PYROPHOSPHATE-MEDIATED REPAIR OF DAMAGED AND MISMATCHED RNA BY A POLYMERASE RIBOZYME

Alexandra D. Kent,[a] Lucy L. Yang,[a] Ivee A. Lee,[b] and Andrej Lupták[a, b, c]*

[a] Department of Chemistry, University of California at Irvine, Irvine, CA 92617
[b] Department of Pharmaceutical Sciences, University of California at Irvine, Irvine, CA 92617
[c] Department of Molecular Biology and Biochemistry, University of California at Irvine, Irvine, CA 92617

* Correspondence: E-mail: aluptak@uci.edu

Abstract

Prior to the emergence of the contemporary biosphere, the early Earth is likely to have progressed through an RNA-based world. In such an evolving world, it would have been critical to have a system of genome replication that allowed heritable information to be transferred with an early catalyst, such as an RNA polymerase ribozyme. Though substantial effort has gone into evolving such a polymerase, many variants suffer from premature termination and low fidelity, resulting in low yields of full-length sequences and small quantities of active sequences, respectively. The ability to replicate and maintain a genome at a higher fidelity would lend an evolutionary advantage to an early evolvable system. Cells today achieve this through a variety of means, including complex damage repair mechanisms and proofreading pathways, but such dedicated machinery would likely not exist in an early evolving world. However, if an early polymerase possessed repair activity under specific conditions, replication would have been improved. Here we demonstrate the first example of ribozyme-mediated repair of damaged and mismatched RNA sequences. Under conditions of high pyrophosphate concentrations, we show that a polymerase ribozyme can repair RNA sequences terminated in a 2′ -3′ cyclic phosphate, a mismatch, or both, to allow further polymerization along a template. In many cases, the fidelity of the repaired RNA extension is significantly higher than priming from a mismatched base-pair; this observation suggests that mutations may have an allosteric effect on the fidelity of downstream replication. Finally, we show that the pyrophosphorolysis reaction directly yields a nucleoside triphosphate, providing the first evidence for ribozyme-catalyzed production of ATP. Increased fidelity and processivity of polymerization advances the longstanding goal of developing a self-replicating polymerase ribozyme.

Introduction

The cellular functions of ribonucleic acids (RNAs) include critical roles in catalysis, translation, and metabolite-dependent gene expression regulation, and as such, it may have been the first biopolymer to store information and catalyze chemical reactions on an early Earth.[1–5] Replication is a process fundamental to life and is the mechanism by which heritable information is stored and propagated. In the extant biosphere, this is achieved through protein-based polymerase enzymes, but in an RNA World—prior to the advent of genetically encoded proteins—replication would likely have been carried out by RNA polymerase ribozymes.[6,7] These ribozymes would need to have sufficient activity and fidelity to build up an information storage system and could serve as the central molecules linking the phenotype of an early protocell to its genotype.[8,9]

Transfer of the α-phosphate of a triphosphorylated nucleoside to a 3′ hydroxyl of another nucleotide, with concomitant release of pyrophosphate, is the central reaction catalyzed by enzymes replicating phosphodiester-based informational polymers. Beginning with the first *in vitro* selection of a ligase ribozyme, a large body of work using molecular evolution and engineering has sought to isolate ribozymes that catalyze this reaction.[10–14] The ligase ribozyme has continued to be evolved into progressively more sophisticated polymerases capable of promoter-based polymerization, production of other functional nucleic acids, reverse transcription, and transcription of non-canonical (xeno-) nucleic acids, among other functions.[15–21] However, due to premature termination and degradation, even the most advanced polymerase ribozymes suffer both from mutation burdens that reduce the activity of produced transcripts and from low yields of full-length products.[22–24] Thus, a mechanism for repairing both degraded and mismatched sequences would lend an evolutionary advantage to an evolving system. Very few mechanisms of RNA repair have been reported, and none have been ribozyme-mediated.[25–27] Additionally, little is known about how pyrophosphate affects the activity of these ribozymes and the related stability of the transcripts.

We present a study of one of the RNA polymerases, the tC19Z ribozyme, and the effect of pyrophosphate on its activity. We demonstrate that addition of pyrophosphate and the polymerase ribozyme to 2′-3′ cyclic phosphate-terminated (damaged) RNA sequences facilitates damage repair and subsequent RNA extension. We also show that pyrophosphorolysis of mismatched RNA sequences yields a nucleoside triphosphate and allows for repair and alteration of extension profiles. Finally, using sequence analysis of mismatched and repaired RNAs, we demonstrate that repair of terminal mismatches improves downstream replication fidelity, providing a mechanism for replication of complex RNAs by a single polymerase ribozyme.

Results

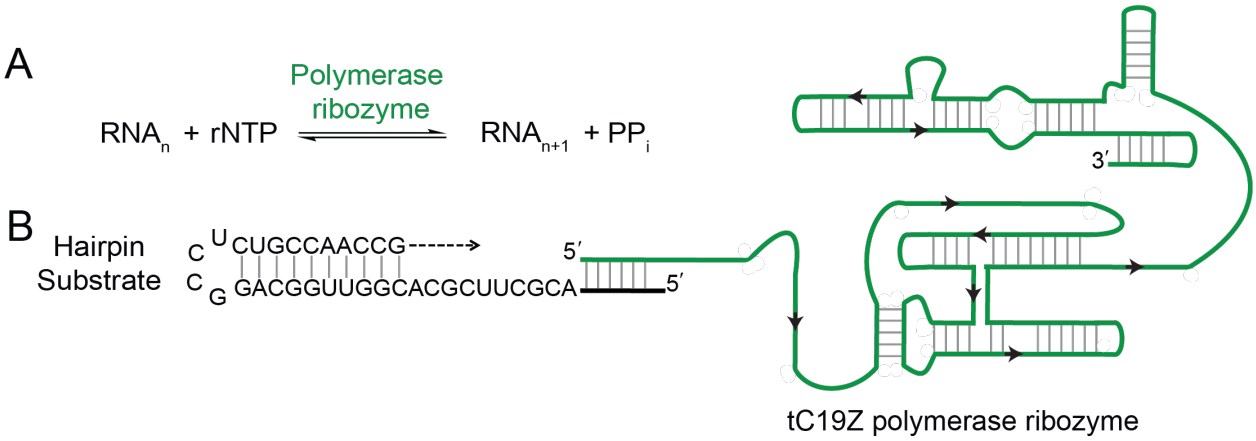

Figure 1. Reversible RNA synthesis by the polymerase ribozyme. (A) Templated addition of one or more nucleotides to the 3′-end of a single-molecule primer–template (hairpin) RNA catalyzed by a polymerase ribozyme with pyrophosphate as a byproduct of the reaction. The reverse reaction, pyrophosphorolysis of RNA, removes the last nucleotide, yielding a nucleoside triphosphate. (B) An *in vitro* evolved polymerase ribozyme, tC19Z (green), shown non-covalently attached to the template strand of a hairpin substrate.

A number of polymerase ribozymes have been developed to extend RNA sequences using nucleoside triphosphates in a templated manner; these reactions generate pyrophosphate as a byproduct (**Fig. 1A**). One of the most active sequences, the tC19Z ribozyme, was selected *in vitro* from an earlier polymerase (R18), which in turn originated from the class I RNA ligase.[16,18] The tC19Z sequence was evolved to have improved sequence generality and polymerase activity, although the template sequence also has a significant influence on replication efficiency. To extend a sequence, the 5′ end of the ribozyme base-pairs with the 5′ end of the template sequence (**Fig. 1B**), which is bound to a primer at its 3′ end. The primer and template are fused to form a hairpin structure (**Fig. 1B**) that serves both as a substrate for the ribozyme at exact stoichiometry and as a practical construct for downstream analysis using high-throughput sequencing.

Pyrophosphate-mediated repair of 2′-3′-cyclic-phosphate-terminated RNA sequences.

