RESEARCH ARTICLE

High-Fidelity Templated Ligation of RNA via 2′,3′-cyclic Phosphate

Adriana Calaça Serrão*[a], Sreekar Wunnava*[a], Avinash V. Dass[a, b], Lennard Ufer[a],
Philipp Schwintek[a], Christof B. Mast[a] and Dieter Braun[a]

[a]   Department of Physics, Center for Nanoscience, Ludwig-Maximilians-Universität München, Amalienstrasse 54, 80799 Munich, Germany
        E-mail: dieter.braun@lmu.de
[b]   Department of Physics and Astronomy, McMaster University, 1280 Main Street West, Hamilton, Ontario, L8S4M1, Canada

        * Equal contribution

        Supporting information for this article is given via a link at the end of the document.

Abstract: The templated ligation of oligonucleotides offers a mode of replication in an RNA world. The 2′,3′-cyclic phosphate (>P) is a prebiotically available activation group for RNA and the product of backbone hydrolysis. Using gel electrophoresis and liquid chromatography, we found that the templated ligation of RNA with >P activation proceeds in alkaline (pH 9-11) low-salt aqueous solutions with 1 mM $MgCl_2$ in temperatures ranging from -20 to 25 °C within a few days. Under the optimum conditions of pH 10 and 5 °C, the ligation yielded 40% after 7 days. No additional catalysts were required. In contrast to previous reports, we found an equimolar mixture of 2′-5′ and 3′-5′ linked oligomers in the used conditions. We probed the nucleotide specificity at the ligation site and found that one mutation reduced the ligation yield by 82-92%. We extrapolated these results to a per-nucleotide replication fidelity of 95-98% when ligating 4- to 6-mers. With splinted oligomers, five ligations created a 96-mer strand, demonstrating a possible assembly pathway for long ribozymes. With the low salt requirements, strand separation will be compatible with the ligation conditions using non-equilibrium settings. The findings suggest that templated ligation mediated by 2′,3′-cyclic phosphate in alkaline conditions offer a slow, but precise replication and elongation reaction for RNA on early Earth.

Introduction

The RNA-world hypothesis suggests that RNA is the key polymer to perform early molecular evolution during the origins of life, serving as both an informational and catalytic polymer. As a central part of molecular evolution, the replication of RNA can therefore be assumed a process of pivotal importance for the emergence of life. To explore this challenge, prebiotic condensation of mononucleotides has been extensively studied with several chemical activations, such as phosphorimidazolides, carbodiimides, or cyclic nucleotides[1–7] Such condensation reactions provide a plausible mechanism for the formation of the first, short RNA fragments. The replication of these emergent strands, however, requires them to be included in a templated copying mechanism.

Primer-extending base-by-base replication has been shown to be feasible for such processes[8–11], although its processivity and the difficulty to realize strand separation under the necessary Mg^{2+} concentrations limit the approach. This method of elongation would, therefore, be inefficient in generating long functional RNA[12]. In contrast, the templated ligation of short oligonucleotides reduces the chemical steps required to generate equally long RNA molecules, effectively speeding up information copying. Thus, templated ligation could have taken over the base-by-base replication processes once an initial pool of short oligonucleotides was generated. Over the years, non-enzymatic templated ligation has been shown for RNA, DNA and XNA[13–15] via chemical activations like imidazolides and its analogues[16–20],

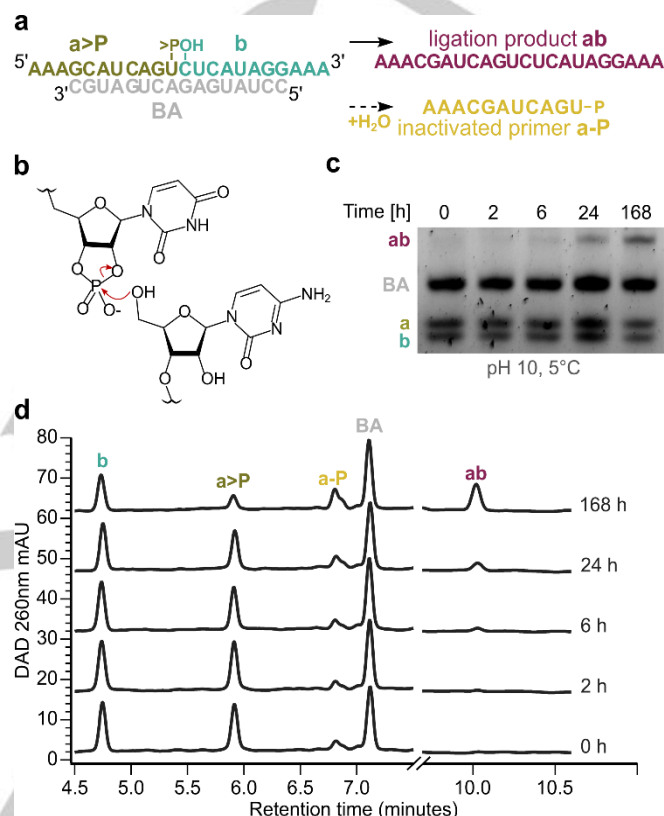**Figure 1: Non-enzymatic template directed ligation of short RNA strands. a**, Schematics of reaction design. Both primers (*a* and *b*) bind on the complementary template, *BA*. The primer *a* has a 2′,3′-cyclic phosphate, while *b* contains a 5′-OH group. **b,** 5′-OH performs a nucleophilic attack on the cyclic phosphate group and forms a phosphodiester bond between the two primers, leading to the ligation product strand *ab*. As a side reaction, the cyclic phosphate in *a* can also hydrolyse, rendering *a* inactive. **c**, Denaturing PAGE analysis of ligation reaction over time. Reaction contained 1 µM primers, 1 µM template, 50 mM CHES, pH 10 and 1 mM $MgCl_2$, at 5 °C. **d**, Stacked HPLC chromatograms (absorption at 260 nm) of the same reaction mixtures as in **c**. The product peak increases over time, as the primers get deplete

