory, including increased polymerization rate, fidelity, and, most importantly, strand displacement, the development of a promoter-dependent RNA polymerase ribozyme with processive clamping ability offers many insights into the dilemmas faced by life in the earliest periods of evolution on this planet.

REFERENCES AND NOTES

- W. Gilbert, *Nature* **319**, 618–618 (1986).
- J. W. Szostak, D. P. Bartel, P. L. Luisi, *Nature* **409**, 387–390 (2001).
- G. F. Joyce, *Nature* **418**, 214–221 (2002).
- D. P. Bartel, J. W. Szostak, *Science* **261**, 1411–1418 (1993).
- W. K. Johnston, P. J. Unrau, M. S. Lawrence, M. E. Glasner, D. P. Bartel, *Science* **292**, 1319–1325 (2001).
- H. S. Zaher, P. J. Unrau, *RNA* **13**, 1017–1026 (2007).
- U. F. Müller, D. P. Bartel, *RNA* **14**, 552–562 (2008).
- A. Wochner, J. Attwater, A. Coulson, P. Holliger, *Science* **332**, 209–212 (2011).
- J. Attwater, A. Wochner, P. Holliger, *Nat. Chem.* **5**, 1011–1018 (2013).
- D. P. Horning, G. F. Joyce, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 9786–9791 (2016).
- Q. S. Wang, L. K. L. Cheng, P. J. Unrau, *RNA* **17**, 469–477 (2011).
- J. T. Szczepanski, G. F. Joyce, *Nature* **515**, 440–442 (2014).
- M. S. Lawrence, D. P. Bartel, *Biochemistry* **42**, 8748–8755 (2003).
- T. M. Gruber, C. A. Gross, *Annu. Rev. Microbiol.* **57**, 441–466 (2003).
- H. Maeda, N. Fujita, A. Ishihama, *Nucleic Acids Res.* **28**, 3497–3503 (2000).
- B. A. Young, T. M. Gruber, C. A. Gross, *Cell* **109**, 417–420 (2002).
- K. S. Murakami, S. A. Darst, *Curr. Opin. Struct. Biol.* **13**, 31–39 (2003).
- G. M. T. Cheetham, T. A. Steitz, *Science* **286**, 2305–2309 (1999).
- Y. W. Yin, T. A. Steitz, *Science* **298**, 1387–1395 (2002).
- T. H. Tahirou et al., *Nature* **420**, 43–50 (2002).
- K. J. Durniak, S. Bailey, T. A. Steitz, *Science* **322**, 553–557 (2008).
- S. J. Benkovic, A. M. Valentine, F. Salinas, *Annu. Rev. Biochem.* **70**, 181–208 (2001).
- Y. Wang et al., *Nucleic Acids Res.* **32**, 1197–1207 (2004).
- G. Hu, *DNA Cell Biol.* **12**, 763–770 (1993).
- D. M. Shechner et al., *Science* **326**, 1271–1275 (2009).
- J. A. Steitz, *Nature* **224**, 957–964 (1969).
- L. K. L. Cheng, P. J. Unrau, *Cold Spring Harb. Perspect. Biol.* **2**, a002204 (2010).
- A. M. Weiner, N. Maizels, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 7383–7387 (1987).
- G. F. Joyce, J. W. Szostak, *Cold Spring Harb. Perspect. Biol.* **10**, a034801 (2018).

ACKNOWLEDGMENTS

We thank I. Yaseen and members of the Unrau laboratory; M. Leroux (SFU), D. Sen (SFU), and U. Müller (UC San Diego) for critical reading of the manuscript; and E. Yingqi Han (software

engineer) for help with processing high-throughput sequencing data. **Funding:** This work was supported by a NSERC Discovery Grant to P.J.U. **Author contributions:** R.C. and P.J.U. designed the experiments, analyzed the data, and wrote the manuscript. R.C. performed the experiments. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/371/6535/1225/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S19
Tables S1 to S4
References (30, 31)

Sequenced DNA Libraries
MDAR Reproducibility Checklist

[View/request a protocol for this paper from Bio-protocol.](#)

3 August 2020; accepted 4 February 2021
10.1126/science.abd9191

Processive RNA polymerization and promoter recognition in an RNA World

Razvan Cojocaru and Peter J. Unrau

Science **371** (6535), 1225-1232.
DOI: 10.1126/science.abd9191

A processive RNA replicator

The RNA World Hypothesis suggests that, before modern life, there were RNA molecules that were capable of carrying genetic information and driving chemical reactions, a task gradually replaced by DNA and enzymes in modern biology. Central to this theory is an RNA replicase capable of mediating general replication of RNA. Using laboratory evolution, Cojocaru *et al.* isolated a promoter-based RNA polymerase ribozyme that, analogous to modern-day protein polymerases, clamps onto templates to increase its processivity, making it a potential model for replication in early biology.

Science, this issue p. 1225

ARTICLE TOOLS

<http://science.sciencemag.org/content/371/6535/1225>

SUPPLEMENTARY MATERIALS

<http://science.sciencemag.org/content/suppl/2021/03/17/371.6535.1225.DC1>

REFERENCES

This article cites 31 articles, 16 of which you can access for free
<http://science.sciencemag.org/content/371/6535/1225#BIBL>

PERMISSIONS

<http://www.sciencemag.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of Service](#)

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

Copyright © 2021 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works