The centrality of phosphate in the origin of life stems from its presence in the backbone of RNA and DNA and is thus inextricably coupled to genome replication in contemporary biology.[28] Pyrophosphate is the byproduct of RNA polymerization but has also been suggested as a source of available phosphate on an early Earth.[29–32] We hypothesized that addition of pyrophosphate to the polymerization reaction would shift the equilibrium of the reaction and induce a low level of pyrophosphorolysis (**Fig. 1A**), which would be beneficial for damaged sequences that cannot be extended further. To investigate this mechanism of RNA damage repair, we employed sequences terminated with a 2′-3′-cyclic-phosphate (2′-3′-cP). Such sequences are degradation products formed *via* a nucleophilic attack by a 2′ oxyanion onto the adjacent phosphodiester. This reaction is common to self-cleaving ribozymes, many RNases, and base-catalyzed RNA degradation,[33–35] and is responsible for the relatively low stability of RNA, when compared with DNA or with nucleic

acids that lack a phosphate-adjacent hydroxyl group. Because this type of damage is the most common form of RNA degradation, we devised an RNA hairpin with terminal 2′-3′-cP damage to serve as a primer–template substrate of the polymerase ribozyme and tested the repair ability of the polymerase ribozyme. We generated the hairpin *via* an *in vitro* transcription reaction of a self-cleaving HDV-like ribozyme (drz-Mbac-1) fused to the 3′ terminus of the target sequence (**Fig. S1**).[36] The ribozyme self-cleaves, yielding a 2′-3′-cP on the 3′ end of the "primer" segment. Because the 2′-3′-cP acts as a chain terminator and prevents the 3′ hydroxyl from extending during polymerization, it represents a replication dead end (**Figs. 2A** and **2B**). The cyclic phosphate can be removed enzymatically by T4 polynucleotide kinase (PNK), and this combination was used as a positive control for testing the extension efficiencies of the hairpins (**Fig. S1**).[37,38]

We prepared RNA hairpins terminated with a 2′-3′-cP on each of the four canonical nucleotides base-paired to the matching nucleotide on the template strand. As expected, no polymerization by the ribozyme was observed when extension of the 2′-3′-cP-damaged sequence was attempted. However, upon addition of pyrophosphate, the terminal damaged nucleotide was removed, and each of the sequences could be extended by the polymerase ribozyme (**Figs. 2D-G, S1** and **S2**). The sequences could also be extended following removal of the 2′-3′-cP by PNK. The extension efficiency, as well as repair efficiency, was largely dependent on the sequence; U/A (U on the primer strand and A on the template strand) was repaired the most efficiently. However, we observed that addition of pyrophosphate did inhibit the forward polymerization reaction (**Figure S3** and **S4**). We measured the repair and extension by both high-resolution polyacrylamide gel electrophoresis (PAGE) and high-throughput sequencing (HTS) (**Fig. 2C** and **S2**).

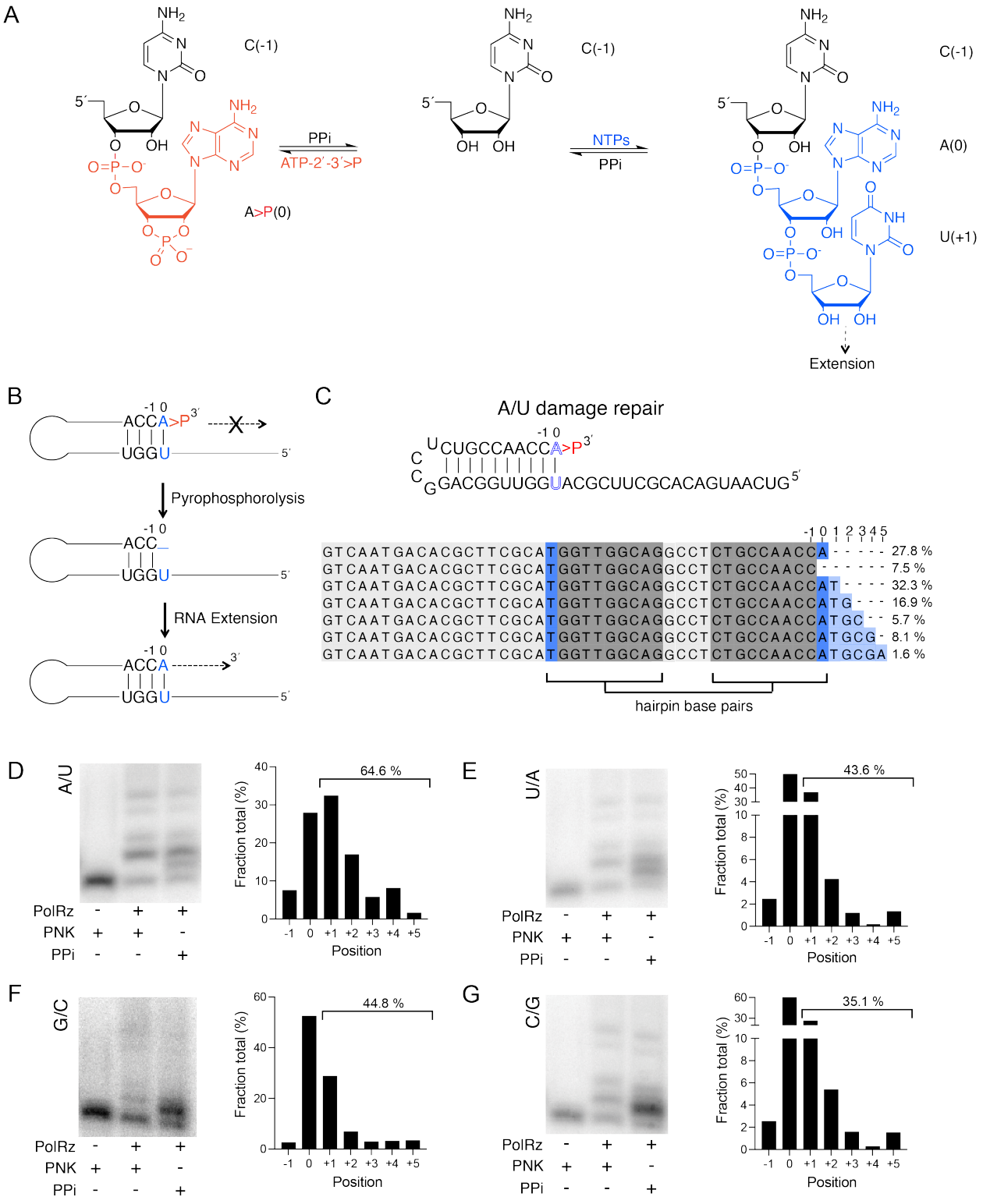

Figure 2. Pyrophosphate-facilitated repair of an RNA damage site by a polymerase

ribozyme. (**A**) A hairpin substrate terminated with a 2′-3′-cyclic phosphate (2′-3′-cP; red nucleotide) cannot be extended by the polymerase ribozyme. However, upon removal of the terminal adenosine by ribozyme-mediated pyrophosphorolysis, the revealed 3′-OH becomes available for further addition of nucleoside monophosphate (blue). (**B**) Graphical representation of an A/U-terminated hairpin with 2′-3′-cP damage and subsequent repair *via* pyrophosphorolysis. (**C**) Full sequence of an A/U hairpin and aligned reads from high-throughput sequencing analysis showing removal of the terminal adenosine (dark blue), and subsequent extension up to five bases (light blue). Percentages on the right indicate the fraction of each sequence in the dataset. The numbering above the alignment represents terminal nucleotides defined in (**A**). (**D**–**G**) Analysis of ^{32}P-ATP-labeled 2′-3′-cP-damaged hairpins terminated in all four base-pairs; terminal base-pairs are indicated alongside each PAGE analysis (left-hand image of each set). Lane 1 shows the starting material after removal of 2′-3′-cP by PNK. Lane 2 is a positive control showing extension of the hairpin by the polymerase ribozyme (PolRz) following 2′-3′-cP removal by PNK. Lane 3 shows the extension products that result from repair of the hairpins by the polymerase ribozyme using pyrophosphate. Graphs on the right of each panel show the percentage of each sequence as determined from HTS data, starting with sequences in which the damaged terminal nucleotide has been removed (fraction of sequences ending at position –1) and proceeding to extensions of the repaired RNA (0–5). The starting RNA containing the 2′-3′-cP cannot be ligated to the primer used for sequencing and is therefore not included in this analysis.

We hypothesized that a plausible mechanism of repair involves pyrophosphorolysis of the 3′ terminal 2′-3′-cP-damaged nucleotide. Therefore, we analyzed the pyrophosphate-mediated removal of a nucleotide by high-resolution PAGE, but did not observe a gel mobility change that could unequivocally be attributed to removal of a single nucleotide from the 2′-3′-cP-terminated hairpin. Comparison of relative gel mobility of the 2′-3′-cP-terminated hairpins treated with acid

(which initially hydrolyzes just one of the phosphodiesters of the cyclic phosphate) with PNK-treated hairpins confirmed that hairpin sequences migrate faster upon removal of the 2′-3′-cP (**Fig. S5**). Previous reports showed that treatment of 2′-3′-cP-terminated oligonucleotide sequences with PNK to remove the cyclic phosphate leads to a gel mobility retardation, and treatment of a 2′-3′-cyclic-phosphate-terminated sequence with acid to generate a linear monophosphate leads to a faster-migrating species.[39–41] We consistently observed the opposite effect and have concluded that the anomalous migration is a property of the hairpin sequence, which differs structurally from the linear RNA sequences used in previously reported experiments.