carbodiimides[21–23], pyrophosphate[19,20,24]. Without these chemical activations, the formation of 40-200nt long ligating ribozymes seems impossible[25–27]. However, many of these reactions require multiple synthesis steps to generate the required activated molecules such as phosphorimidazolides.

On the other hand, activated phosphates, particularly cyclic phosphates, present a simple and endogenous activated group that reduces the need for complex multi-step reactions and the use of additional activating agents. RNA oligonucleotides with 2′,3′-cyclic phosphate (>P) ends are readily available in the prebiotic pool through the polymerization of 2',3'-cyclic nucleotides [5,7,28] or the hydrolysis of longer RNA strands [29–31]. Overall, employing >P

for ligation provides a convenient route for recycling without the need for complex processes involving the attachment and removal of leaving groups. Instead, it involves the straightforward circularization of a >P derived from a 2'- or 3'-phosphate [27,32–34]. Biochemical protocols involving EDC[35] or DAP+imidazole[27] have been established, which gives us reason to be optimistic about discovering prebiotically feasible alternatives for the ring closure reaction.

Vlassov and co-workers explored the potential of RNA with >P as a prebiotic activation group for ligation reactions [36–42]. They demonstrated that a DNA-RNA chimera ligates at mildly alkaline pH and 37°C, reaching maximum yield of 16% at pH 8.8. Further investigations by Lutay et al. [41,43] showed that an RNA based system with single strand overhangs leads to concomitant cleavage and ligation with a maximum combined cleavage-ligation yield of 6% of at 25°C pH 9.5, with predominantly 2′–5′ linkages. The authors report that this reaction proceeds through the formation of >P which can either be attacked by the 5′-OH of the neighboring oligonucleotide (Figure 1 **b**) or reacts with a water molecule resulting in 2′ or 3′ phosphate, which is inactive.

In this study, we build on these previous works by determining the kinetics of non-enzymatic ligation using short RNA with >P (Figure 1 **a**), and their dependence on pH and temperature. We did not only simplify, but significantly optimize the reaction and report a maximal yield of ca. 40% at 5 °C and pH 10. Moreover, in contrast to previous studies in aqueous conditions[43,44], we obtained a 50% ratio of canonical 3′–5′ linkages at the ligation site. We also explore the influence of sequence variation at the ligation site and found that the reaction is very sequence specific for corresponding bases, especially for ligating 4 to 6 mers. Finally, we could demonstrate that multistep ligations within a splinted RNA system using >P can generate long RNA molecules on the length scale of 100 nucleotides. This work serves as a proof-of-principle for non-enzymatically generating long RNA, in a sequence-specific manner using a most simple and ubiquitous phosphate chemistry on early Earth.

Results and Discussion

Learning from our previous work[28] and other studies on ligation with >P[39–41,43] we decided to investigate the ligation reaction with >P under alkaline conditions. As such conditions allow for RNA oligomerization in a dry state [28], we demonstrate here that similar conditions are compatible with ligation in the wet state suggesting a shared emergence. This was also corroborated by a broader pH screen (Supporting Information, SI 4 Figure S4.1). All the reactions were done in aqueous alkaline conditions with 50 mM CHES buffer with equimolar concentration of the participating strands. A possible influence of the buffer itself was not supported in a buffer screen at pH 9 (SI 4 Figure S4.2). The two ligating strands are labelled as primer *a* and *b* and the template, *BA*. All the reactions were performed in the presence of 1 mM Mg^{2+} as it was found to be more efficient (SI 4 Figure

S4.3). The analysis of the ligation reaction was done by either polyacrylamide gel electrophoresis (PAGE, Figure 1 **c**, see SI 1.3 and 2.2) or by high performance liquid chromatography (HPLC, Figure 1 **d**, SI 1.4 and 2.1), or a combination of both.
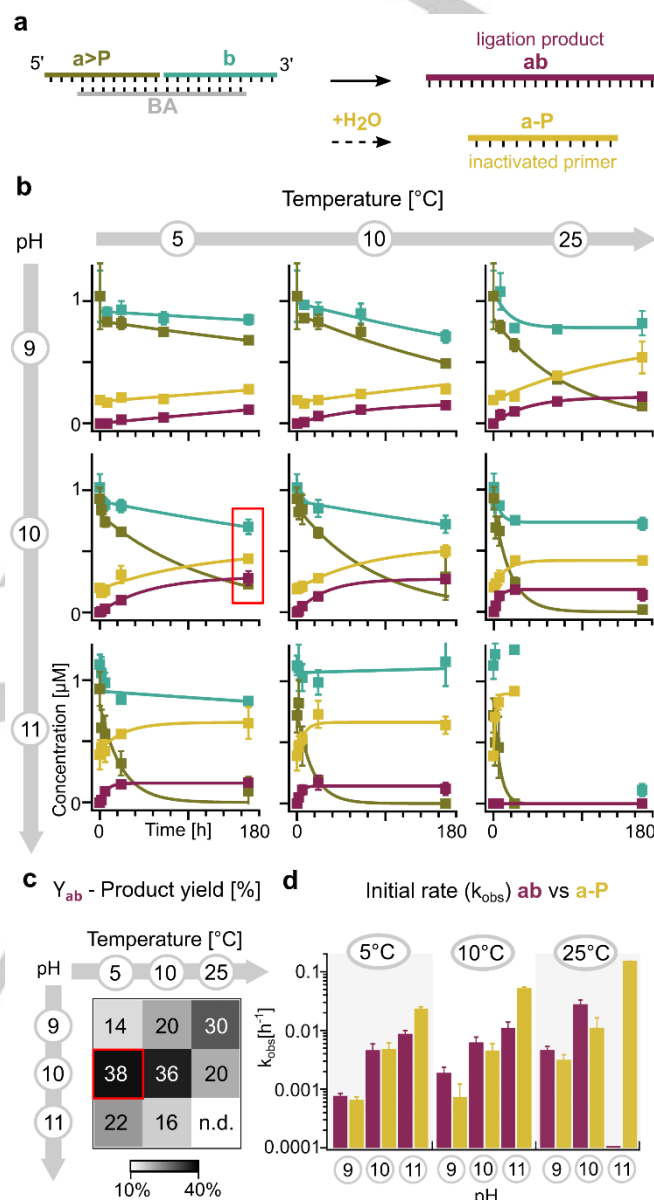

Figure 2: Kinetic study of temperature and pH influence in ligation reaction. a, Schematics of the ligation reaction. The hydrolysis of the cyclic phosphate competes with the ligation reaction. **b,** Screening the concentration of primers and product over a 7 days' period for varied pH (9-11) and temperature (5, 10, 25°C). Highest yield at 7 days are highlighted in red, for pH 10 and 5°C. At both high temperature and pH (25°C, pH 11) there is additional hydrolysis of the backbone, particularly after 7 days (bottom, right-most graph). The full lines correspond to the exponential fit of the data as a guide to the eye. **c**, Product yield obtained (%) at 7 days for all the conditions tested. Maximum obtained yield was for pH 10 at 5°C (red square). This reported yield has been corrected for the limiting concentration of a>P. **d**, Observed initial pseudo-first order rate of product and inactive primer formation for all the conditions (SI 7). Reactions contained 1 µM primers, 1 µM template, 50 mM CHES, with varied pH and 1 mM MgCl$_2$. Concentrations were measured with HPLC UV detection at 260 nm. Data are represented as mean ± standard deviation of three independent replicates.

To study the impact of temperature and pH on the ligation reaction, three temperatures 5-25 °C and pH 9-11 were tested in Figure 2. The reaction yielded both the ligation product (*ab*)

and the inactivated primer side product (*a-P*) while consuming the two primers *a>P* and *b*. The template (*BA*) was not consumed by the reaction. Thus, its concentration was used to correct for small pipetting errors through normalization.

The concentration over time of the strands *a>P, b, ab, a-P*, for different temperature-pH combinations are plotted in Figure 2 **b**, (see Figure 2 **a** for color-coded schematics). It is important to note that in all the experiments, the initial concentration of *a-P* is on average 26% of total *a,* suggesting that a part of *a>P* was already hydrolyzed in the stock solutions. This capped the maximum possible yield at 0.74 μM, the concentration of the initial *a>P*. The yield of *ab* depends on the temperature and pH combination of the reaction. The maximum yield at 7 days, of 28% was obtained at 5 °C pH 10 (Figure 2 **b** and **c**, red rectangle). Correcting for the available amount of *a>P,* the yield was 38%.