To analyze pyrophosphate-mediated repair of damaged RNAs, we began by performing repair experiments that could indirectly indicate the removal of the terminal nucleotide. First, we omitted the rNTP that corresponded to the damaged nucleotide in the extension reaction. For example, in the case of an A/U hairpin, we omitted ATP from the rNTPs mixture, because following excision of the terminal AMP-2′-3′-cP, the ribozyme would need to replace it with undamaged AMP to continue the extension reaction. As expected, we found that if ATP was not present in the reaction, the sequence was not extended (**Fig. S6**). To gather additional evidence of damage repair, we transcribed hairpins without ^{32}P and subsequently performed the polymerization reaction with α-^{32}P-NTPs to reveal incorporation of the ^{32}P tracer. We observed that the α-^{32}P-NTP is only utilized in sequences that have been repaired, and although the ribozyme can incorporate it as a mismatched nucleotide during the extension, this occurs with lower efficiency (**Fig. S7**). Finally, we observed no reaction in the presence of only uridine and pyrophosphate, indicating that a non-canonical reaction, such as opening of the 2′-3′-cP by pyrophosphate to form a triphosphate that can subsequently react with unphosphorylated uridine, did not take place (**Fig. S8**).

Pyrophosphate-mediated mismatch repair

To maintain fidelity of genome copying at non-deleterious levels, modern replication machinery uses distinct active sites or entire protein complexes to correct mismatched base-pairs, but at the onset of evolvable systems, it is unlikely that mismatch repair was available as an independent biochemical activity.[42,43] In addition to damage repair, pyrophosphate-mediated mismatch repair would also have been a desirable feature of an early replicating polymerase, lending itself to increased replication fidelity. To test whether pyrophosphate could be used by a polymerase ribozyme to remove a mismatched nucleotide, we prepared hairpin sequences with 3′-terminal mismatches. Using an A/C mismatch, we added pyrophosphate to the 2′-3′-cP-terminated hairpin, but observed neither ribozyme-catalyzed repair nor subsequent extension of the sequence. RNA polymerization by the ribozyme was still not observed upon removal of the 2′-3′-cP by PNK and addition of NTPs, demonstrating that an A/C mismatch acts as a chain-terminator. However, after both removal of the 2′-3′-cP and addition of pyrophosphate, we observed ribozyme-mediated extension of the sequence (**Fig. 3A-C**).

Removal of the terminal nucleotide through pyrophosphorolysis yields a nucleoside triphosphate. Unlike in 2′-3′-cP-terminated hairpins, an A/C-terminated hairpin would yield a molecule of ATP upon pyrophosphorolysis. To detect this critical metabolite, we used an ultrasensitive firefly-luciferase-based assay to measure the ATP generated in the reaction (**Fig. 3D**).[44] Because pyrophosphate inhibits the luciferase reaction, insoluble pyrophosphate was first removed from the samples by centrifugation, and the remaining pyrophosphate was subsequently hydrolyzed by addition of pyrophosphatase. The resulting samples were incubated with the firefly luciferase and D-luciferin and then were imaged using an ultracooled EM-CCD camera. The integrated luminescence was measured for both the pyrophosphorolysis reaction and a control containing no pyrophosphate. In the absence of pyrophosphate we observed a baseline level of luminescence; however, when pyrophosphate was added to the sample the luminescence was significantly greater indicating the presence of ATP. These experiments provide direct evidence

for the pyrophosphorolysis reaction and, for the first time, demonstrate ribozyme-catalyzed production of ATP.

To further analyze the repair reactions, we compared our PAGE analyses and ATP detection with sequencing of each of the mismatched-terminated hairpin reactions. HTS analysis revealed that pyrophosphorolysis of the substrate yields approximately 100-times higher concentrations of a hairpin lacking the terminal nucleotide than the no-pyrophosphate controls (**Fig. S9**), providing a robust experimental confirmation for the pyrophosphorolysis of the terminal nucleotide. Extensions of the hairpins correlated well with the gel electrophoresis analysis; however, the HTS data allowed us to measure the fractions of the sequences that were repaired and subsequently extended (**Fig. 3C**). We found that the efficiency of repair and extension depends on the type of mismatch: purine/purine mismatches were repaired and extended at low frequency upon the addition of pyrophosphate, but were not extendable by the ribozyme alone. In contrast, pyrimidine/pyrimidine mismatches were extended well by the ribozyme, but almost all extension products retained the mismatch (**Figs. 3E** and **3F**; **Fig.S10**). Purine/pyrimidine mismatches were repaired the most efficiently. Overall, the fraction of the mismatches that was repaired in each case corresponded well to the extensions observed by PAGE (**Figs. 4A** and **4C; Fig. S11**).

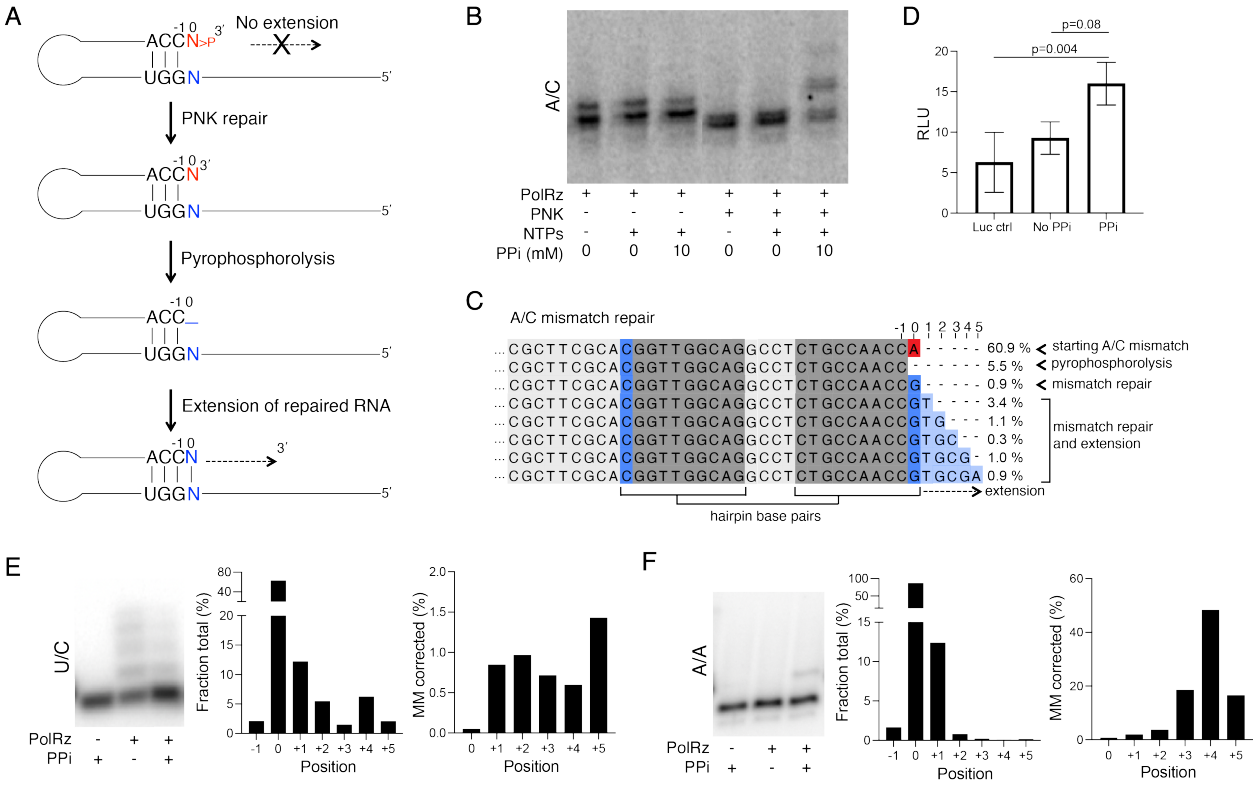

Figure 3. Pyrophosphate-mediated mismatch repair and ATP generation by a polymerase ribozyme. (**A**) Graphical representation of a mismatched hairpin (shown as red and blue nucleotides) with 2′-3′-cP damage (>P), repair by PNK to remove the 2′-3′-cP and ribozyme-catalyzed pyrophosphorolysis to remove the mismatched nucleotide, and subsequent RNA polymerization with the correct nucleotide. (**B**) High-resolution PAGE analysis of a ^{32}P-labeled A/C mismatched hairpin. The left three lanes show that no extension is observed while the 2′-3′-cP remains on the hairpin. Upon removal of the 2′-3′-cyclic phosphate by PNK, the ribozyme can extend the RNA, but only if both pyrophosphate and rNTPs are present. (**C**) Sequences of an A/C mismatched hairpin generated from HTS data showing the starting A/C mismatch (red A), the removal of the mismatched nucleotide by phosphorolysis (second row sequence comprising 5.5% of total hairpin reads), the addition of the correct nucleotide (blue G at position 0), and the extension of the corrected sequence (light-blue boxes at positions 1–5). Percentages indicate the fraction of each sequence determined by HTS. Total percentage of sequences extended are