Since the formation of the primer-primer-template complex through base-pairing proceeds at much faster timescales ($k_{on} \approx 1$ μM^{-1}s^{-1}, literature values for similar sized oligonucleotides[45–47]) than the nucleophilic attack on the cyclic phosphate, it can be assumed that the reaction is of pseudo-first order. Thus, first order kinetic rate constants (k_{obs}) for product formation and hydrolysis of cyclic phosphate were fitted to the data, see Figure 2 **d** (SI 7 Figure S7.1).
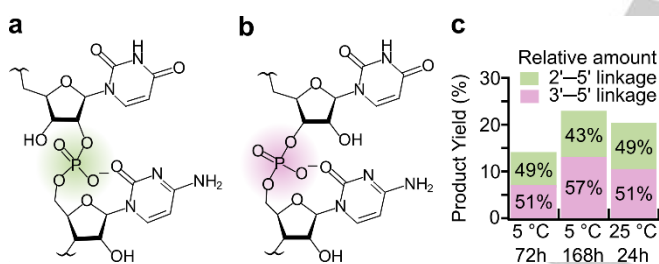

Figure 3: Linkage analysis of the reaction product *ab* through digestion with Nuclease P1. Scheme of the ligation site with a 2′–5′ linkage (**a**, green) and a 3′–5′ linkage (**b**, pink). **c,** Total product yield obtained for three condition sets, with the corresponding relative amount of 2′–5′ and 3′–5′ linkage. The ligation with 2′,3′-cyclic phosphates does not exhibit regioselectivity, as both linkages are equally represented for the studied conditions. After the reactions with 10 μM primers, 10 μM template, 50 mM CHES pH 10 and 1 mM MgCl₂, they were digested with Nuclease P1 (SI 1.5). The concentrations of the samples before and after digestion were measured with HPLC UV detection at 260 nm (SI 2.1)). Data represents the mean of independent duplicates.

A few salient features of the plots in Figure 2 **b, c, d** are the rates and yields of ligation (*ab*) and hydrolysis (*a-P*). Both the rates of ligation and hydrolysis increase with an increase in either pH or temperature, keeping the other parameter constant. Especially since the hydrolysis of the cyclic phosphate renders the primer *a>P* inactive, we scoped the reaction conditions where the ligation reaction is favored. Figure 2 **d** shows that while the observed ligation rate mostly increases from pH 9 to 10, the inactivation rate increases from pH 10 to 11. At higher pH, the ligation kinetics is slightly faster, however the final yield drops to the competing inactivation rate.

The rate of ligation is in general higher than the rate of the hydrolysis at 5 and 10 °C while 25 °C favor the inactivation.

The result is a broad window of ligation kinetics with a maximum 7-day yield obtained at pH 10 and 5°C. It has been reported that low temperatures reduce the entropic cost of the ligation reaction and shift the reaction equilibrium from hydrolysis to ligation[27]. One should note that only when both temperature 25 °C and pH 11 are maximal, significant RNA backbone hydrolysis is observed.

The attack of the 5′-OH on the cyclic phosphate (Figure 1 **b**) can form either a 2′–5′ or a 3′–5′ phosphodiester bond (Figure 3**a** and **b** respectively). Previous studies on the polymerization and ligation with cyclic phosphates have reported varying ratios of 3′–5′ to 2′–5′ linkages, depending largely on the experimental conditions. For example, dry state polymerization resulted in a natural linkage enrichment ratio of 2:1[5,48], while an aqueous state (with 0.5M diamine, pH 8 and 0°C) was reported to lead to at least 97% of 2′–5′ [44]. Templated cleavage and ligation at 25°C, pH 9 and 5 mM MgCl₂ in an aqueous solution was reported to also show a predominance of 2′–5′ linkages (about 95%)[43]. Conversely, templated ligation in the eutectic phase resulted in an excess of 3′–5′ linkages[36]. For templated ligation reaction described here, we found no regioselectivity under the tested conditions (Figure 3 **c**). To investigate this, the reaction was quenched by ethanol precipitation and the samples were digested with Nuclease P1 following the manufacturer's protocol (SI 1.5). Nuclease P1 specifically lyses the 3′–5′ linkages, which in this case would digest all the strands *a, b* and *BA* completely but digest *ab* either completely, or result in a UC dimer.

The concentration of total product pre-digestion and UC dimer post-digestion were determined using HPLC UV absorption (SI 2.1). We observed that both types of linkages were formed equally (Figure 3 **c**) indicating that the reaction is not regioselective. However, a slight enrichment of the canonical linkage over time for the 5 °C conditions can be seen, possibly due to the favored hydrolysis of 2′–5′ linkages, particularly in double-stranded RNA in alkaline solutions[20]. Nevertheless, such mixed backbone RNA have been demonstrated to still fold into functional structures[49]. Furthermore, SI 6.2 Figure S6.3 **a** shows that *ab* with 2′–5′ linkage can still template the formation of BA through >P mediated ligation, highlighting the possibility of a replication cycle.

To answer the question if the ligation reaction is sequence-specific, we conducted the reaction with each of the 16 different nucleotide combinations (four each on 3′ end of *a* and 5′ end of *b*) at the ligation site. Additionally, we tested two different templates, one with GA and the other with UA at the ligation site. Except for the ligation site, the remainder of the sequence was fully complementary to the template. These reactions were carried out at pH 10 and 5°C for 168 hours i.e. condition with the highest yield of ligation among the ones tested (Figure 2 **c**). Figure 4 **b** and **c** show that of all the combination of the primers tested, the highest yield of *ab* is obtained for the sequence with the correct nucleotides at the ligation site (marked in red, CU for the template GA and AU for the template UA). However, mismatched ligations did

occur, albeit with much lower relative yields, which were especially reduced for the template UA.

Interestingly, G:U wobble pairing of the 5′ U of primer *b* at the ligation site led to a high relative yield (78%) compared to the complementary primers for the template GA. The average relative yield of wrong ligation (including the G:U wobble) is 15% for template GA, while it amounts to 5% for the template UA. When considering one single mutation at the ligation site either on *a* or *b*, the reaction yield drops on average 91% or 82% respectively, relative to the non-mutated complex.
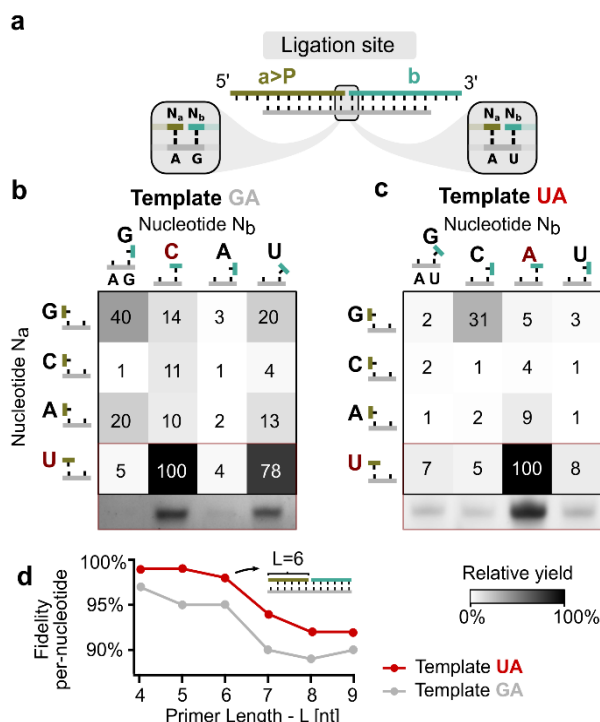

NUPACK with the experimental ligation error rates (Figure 4 **b** and **c**).

In combination, per nucleotide fidelities reached 89-92% for the tested system and 95-98% for primers up to a length of 6-mer each. This is especially interesting since these shorter primers are the typical lengths in dry oligomerization from 2′,3′-cyclic nucleotides [28]. Longer primers would not discriminate sufficiently enough against single point mutations for the given temperature of 5°C. For shorter strands, it is likely that one single mutation would destabilize the complex thus lowering the tolerance for mutations. Consequently, the base-by-base replicator must have increased fidelity for shorter lengths.
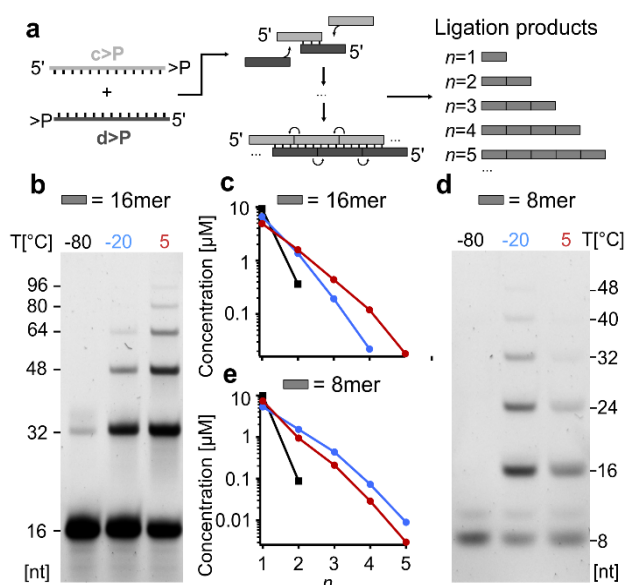

Figure 4: Ligation site specificity for two different templates sequences. Reaction yields at 7 days were quantified for reactions with primers containing each of the possible four nucleotides the 3′ end of *a* (N_a) and 5′ end of *b* (N_b) leading to 16 primer combinations. **a**, Schematics of the ligation sites and the two templates tested. The templates differed only at the dimer complementary to the ligation site, with either 5′ GA 3′ (**b**) or 5′ UA 3′ (**c**). The maximum yield obtained in both cases was for the correct combination of complementary primers (nucleotides highlighted in red). G:U wobble pairing is represented as a tilted nucleotide in the cartoon representation. For most combinations, one single mutation at the ligation site reduced its relative yield considerably or prevented ligation, even though the second primer is fully bound to the template. The snippet below the heat map in (**b** and **c**) corresponds to the PAGE of the bottom row, showing the ligated product *ab*. Reactions were performed with 10 μM primers, 10 μM template, 50 mM CHES pH 10 and 1 mM MgCl₂ for 7 days at 5 °C. Data are represented as mean of independent triplicates. **d,** The fidelity of ligation was extrapolated to a per-nucleotide replication fidelity using primers of varying length using single-mutation sensitive binding calculation of the primers with NUPACK (SI 8.1). The fidelity dropped for longer primers. Also, a G at the ligation site lead to lower fidelity due to the G-U wobble pairs (SI 8.3).