shown in Figure S10. (**D**) Firefly luciferase assay measuring ATP generated by the removal of the terminal adenosine of an A/C mismatched hairpin. Luc control is a reaction containing only the luciferase enzyme and D-luciferin, no-PPi is a polymerase reaction lacking pyrophosphate, and PPi is a polymerase-catalyzed pyrophosphorolysis of the terminal adenosine monophosphate to yield ATP. Error bars represent s. e. m. of three imaging replicates. The p-values are calculated from a two-tailed analysis. (**E, F**) High-resolution PAGE analysis of ^{32}P-labeled hairpins terminated in a U/C (**E**) or A/A (**F**) mismatch (shown left in each panel). Left lane shows no extension in the absence of the polymerase ribozyme, middle lane shows extension of the hairpin by the polymerase ribozyme in the absence of pyrophosphate, and right lane shows the reaction in the presence of pyrophosphate. The graphs show the distribution of sequences upon pyrophosphorolysis (middle panels) and the percentage of the extended products in which the mismatch was repaired (right panels), as derived from HTS data.

For certain hairpins, we observed both damage and mismatch repair upon addition of pyrophosphate. Specifically, the purine/purine mismatches (A/A, G/A, and A/G) and the G•U wobble pairs (G/U and U/G) could be repaired by the ribozyme when pyrophosphate was added to the extension reaction of these 2′-3′-cP-terminated mismatches (**Fig. 4B** and **Fig. S11**). As was observed with the matched hairpins, the extension efficiencies for the mismatched hairpins depended on the terminal nucleotide, both after 2′-3′-cP removal by PNK or the removal of the terminal nucleotide following treatment with PNK and pyrophosphate. The pyrimidine/pyrimidine (C/C, U/U, C/U, and U/C) mismatches were best extended by the ribozyme following 2′-3′-cP removal by PNK. The extension was unaffected by the addition of pyrophosphate, indicating that the polymerase is more tolerant of these mismatches over others. In contrast, sequences containing terminal C/A, U/G, or G/U mismatches behave like the A/C mismatch and only showed extension (with varying efficiencies) after cyclic phosphate removal and pyrophosphorolysis.

Critically, the sequencing data provided quantitative results for the fidelity of extension products. We measured the fidelity at each position of the extension, for both long and short extension products, and found that for all mismatches, repair of the mismatched base led to an increase in the fidelity of downstream extensions (**Fig. 4D**). In some cases, such as the pyrimidine/pyrimidine mismatches, the effect is strongly pronounced: although the frequency of repair is low, when a hairpin is repaired, the extension occurs with very high fidelity. Therefore, we conclude that mismatched sequences that do not terminate the extension are more likely to acquire additional mutations, implying a mechanism for error propagation that must involve an allosteric modulation of the fidelity of incoming nucleotide incorporation by upstream mismatches.

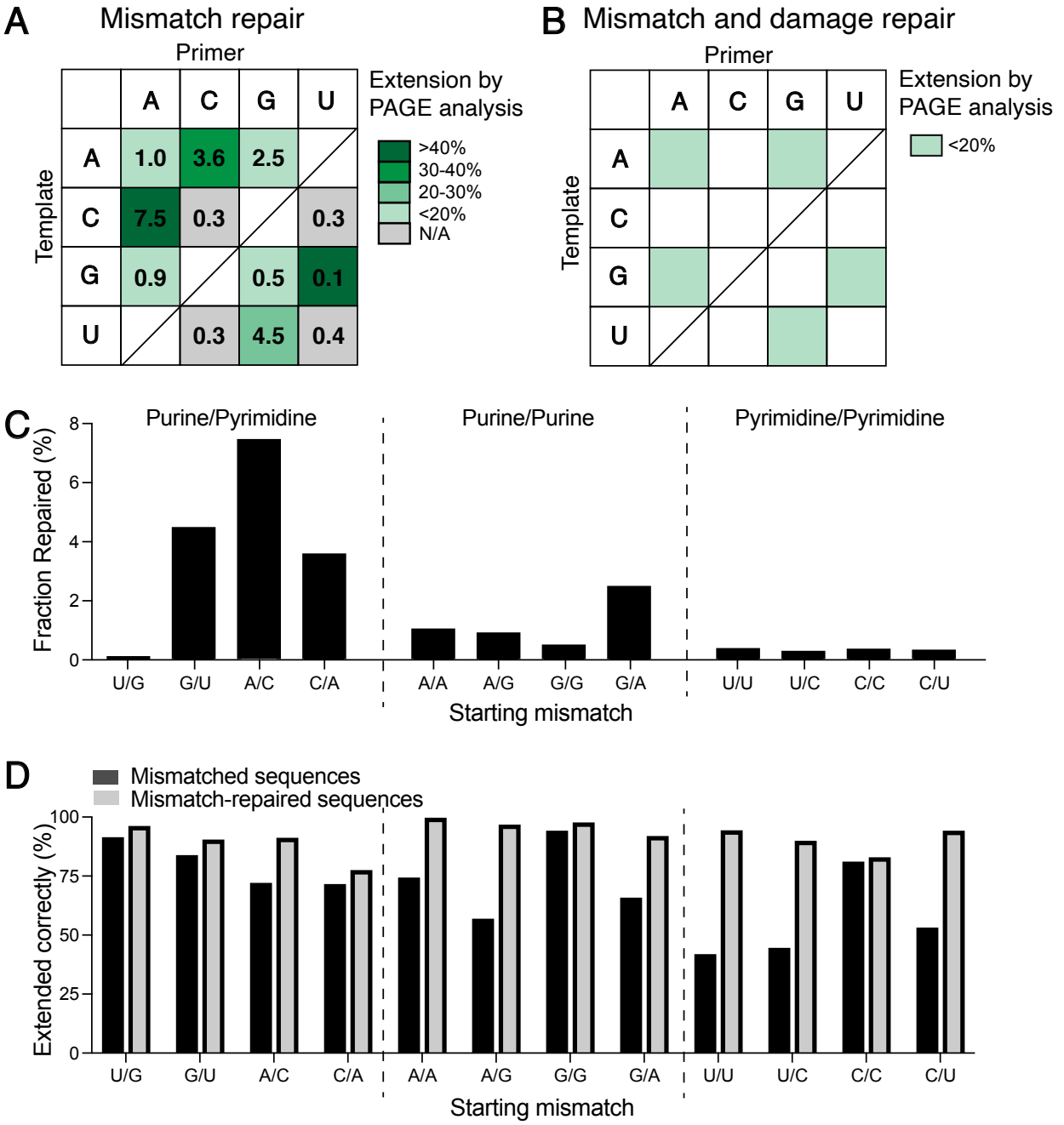

Figure 4. Trends in mismatch repair and fidelity. (**A**) Grid showing the correlation between mismatch repair and overall mismatch extension. The box colors indicate what percentage of each mismatch is extended when repaired by the ribozyme in the presence of pyrophosphate, compared to a no-pyrophosphate control, as determined by PAGE analysis. The gray color indicates no significant difference observed in extension reactions in the presence/absence of

pyrophosphate. The numbers correspond to the total fraction of mismatch repair derived from HTS data. (**B**) Grid showing mismatches in which both 2′-3′-cP phosphate damage and the mismatch could be repaired *via* ribozyme-catalyzed pyrophosphorolysis of the terminal nucleotide. (**C**) Graph showing the fraction of each mismatch repaired via ribozyme-catalyzed pyrophosphorolysis. The mismatches are grouped by base pair type. (**D**) Graph of the fidelity of extensions, comparing the fractions of the correctly extended sequences in sequences containing a mismatch (black bars) and the correctly extended sequences that are extended after mismatch repair (gray bars, indicating the fidelity of all sequences that follow the mismatch-repaired position). In all cases, pyrophosphate-mediated mismatch repair leads to an increase in fidelity of extended sequences. The large differences between the dark and light bars imply that unrepaired mismatches lead to further errors in RNA polymerization.

Discussion

Modern biological systems have numerous, distinct complexes and pathways dedicated to repairing each type of genome damage. Early replicating systems in an RNA world would not have had those functionalities or enzymatic diversity available to them. However, in order to replicate long genomes, damage and mismatch repair would likely have been necessary. It has been widely hypothesized that in order for known polymerase ribozymes to truly achieve efficient self-replication capability, there would need to be an increase in the current standard of fidelity of replication. Unfortunately, this has been an elusive trait for even the most evolved polymerase ribozymes. Though early studies on the first polymerase ribozymes hypothesized that pyrophosphorolysis could improve fidelity, this reactivity was never detected.[45] We address this gap by demonstrating that a highly evolved polymerase ribozyme does possesses repair activity in the presence of pyrophosphate. The tC19Z polymerase ribozyme was originally selected to

extend a primer along a template in an effort to generate a system of self-sustained replication.[18] We have demonstrated that the tC19Z polymerase ribozyme can use pyrophosphate to repair 2′-3′-cyclic-phosphate-damaged RNA sequences that are formed *via* base-catalyzed RNA cleavage, lending itself to a more stable early genome. In addition, tC19Z can repair mismatched sequences, which may present a solution to the fidelity problems faced by early replication ribozymes. Although many ribozymes exhibit varying levels of replication efficiency, tC19Z is the first capable of repairing RNA. As such, this study demonstrates that a single ribozyme can perform several distinct biochemical functions, likely with the same active site.[16,19,23] Such economy of function may be a hallmark of emerging genome-encoded functional macromolecules before dedicated synthesis and repair machinery evolved *via* gene duplication. These results indicate that we may be one step closer to generating truly self-replicating systems that could have evolved in an RNA world.

Although our findings established that tC19Z can repair terminal mismatched sequences, the extent of ribozyme-mediated repair is highly dependent on mismatch identity, and distinct trends are present. Previously published findings suggest that even less-evolved polymerase ribozymes (more closely resembling the class I ligase) have nucleotide preferences and add mismatches or extend mismatches with varying efficiencies.[15] Our own observations on single nucleotide incorporation and patterns of extended sequences show that tC19Z has preferences for some nucleotides and sequences over others. For example, the ribozyme will add to and repair damaged uridine nucleotides in preference to the other nucleotides. Based on the mismatch repair trends summarized in Figure 3, we hypothesize that there is a structural basis for the trends in ribozyme-mediated extension of mismatched sequences and the effect of pyrophosphate on extension. Specifically, the width of the terminal base-pair appears to affect the recognition and polymerization capabilities of the ribozyme.[46–50] However, additional structural data for this ribozyme-hairpin RNA complex are needed to definitively conclude what the ramifications of

mismatches are on the ribozyme's polymerization reactivity.

Our sequencing analysis reveals a connection between downstream fidelity and mismatch repair: a substantial increase in downstream fidelity of extended sequences is observed when mismatches are corrected. Some mismatches negatively affect the fidelity of downstream polymerization more prominently than others, but the fidelity can be restored by correcting the terminal mismatch. However, this result also indicates that during replication, if the ribozyme inserts one incorrect base, it is more likely to add additional mismatches. For example, if the U/C mismatched hairpin is not repaired, over 60% of subsequently extended sequences have a second mismatch in the +1 position, and the majority involve addition of a cytosine to make a C/A mismatch. Ultimately, compounding downstream mismatches leads to termination of replication, which would result in genome catastrophe in an evolving world.

Finally, we were able to visualize the formation of ATP following pyrophosphorolysis of an A/C mismatched hairpin. Although non-enzymatic NTP synthesis has been widely explored in an origins-of-life context, there have been few examples of ribozyme-generated nucleoside triphosphates, and this has been a goal separate from that of replication.[44,51,52] With the ultrasensitive ATP assay, we were able to detect ATP concentrations above those found in a no-pyrophosphate control. Taken in concert, these findings suggest that the ribozyme is using pyrophosphate to remove the 2′-3′-cP-terminated nucleotide or mismatched nucleotide from the hairpin, thereby repairing the sequence prior to extension and generating NTPs in the process. Therefore, RNA generated from other activated monomers, such as methylimidazole-based nucleosides, could be degraded by a tC19Z-like polymerase and pyrophosphate to yield nucleoside triphosphates. These newly created, high-energy NTPs could foster the genetic takeover necessary to transition to a contemporary triphosphate-based metabolism and replication system.

Methods

Cloning tC19Z into a pUC19 vector.

The insert, tC19Z, was synthesized and obtained from Integrated DNA Technologies and then was subsequently amplified with primers containing the restriction enzyme sites BamHI and EcoRI (**Table S1**).[18] The tC19Z insert and pUC19 vector were double digested with BamHI and EcoRI for four hours. Additionally, the pUC19 vector was further treated with calf-intestinal, alkaline phosphatase (CIP) for one hour following the double digestion. The insert was ligated into the vector using T4 DNA ligase and incubating at 16 ºC for 16 hours. The resulting ligation was transformed into DH5α competent cells and subsequently plated on ampicillin-containing agar plates which were grown at 37 ºC for 14 hours. Single colonies were picked and grown in 3 mL of ampicillin containing LB broth for 14 hours. Cultures were subjected to alkaline lysis and the plasmids were purified. The presence of tC19Z in the vector was confirmed by Sanger sequencing. Future PCR reactions of tC19Z were performed fresh from the plasmid for each batch of ribozyme.

PCR and transcription of tC19Z and generation of hairpin constructs.

Polymerase chain reaction (PCR) was performed to isolate tC19Z from the plasmid prepared above. 5 ng of the plasmid, 0.5 μM of each primer (**Table S1**), and *Taq* DNA polymerase were used to amplify the sequence at an annealing temperature of 49 ºC. PCR products of tC19Z from the plasmid prepared above were used for *in vitro* transcription. Transcription was performed under the following conditions: 50 mM Tris HCl, 2 mM spermidine, 12 mM $MgCl_2$, 8 mM DTT, 2 mM each rNTP, and 0.01% Triton X-100 in a total volume of 100 μL. T7 RNA polymerase was added to the reaction and allowed to react at 37 ºC for 3 hours. The transcriptions were desalted with G15 Sephadex and purified using PAGE [18% w/v acrylamide/bisacrylamide 19:1 solution, 7 M urea, and 0.5x TBE Buffer (50 mM tris(hydroxymethyl)aminomethane hydrochloride, 45 mM boric acid, and 0.5 mM EDTA, pH 8.4)]. RNA was detected by UV shadowing and then was (i)

eluted in 300 μL of 300 mM KCl from gel pieces, (ii) precipitated by adding 700 μL of ethanol, and (iii) incubated at –20 ºC overnight. The samples were then centrifuged to pellet the precipitate, which was subsequently dried in air and resuspended in water.

To prepare the RNA hairpin substrates for the tC19Z polymerase ribozyme (**Table S1**), the hairpin sequence (0.5 μM) was generated with the drz-Mbac-1 HDV-like ribozyme (0.5 μM) using primer extension (**Table S1**).[36] Following amplification, the sequence was transcribed in the presence of α-^{32}P-ATP by adding 50 mM Tris HCl, 2 mM spermidine, 12 mM $MgCl_2$, 8 mM DTT, 2 mM GTP, 2 mM CTP, 2 mM UTP, 200 μM ATP, and 0.01% Triton X-100 to a 50 μL reaction containing T7 RNA polymerase, and the mixture was allowed to react at 37 ºC for 2 hours. The HDV-like ribozyme self-cleaves co-transcriptionally to yield cyclic-phosphate-terminated hairpins. The transcriptions were desalted with Sephadex G10 spin columns, purified using PAGE, and worked up as described above to generate internally ^{32}P-labeled polymerase substrates.

Adding ^{32}P-phosphate to an RNA primer.

An RNA primer 10 nucleotides in length was obtained from Dharmacon (**Table S1**). A ^{32}P radiolabel was added to the 5′-end of the sequence by adding the RNA primer at a concentration of 5 μM to a solution containing γ-^{32}P-ATP, 1x polynucleotide kinase buffer, and 5 units polynucleotide kinase. The reaction was incubated at 37 ℃ for 2 hours, purified using PAGE, and worked up as described above.