A prebiotic replication through ligation would take place from a diverse pool of oligonucleotides where the sequence outside the ligation site would also influence the replication yield. To gauge the performance ligation for replication, we calculated the per-nucleotide replication fidelity as detailed in the SI 8. This corresponds to the minimum fidelity that a base-by-base replicator would require to create the same number of errors within the ligated strand. We combined binding analysis with one mutation in the ligated strands using

Figure 5: Assembly of long RNA via splinted ligation of 2′, 3′ cyclic phosphate containing oligonucleotides. a, Schematics of sequence design. Two strands (*c* and *d*) with complementary sub-regions and not corresponding to the complete reverse complement, bind to form a long chain with repeating units of each. Both *c* and *d* contain 2′, 3′-cyclic phosphate, and the ligation can yield all possible length multiples of the initial strands. For the case where *c* and *d* are 16-mer long, denaturing PAGE of ligation reaction at -80 °C, -20 °C(frozen) and 5 °C (**b**) revealed products of up to five concatenations. The concentration of each ligation product (up to n=5), obtained from the SYBR Gold fluorescence analysis, is plotted in **c.** Similarly, for an initial strand size of 8-mer, denaturing PAGE (**d**) and product concentration (**e**) are shown. The optimal temperature for splinted ligation depends on oligomer length, as 5 °C yields higher concentration for 16-mer while -20 °C is better for 8-mer. Reaction were performed with 10 μM primers, 10 μM template, 50 mM CHES pH 10 and 1 mM MgCl2 for 7 days. Data are represented as mean of three independent replicates.

The literature reports a decrease in sequence specificity at lower temperatures [18,50], due to the stabilization of weak interactions within the duplex. This would enable the two moieties (>P and 5′-OH) to be in close proximity for a favorable attack. Therefore, at higher temperatures specificity would likely increase, while shifting the equilibrium towards the hydrolysis of the cyclic phosphate. While the non-enzymatic ligation studied here has higher error rate in comparison to ribozyme or enzyme-catalyzed reactions, this results in a higher phenotypic diversity, which could help

explore the sequence space and select for ribozymes that could eventually perform with a better fidelity[51].

After understanding that the non-enzymatic ligation with 2′,3′-cyclic phosphates is a reliable copying mechanism on the sequence level, we aimed to investigate the potential for elongation. This would have been an important characteristic of a potential prebiotic replication mechanism, as it would establish a link between non-templated nucleotide condensation and the faster ribozymatic replication. However, the majority of ribozymes, and particularly those catalyzing replication reactions are long, with tens or hundreds of nucleotides[52–54].

We explored the possibility of bridging these two oligonucleotide length regimes by designing splint strands with short (4- or 8-mer) binding regions that can both cross-template and ligate (Figure 5, **a**). Each system has two strands (labelled *c* and *d*) with >P. The strands are designed such that the 5′ half of the strand *d* is a reverse complement to the 5′ half of *c* and the same goes for the 3′ halves such that they bind and form a long network of *ccc…* bound to *ddd…* (Figure 5 **a**, see SI 3 for sequence information). The formed secondary structure allows for multiple ligations resulting in homopolymers of *c* and *d*. Figures 5 **b** and **d** show that up to 5 concatenations (n=6) could be detected for both the 16-mer system (Figure 5 **b**) and 8-mer system (Figure 5 **d**) resulting in 96- and 48-mer RNA respectively. Figures 5 **c** and **e** show the gel quantification of the respective gels in Figures 5 **b** and **d**. For the shorter system, the yield was reduced by about one order of magnitude for each additional concatenation, and was generally lower for the shorter system, which was likely due to the slower kinetics on ice.

It is interesting to note that the maximum yield for the 16-mer system was obtained at 5 °C whereas for the 8-mer system it was at -20 °C. We believe this to result from the low duplex stability of the 4-mer duplex region at 5 °C. We tested this by studying the ligation of shorter systems (SI 5). Systems with 4- and 6-mer binding regions were designed to have the same sequence at the ligation site and the same GC content as the strands used in Figures 1 and 2. While there was a significant reduction in yield for both the systems, we found that for the 6-mer system, the yields were higher at 5 °C, whereas for the 4-mer system -20 °C was more favorable (Figure S5.1). Additionally, the obtained yields at 7 days were very low (about 3%) for the latter. This suggests that to stabilize the duplex for short oligonucleotides, a compromise between the slower rate of ligation and the low probability of duplex formation must be made.

With the shown results, we demonstrate that the ligation reaction with 2′,3′-cyclic phosphate is highly robust, encompassing orders of magnitudes in length, essentially bridging the non-enzymatic monomer polymerization to the regime where long, functional ribozymes can evolve.