Kinetic analysis of pyrophosphate inhibition of RNA primer extension.

Kinetic data for pyrophosphate inhibition were obtained by monitoring the extension of a ^{32}P-labeled primer by tC19Z over time and varying concentrations of pyrophosphate. Equimolar concentrations (0.5 μM) of tC19Z, primer, and template were mixed and heated to 82 ºC for 2 minutes, then cooled slowly to 17 ºC for 10 minutes. Varying concentrations of pyrophosphate were then added to the reactions followed by buffer containing 200 mM $MgCl_2$, 50 mM Tris HCl buffer (pH 8.3), and 4 mM rNTPs. The solution was then allowed to incubate at 17 ºC, 5 μL aliquots were withdrawn at various time points, and the reaction was quenched with 250 mM EDTA and 7

M urea gel-loading buffer. The reactions were analyzed using an 18% denaturing PAGE to obtain single-nucleotide-resolution images. The gels were imaged on a Typhoon 9410 Variable Mode Imager, and the pixel density of the bands was analyzed using ImageJ to generate the kinetic data. Initial rates were calculated using the first four time points and used to demonstrate pyrophosphate inhibition.

Repair of cyclic phosphate damaged RNA hairpins.

Hairpin substrate repair was visualized and quantified by imaging the extension of an internally ^{32}P-labeled hairpin by tC19Z. For negative controls, equimolar concentrations (0.5 μM) of tC19Z and 2′-3′-cyclic-phosphate-terminated hairpin were mixed and heated to 82 ºC for 2 minutes, and then cooled slowly to 17 ºC for 10 minutes. At this point, one of two buffers was added (1st buffer: 200 mM $MgCl_2$ and 50 mM Tris buffer; 2nd buffer: 200 mM $MgCl_2$, 50 mM Tris buffer, and 4 mM rNTPs), and the reactions were incubated at 17 ºC for 24 hours. The reactions were quenched with 250 mM EDTA and 7 M urea gel-loading buffer, and then were analyzed on a 15% PAGE gel. For the repair reaction, equimolar concentrations (0.5 μM) of tC19Z and cyclic-phosphate-terminated hairpin were mixed and heated to 82 ºC for 2 minutes, followed by slow cooling to 17 ºC for 10 minutes and subsequent addition of 10 mM pyrophosphate. A buffer containing 200 mM $MgCl_2$, 50 mM Tris buffer, and 4 mM rNTPs was added, and the reaction was allowed to incubate at 17 ºC for 24 hours. The reactions were quenched with 250 mM EDTA and 7 M urea gel-loading buffer and were analyzed on a 15% denaturing polyacrylamide gel. To generate a positive control sample, the cyclic phosphate was removed from the hairpin by treating with polynucleotide kinase in 1x PNK buffer for 1 hour at 37 ºC then inactivated at 65 ºC for 10 minutes. The PNK-repaired hairpin substrate was then added to the tC19Z polymerase ribozyme, both at concentrations of 0.5 μM, and the solution was heated to 82 ºC for 2 minutes and then cooled slowly to 17 ºC for 10 minutes. Next, one of two buffers was added (1st buffer: 200 mM $MgCl_2$ and 50 mM Tris buffer; 2nd buffer: 200 mM $MgCl_2$, 50 mM Tris buffer, and 4 mM rNTPs), and the reactions were allowed to incubate at 17 ºC for 24 hours. The reactions were quenched

with 250 mM EDTA and 7 M urea gel-loading buffer and were analyzed on a 15% denaturing polyacrylamide gel. The gels were imaged on a Typhoon 9410 Variable Mode Imager.

Kinetic analysis of cyclic-phosphate damage repair.

The kinetics of tC19Z-mediated extension of pyrophosphate-repaired hairpin, PNK-repaired hairpin, and pyrophosphate inhibited PNK-repaired hairpin were obtained under similar conditions as those described above. Either hairpin or PNK-repaired hairpin was incubated with tC19Z at 82 ºC for 2 minutes and then cooled slowly to 17 ºC for 10 minutes. For the reactions containing pyrophosphate-repaired hairpin and pyrophosphate-inhibited hairpin, pyrophosphate was added at a concentration of 10 mM. A buffer containing 200 mM $MgCl_2$, 50 mM Tris buffer, and 4 mM rNTPs was added to each reaction, and the solutions were allowed to incubate at 17 ºC. Time points were removed from the reactions in 5 µL aliquots and quenched with 250 mM EDTA and 7 M urea gel-loading buffer. The reactions were analyzed on a 15% denaturing polyacrylamide gel that was imaged on a Typhoon 9410 Variable Mode Imager. ImageJ was used to quantify band density and generate the kinetic data for tC19Z extension of hairpins under varying pyrophosphate and repair conditions.

Hairpin migration and extension under various conditions.

To better understand hairpin migration and reactivity, an A/U hairpin was either left untreated, or was allowed to react with an alkaline phosphatase [5 µM hairpin, 1x Cutsmart buffer (NEB), and shrimp alkaline phosphatase] for 30 mins at 37 ºC. The two conditions were then either left untreated, treated with PNK as described above, or incubated with acid (25 mM HCl) followed by quenching. Each solution was then analyzed on a 15% denaturing polyacrylamide gel. Furthermore, each treated hairpin (0.5 µM) was heated with tC19Z (0.5 µM) for 2 minutes at 85 ºC and then cooled to 17 ºC over 10 minutes. A buffer containing 200 mM $MgCl_2$ and 50 mM Tris HCl (and, in some reactions, 4 mM rNTPs) was added. Additionally, 4 mM uridine and/or 10 mM pyrophosphate were added to some reactions. The mixtures were incubated at 17 ºC for 24 hours followed by quenching with 250 mM EDTA and 7 M urea gel-loading buffer. The reactions

were analyzed on a 15% denaturing polyacrylamide gel that was imaged on a Typhoon 9410 Variable Mode Imager.

Incorporation of ^{32}P-rNMPs into non-labeled, repaired hairpins.

To further demonstrate repair, an A/U hairpin was transcribed with an α-^{32}P[ATP] or without (cold). Some of the hairpins were then treated with PNK as described above. The ^{32}P-labeled hairpin (0.5 μM) was incubated with tC19Z (0.5 μM) at 82 ºC for 2 mins and cooled to 17 ºC over 10 minutes. A buffer containing 200 mM $MgCl_2$ and 50 mM Tris HCl (and, in some reactions, 4 mM rNTPs) was added. To some reactions was added either 10 mM pyrophosphate or α-^{32}P[ATP] and 10 mM pyrophosphate. The cold hairpin was treated in the same way. The mixtures were incubated at 17 ºC for 24 hours followed by quenching with 250 mM EDTA and 7 M urea gel-loading buffer. The reactions were analyzed on a 15% denaturing polyacrylamide gel that was imaged on a Typhoon 9410 Variable Mode Imager.

A U/A hairpin substrate was transcribed without α-^{32}P-ATP and either left untreated or allowed to react with PNK as described above. The hairpin (0.5 μM) was heated with tC19Z (0.5 μM) for 2 minutes at 85 ºC and then cooled to 17 ºC over 10 minutes. A buffer containing 200 mM $MgCl_2$, 50 mM Tris HCl and, in some reactions, 10 mM pyrophosphate was added. Additionally, either α-^{32}P[ATP], α-^{32}P[UTP], or α-^{32}P[CTP] was added. The reactions were incubated at 17 ºC for 24 hours followed by quenching with 250 mM EDTA and 7 M urea gel-loading buffer. The reactions were analyzed on a 15% denaturing polyacrylamide gel that was imaged on a Typhoon 9410 Variable Mode Imager.

Extension of mismatched RNA hairpin sequences.