Conclusion

A non-enzymatic replicator chemistry on the early Earth should have the capacity, under plausible conditions, to elongate strands and undergo further replication steps all while being highly accurate and processive. We demonstrate with this work that ligation with >P RNA can fulfill these criteria.

Firstly, we show that the template-directed replication mechanism, with >P RNA, is robust, reproducible and achieves high conversion yields in both aqueous and frozen alkaline solutions. For the former case, 40% yield was observed in contrast to previously reported yield of 16% for similar reactions. It was found that a combination of high pH and low temperature promotes ligation over the hydrolysis of the cyclic phosphate moiety. However, a fraction of the >P still hydrolyzes, which contributes to its incomplete conversion. While the yield could be further improved by adding reagents that aid in the re-cyclization of the monophosphate moiety, such as diamidophosphate in combination with imidazole[27], this would increase the complexity of the system. Moreover, this reaction has low salt requirements (1 mM $MgCl_2$, Figure S4.3) ensuring RNA backbone integrity and thus can proceed in a wide pH range, even un-buffered (SI 4).

The elongation of short RNA was demonstrated with splinted systems that yielded up to six-copy concatemers of short RNA strands of either 16- or 8-mer, resulting in long RNA on the scale of 100-mer (Figure 5). This length range approaches the average length of replicating ribozymes [52–54], representing a significant step toward assembling functional RNAs by plausible means. Even very short >P RNA fragments (with 4-mer base-pairing regions) ligate under frozen conditions (SI 5), establishing a bridge from the single nucleotide condensation reactions, yielding very short RNA strands, to a regime where templated ligation reactions could dominate.

Furthermore, we evaluate the copying accuracy with two templates with varying nucleotides at the ligation site. One single mutation at either the 3′ or 5′ end nucleotide resulted in a reduction by more than 82% in yield (Figure 4), even when the remaining primer was entirely complementary. We estimate a per-nucleotide fidelity of at least 89% for the studied system. If we would have chosen primers with length 4 to 6 nucleotides, we estimated that a base-by-base replicator offers a fidelity between 95% and 98% depending on whether the ligation includes a G-base or a U-base on the template, respectively.

Contrary to previous studies on >P, we found that under the tested conditions, the reaction was not regioselective, producing equal amounts of 2′–5′ and 3′–5′ linkages at the ligation site (Figure 3). While approximately half of the linkages were non-canonical, we argue this does not diminish the applicability of the reaction in a prebiotic context. Strands with 2′–5′ linkages have been shown to fold into functional structures[49], and these non-canonical linkages have also been demonstrated to be more labile than 3′–5′ and have

5

potential for interconversion[20,55]. Furthermore, Figure S6.3 shows that the product *ab* with a 2′–5′ linkage at the ligation site could still template the reverse ligation reaction. The non-canonical linkage in the template at the ligation site did not impede ligation, paving the way for exponential replication cycles. Especially the low-salt requirements, particularly the low amount of Mg^{2+}, allow for heat-driven strand separation, while limiting backbone hydrolysis. Thus, one can envision a ligation chain reaction triggered by thermal cycling of the reaction mixture similar to Edeleva et al., with the benefit of not generating deleterious side-products by EDC.[21] Additionally, a ligating ribozyme such as sunY [56,57] having typically high (50 mM) magnesium requirements, has been shown to ligate in the low (5 mM) Mg^{2+} solutions under thermal non-equilibrium conditions of an air-water interface which offers continuous feeding [26]. This suggests that such scenarios could provide a niche where the ligation reactions by 2′,3′-cyclic phosphate could evolve towards a ribozymatic replicator.

Considering these results, ligation with >P is an interesting framework to produce diverse pools of long RNA that could undergo molecular evolution. We show that the system described in the current study enables the robust generation of long RNA, with high fidelity. This was demonstrated for a range of lengths, sequence combinations, reaction conditions and temperatures suggesting that ligation of RNA with >P holds a central position in the general conception of the RNA world.

Acknowledgments

Keywords: cyclic nucleotide • non-enzymatic ligation • prebiotic chemistry

Bibliography

[1] G. Costanzo, S. Pino, A. M. Timperio, J. E. Šponer, J. Šponer, O. Nováková, O. Šedo, Z. Zdráhal, E. Di Mauro, *PLOS ONE* **2016**, *11*, e0165723–e0165723.

[2] G. Costanzo, S. Pino, F. Ciciriello, E. Di Mauro, *J. Biol. Chem.* **2009**, *284*, 33206–33216.

[3] M. Morasch, C. B. Mast, J. K. Langer, P. Schilcher, D. Braun, *ChemBioChem* **2014**, *15*, 879–883.

[4] J. P. Ferris, A. R. Hill, R. Liu, L. E. Orgel, *Nature* **1996**, *381*, 59–61.

[5] M. S. Verlander, R. Lohrmann, L. E. Orgel, *J. Mol. Evol.* **1973**, *2*, 303–316.