Mismatched RNA hairpins were extended by the tC19Z polymerase ribozyme under conditions of PNK repair or pyrophosphate. Pyrophosphate facilitated extension of both a cyclic-phosphate-damaged hairpin and a mismatched hairpin when 0.5 μM hairpin was added to 0.5 μM tC19Z. The solution was incubated at 82 ºC for 2 minutes and then cooled slowly to 17 ºC for 10

minutes. Pyrophosphate was added at a final concentration of 10 mM, followed by a buffer containing 200 mM $MgCl_2$, 50 mM Tris•Cl, and 4 mM rNTPs, and the solution was incubated at 17 ºC for 24 hours. A reaction containing a PNK-repaired mismatched hairpin and tC19Z at concentrations of 0.5 μM each was incubated at 82 ºC for 2 minutes and then cooled slowly to 17 ºC for 10 minutes. Pyrophosphate was added at a concentration of 10 mM, followed by a buffer containing 200 mM $MgCl_2$, 50 mM Tris buffer, and 4 mM rNTPs, and the reaction was allowed to proceed at 17 ºC for 24 hours. All reactions were quenched with 250 mM EDTA and 7 M urea gel-loading buffer and then analyzed on a 15% denaturing polyacrylamide gel that was imaged on a Typhoon 9410 Variable Mode Imager. ImageJ was used to quantify band density to generate fraction extended graphs.

High-throughput sequencing of RNA hairpins.

RNA hairpins were analyzed by Illumina Sequencing to determine sequence identity. Cyclic phosphate repair and mismatch extension reactions, as described above, were performed and fractionated by PAGE. The images obtained were analyzed for extension products, and the corresponding bands were excised from the gel, eluted, and precipitated with ethanol. The RNA sequences were resuspended in water and then ligated to a DNA sequence containing an Illumina forward adaptor (**Table S1**). The DNA adaptor sequence was designed with a 5′-phosphate, followed by four random nucleotides to reduce ligation bias at the 3′-RNA terminus. The adaptor also contained a 3′-amine to prevent self-ligation. The oligo (10 μM) was added to a mixture containing PEG 8000, 1x T4 RNA Ligase buffer, 15 μM ATP, and T4 RNA Ligase 1 to a total volume of 100 μL in order to adenylylate the 5′-end of the DNA sequence (5′-App oligo). G15 Sephadex was used to desalt and remove residual ATP from the 5′-App oligo.

To anneal the sequences, the 5′-App oligo was added in approximately three times excess of the RNA hairpin, and the resulting solution was heated to 95 ºC for 3 mins and then cooled to 25 ºC at a rate of 1 ºC/min. To ligate, the annealed oligos were added to 1x T4 RNA ligase buffer,

PEG 8000, and T4 RNA Ligase 2, truncated KQ to a total volume of 20 μL, and allowed to react at 25 ºC for 18 hours and then at 4 ºC for 2 hours.

To a 5 μL aliquot of the ligation reaction was added 2.5 μM of the RT primer (**Table S1**), which was annealed at 75 ºC for 3 mins, 37 ºC for 10 mins, and 25 ºC for 10 mins. Following annealing, 1x Protoscript II buffer, 0.5 mM each dNTP, 5 mM $MgCl_2$, 10 mM DTT, and Protoscript II was added and the reverse transcription reaction was heated to 50 ºC for 2 hours to generate cDNA.

The Illumina reverse adaptor was added to the cDNA through a PCR reaction (**Table S1**). Forward and reverse primers (1 μM) were added to the cDNA, as well as 1x Q5 reaction buffer, 5 mM of each dNTP, and Q5 DNA polymerase. The reaction was annealed at 52 ºC and extended at 72 ºC. The Illumina barcodes were added in a second PCR reaction containing 1 μM of each primer, 1x Q5 buffer, 5 mM each dNTP, and Q5 DNA polymerase. The reaction to install the Illumina barcodes was annealed at 65 ºC and extended at 72 ºC, and the reaction was purified via agarose gel. The libraries were submitted for Illumina Sequencing, and the resulting reads were counted and primer regions were trimmed using FastAptamer and CutAdapt, respectively. Starting and extended sequences were isolated from data for analysis using FastAptamer.[53]

Measuring ATP concentrations using a firefly luciferase assay.

A solution containing a PNK-repaired A/C hairpin and tC19Z polymerase ribozyme at concentrations of 0.5 μM each was incubated at 82 ºC for 2 minutes and then cooled slowly to 17 ºC for 10 minutes. Pyrophosphate was added at a concentration of 10 mM followed by a buffer containing 200 mM $MgCl_2$, and 50 mM Tris buffer, and the reaction was allowed to proceed at 17 ºC for 24 hours. Control reactions were incubated as above, but in the absence of pyrophosphate. After 24 hours, the reaction containing pyrophosphate was centrifuged, and the supernatant was collected. Inorganic pyrophosphatase (NEB) was added to both reactions and allowed to react at 37 ºC for 1 hour. Luminescence was generated using the StayBrite Highly Stable ATP

Bioluminescence Assay Kit (BioVision), as described previously.[44] Briefly, to a reaction tube was added 1 μL of either a control or a pyrophosphate reaction, 1 μL of 50 μM D-luciferin, 1 μL of 4x StayBrite reaction buffer, and 1 μL StayBrite luciferase enzyme. A Frame-Seal chamber (Bio-Rad) was placed between two glass slides, and then to the glass slide were added 1 μL of an enzyme-only control, 1 μL no-pyrophosphate control, and 1 μL pyrophosphorolysis reaction. The Frame-Seal chamber was then mounted to a mirror in front of a macro lens constructed by mounting two Nikon f/1.2 lenses in opposition. The macro lens system images an area of approximately the same size of the CCD chip on the camera ensuring that most of the luminescence output is directed to the chip of the EMCCD camera (iXon Ultra 888, Andor). The temperature of the camera chip was maintained at –100 ºC during detection using a thermoelectric recirculating chiller (Thermocube 300, Solid State Cooling Systems). During imaging, 10 one-minute exposures were collected. The images were then stacked, a Z-projection was generated, and relative luminescence was measured using Fiji.

References

1. Orgel, L. E. Prebiotic Chemistry and the Origin of the RNA World. *Crit. Rev. Biochem. Mol. Biol.* **39**, 99–123 (2004).

2. Robertson, M. P. & Joyce, G. F. The origins of the RNA world. *Cold Spring Harb. Perspect. Biol.* **4**, a003608 (2012).

3. Schrum, J. P., Zhu, T. F. & Szostak, J. W. The Origins of Cellular Life. *Cold Spring Harb. Perspect. Biol.* **2**, (2010).

4. Bartel, D. P. & Unrau, P. J. Constructing an RNA world. *Trends Cell Biol.* **9**, 9–13 (1999).

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

6. Ma, W. & Yu, C. Intramolecular RNA replicase: Possibly the first self-replicating molecule in the RNA world. *Orig. Life Evol. Biosph.* **36**, 413–420 (2006).

7. Chen, X., Li, N. & Ellington, A. D. Ribozyme Catalysis of Metabolism in the RNA World. *Chem. Biodivers.* **4**, 633–655 (2007).

8. Benner, S. A., Ellington, A. D. & Tauer, A. Modern metabolism as a palimpsest of the RNA world. *Proc. Natl. Acad. Sci. U. S. A.* **86**, 7054–7058 (1989).

9. Neveu, M., Kim, H. J. & Benner, S. A. The “Strong” RNA World Hypothesis: Fifty Years Old. *Astrobiology* **13**, 391–403 (2013).

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

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

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

13. Jaeger, L., Wright, M. C. & Joyce, G. F. A complex ligase ribozyme evolved in vitro from a group I ribozyme domain. *Proc. Natl. Acad. Sci.* **96**, 14712–14717 (1999).

14. Shechner, D. M. *et al.* Crystal Structure of the Catalytic Core of an RNA-Polymerase Ribozyme. *Science* **326**, 1271–1275 (2009).

15. Ekland, E. H. & Bartel, D. P. RNA-catalysed RNA polymerization using nucleoside triphosphates. *Nature* **382**, 373–376 (1996).

16. Johnston, W. K., Unrau, P. J., Lawrence, M. S., Glasner, M. E. & Bartel, D. P. RNA-Catalyzed RNA Polymerization: Accurate and General RNA-Templated Primer Extension. *Science* **292**, 1319–1325 (2001).

17. Samanta, B. & Joyce, G. F. A reverse transcriptase ribozyme. *Elife* **6**, e31153 (2017).

18. Wochner, A., Attwater, J., Coulson, A. & Holliger, P. Ribozyme-Catalyzed Transcription of an Active Ribozyme. *Science* **332**, 209–212 (2011).

19. Horning, D. P. & Joyce, G. F. Amplification of RNA by an RNA polymerase ribozyme. *Proc. Natl. Acad. Sci.* **113**, 9786–9791 (2016).

20. Horning, D. P., Bala, S., Chaput, J. C. & Joyce, G. F. RNA-Catalyzed Polymerization of Deoxyribose, Threose, and Arabinose Nucleic Acids. *ACS Synth. Biol.* **8**, 955–961 (2019).

21. Cojocaru, R. & Unrau, P. J. Processive RNA polymerization and promoter recognition in an RNA World. *Science* **371**, 1225–1232 (2021).

22. Tjhung, K. F., Shokhirev, M. N., Horning, D. P. & Joyce, G. F. An RNA polymerase ribozyme that synthesizes its own ancestor. *Proc. Natl. Acad. Sci.* **117**, 2906–2913 (2020).

23. Zaher, H. S. & Unrau, P. J. Selection of an improved RNA polymerase ribozyme with superior extension and fidelity. *RNA* **13**, 1017–1026 (2007).

24. Attwater, J. *et al.* Chemical fidelity of an RNA polymerase ribozyme. *Chem. Sci.* **4**, 2804–

2814 (2013).

25. Beate, S., Rana, S., Kiong, H. C. & Stewart, S. Portability and fidelity of RNA-repair systems. *Proc. Natl. Acad. Sci.* **101**, 2788–2793 (2004).

26. Burroughs, A. M. & Aravind, L. RNA damage in biological conflicts and the diversity of responding RNA repair systems. *Nucleic Acids Res.* **44**, 8525–8555 (2016).

27. Nandakumar, J., Schwer, B., Schaffrath, R. & Shuman, S. RNA Repair: An Antidote to Cytotoxic Eukaryal RNA Damage. *Mol. Cell* **31**, 278–286 (2008).

28. Kamerlin, S. C. L., Sharma, P. K., Prasad, R. B. & Warshel, A. Why nature really chose phosphate. *Q. Rev. Biophys.* **46**, 1–132 (2013).

29. Pasek, M. A. Rethinking early Earth phosphorus geochemistry. *Proc. Natl. Acad. Sci.* **105**, 853–858 (2008).

30. Miller, S. L. & Parris, M. Synthesis of pyrophosphate under primitive earth conditions. *Nature* **204**, 1248–1250 (1964).

31. Hørder, M. Complex formation of inorganic pyrophosphate with magnesium: The influence of ionic strength, supporting medium and temperature. *Biochim. Biophys. Acta - Enzymol.* **358**, 319–328 (1974).

32. Holm, N. G. & Baltscheffsky, H. Links Between Hydrothermal Environments, Pyrophosphate, Na+, and Early Evolution. *Orig. Life Evol. Biosph.* **41**, 483–493 (2011).

33. Jimenez, R. M., Polanco, J. A. & Lupták, A. Chemistry and Biology of Self-Cleaving Ribozymes. *Trends Biochem. Sci.* **40**, 648–661 (2015).

34. Li, Y. & Breaker, R. R. Kinetics of RNA Degradation by Specific Base Catalysis of

Transesterification Involving the 2‘-Hydroxyl Group. *J. Am. Chem. Soc.* **121**, 5364–5372 (1999).

35. Seith, D. D., Bingaman, J. L., Veenis, A. J., Button, A. C. & Bevilacqua, P. C. Elucidation of Catalytic Strategies of Small Nucleolytic Ribozymes from Comparative Analysis of Active Sites. *ACS Catal.* **8**, 314–327 (2018).

36. Riccitelli, N. J., Delwart, E. & Lupták, A. Identification of Minimal HDV-Like Ribozymes with Unique Divalent Metal Ion Dependence in the Human Microbiome. *Biochemistry* **53**, 1616–1626 (2014).

37. Wang, L. K., Lima, C. D. & Shuman, S. Structure and mechanism of T4 polynucleotide kinase: an RNA repair enzyme. *EMBO J.* **21**, 3873–3880 (2002).

38. Das, U. & Shuman, S. Mechanism of RNA 2’,3’-cyclic phosphate end healing by T4 polynucleotide kinase-phosphatase. *Nucleic Acids Res.* **41**, 355–365 (2013).

39. Winkler, W. C., Nahvi, A., Roth, A., Collins, J. A. & Breaker, R. R. Control of gene expression by a natural metabolite-responsive ribozyme. *Nature* **428**, 281–286 (2004).

40. Roth, A. *et al.* A widespread self-cleaving ribozyme class is revealed by bioinformatics. *Nat. Chem. Biol.* **10**, 56–60 (2014).

41. Pinto, P. H. *et al.* ANGEL2 is a member of the CCR4 family of deadenylases with 2′,3′-cyclic phosphatase activity. *Science* **369**, 524–530 (2020).

42. Li, G. M. Mechanisms and functions of DNA mismatch repair. *Cell Res.* **18**, 85–98 (2008).

43. Iyer, R. R., Pluciennik, A., Burdett, V. & Modrich, P. L. DNA mismatch repair: functions and mechanisms. *Chem. Rev.* **106**, 302–323 (2006).

44. Chizzolini, F., Kent, A. D., Passalacqua, L. F. M. & Lupták, A. Enzymatic RNA Production from NTPs Synthesized from Nucleosides and Trimetaphosphate**. *ChemBioChem* **22**, 2098–2101 (2021).

45. Glasner, M. E., Yen, C. C., Ekland, E. H. & Bartel, D. P. Recognition of Nucleoside Triphosphates during RNA-Catalyzed Primer Extension. *Biochemistry* **39**, 15556–15562 (2000).

46. Cruse, W. B. *et al.* Structure of a mispaired RNA double helix at 1.6-A resolution and implications for the prediction of RNA secondary structure. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 4160–4164 (1994).

47. Pan, B., Mitra, S. N. & Sundaralingam, M. Crystal Structure of an RNA 16-mer Duplex R(GCAGAGUUAAAUCUGC)2 with Nonadjacent G(Syn)·A+(Anti) Mispairs. *Biochemistry* **38**, 2826–2831 (1999).

48. Sugimoto, N., Kierzek, R. & Turner, D. H. Sequence dependence for the energetics of terminal mismatches in ribooligonucleotides. *Biochemistry* **26**, 4559–4562 (1987).

49. Sydow, J. F. *et al.* Structural basis of transcription: mismatch-specific fidelity mechanisms and paused RNA polymerase II with frayed RNA. *Mol. Cell* **34**, 710–721 (2009).

50. Kahn, J. D. & Hearst, J. E. Reversibility of nucleotide incorporation by Escherichia coli RNA polymerase, and its effect on fidelity. *J. Mol. Biol.* **205**, 291–314 (1989).

51. Akoopie, A., Arriola, J. T., Douglas, M. & Muller, U. F. A GTP-synthesizing ribozyme selected by metabolic coupling to an RNA polymerase ribozyme. *Sci. Adv.* **7**, 7487 (2022).

52. Mohamady, S. & Taylor, S. D. Synthesis of Nucleoside Triphosphates from 2′-3′-

Protected Nucleosides Using Trimetaphosphate. *Org. Lett.* **18**, 580–583 (2016).

53. Alam, K. K., Chang, J. L. & Burke, D. H. FASTAptamer: A Bioinformatic Toolkit for High-throughput Sequence Analysis of Combinatorial Selections. *Mol. Ther. Nucleic Acids* **4**, e230 (2015).

Acknowledgements

We thank the staff at the Genomics High-Throughput Facility (GHTF) Shared Resource of the Cancer Center, Support Grant (P30CA-062203) at the University of California, Irvine and NIH shared instrumentation grants 1S10RR025496-01, 1S10OD010794-01, and 1S10OD021718-01; K. H. Cole for helpful discussion and comments on the manuscript; and S. Flynn for assistance with sequencing data. This work was supported by the John Templeton Foundation through a grant to the Foundation for Applied Molecular Evolution (to A. L.), NSF CBET Biophotonics 1804220 (to A. L.), NASA Exobiology 80NSSC21K0488 (to A. L.), NASA ICAR 80NSSC21K0596 (to A.L.), NSF GRFP (to A. K.), Goldwater Scholarship (to L. Y.), and UCI Undergraduate Research Opportunities Program awards (to L. Y. and I. L.).

Affiliations

Department of Chemistry, University of California, Irvine, Irvine, CA, USA
Alexandra D. Kent, Lucy L. Yang, and Andrej Lupták

Department of Pharmaceutical Sciences, University of California, Irvine, Irvine, CA, USA
Ivee A. Lee and Andrej Lupták

Department of Molecular Biology and Biochemistry, University of California, Irvine, Irvine, CA, USA
Andrej Lupták

Author contributions

A.D.K. and A.L. conceived and designed the study. A.D.K. performed the experiments with assistance from L.L.Y and I.A.L. The data were analyzed by A.D.K. The manuscript was prepared by A.D.K and A.L.