[6] C. V. Mungi, N. V. Bapat, Y. Hongo, S. Rajamani, *Life* **2019**, *9*, 1–11.

[7] M. S. Verlander, L. E. Orgel, *J. Mol. Evol.* **1974**, *3*, 115–120.

[8] S. J. Zhang, D. Duzdevich, D. Ding, J. W. Szostak, *Proc. Natl. Acad. Sci.* **2022**, *119*, e2116429119.

[9] L. Li, N. Prywes, C. P. Tam, D. K. O'Flaherty, V. S. Lelyveld, E. C. Izgu, Ayan Pal, A. Pal, J. W. Szostak, *J. Am. Chem. Soc.* **2017**, *139*, 1810–1813.

[10] S. R. Vogel, C. Deck, C. Richert, *Chem. Commun.* **2005**, 4922–4924.

[11] S. Dagar, S. Sarkar, S. Rajamani, *Orig. Life Evol. Biospheres* **2023**, DOI 10.1007/s11084-023-09636-z.

[12] D. Ding, L. Zhou, S. Mittal, J. W. Szostak, *J. Am. Chem. Soc.* **2023**, jacs.3c00612.

[13] Y. Brudno, M. E. Birnbaum, R. E. Kleiner, D. R. Liu, *Nat. Chem. Biol.* **2010**, *6*, 148–155.

[14] Jesse J. Chen, J. Chen, X. Cai, J. W. Szostak, *J. Am. Chem. Soc.* **2009**, *131*, 2119–2121.

[15] K. Murayama, H. Okita, T. Kuriki, H. Asanuma, *Nat. Commun.* **2021**, *12*, 804–804.

[16] L. Zhou, L. Zhou, D. K. O'Flaherty, J. W. Szostak, *J. Am. Chem. Soc.* **2020**, *142*, 15961–15965.

[17] F. Wachowius, P. Holliger, *ChemSystemsChem* **2019**, *1*, 1–4.

[18] L.-F. Wu, Z. Liu, S. J. Roberts, M. Su, J. W. Szostak, J. D. Sutherland, *J. Am. Chem. Soc.* **2022**, *144*, 13920–13927.

[19] R. Rohatgi, Rajat Rohatgi, David P. Bartel, D. P. Bartel, D. P. Bartel, J. W. Szostak, *J. Am. Chem. Soc.* **1996**, *118*, 3332–3339.

[20] R. Rohatgi, Rajat Rohatgi, David P. Bartel, D. P. Bartel, D. P. Bartel, J. W. Szostak, *J. Am. Chem. Soc.* **1996**, *118*, 3340–3344.

[21] E. V. Edeleva, A. Salditt, J. Stamp, P. Schwintek, J. H. van Esch Jh, J. Boekhoven, D. Braun, *Chem. Sci.* **2019**, *10*, 5807–5814.

[22] G. von Kiedrowski, *Angew. Chem.* **1986**, *25*, 932–935.

[23] C. T. Obianyor, G. P. Newnam, B. E. Clifton, M. A. Grover, N. V. Hud, *ChemBioChem* **2020**, *21*, 3359–3370.

[24] K. Gao, L. E. Orgel, *Orig. Life Evol. Biospheres* **2000**, *30*, 45–51.

[25] K. Le Vay, E. Y. Song, B. Ghosh, T.-Y. D. Tang, Hannes Mutschler, H. Mutschler, *Angew. Chem.* **2021**, DOI 10.1002/anie.202109267.

[26] A. Salditt, L. Karr, E. Salibi, K. Le Vay, D. Braun, H. Mutschler, *Nat. Commun.* **2023**, *14*, 1495.

[27] E. Y. Song, E. I. Jiménez, H. Lin, K. Le Vay, R. Krishnamurthy, H. Mutschler, *Angew. Chem.* **2020**, *60*, 2952–2957.

[28] A. V. Dass, S. Wunnava, J. Langlais, B. von der Esch, M. Krusche, L. Ufer, N. Chrisam, R. C. A. Dubini, F. Gartner, S. Angerpointner, C. F. Dirscherl, P. Rovó, C. B. Mast, J. E. Šponer, C. Ochsenfeld, E. Frey, D. Braun, *ChemSystemsChem* **2023**, *5*, DOI 10.1002/syst.202200026.

[29] J. M. Buzayan, W. L. Gerlach, G. Bruening, *Nature* **1986**, *323*, 349–353.

[30] R. Breslow, M. Labelle, *J. Am. Chem. Soc.* **1986**, *108*, 2655–2659.

[31] Y. Li, R. R. Breaker, *J. Am. Chem. Soc.* **1999**, *121*, 5364–5372.

[32] R. Lohrmann, L. E. Orgel, *Science* **1968**, *161*, 64–66.

[33] M. A. Crowe, J. D. Sutherland, *Chembiochem* **2006**, *7*, 951–956.

[34] K. J. Hertel, D. Herschlag, O. C. Uhlenbeck, *Biochemistry* **1994**, *33*, 3374–3385.

[35] K. K. Le Vay, H. Mutschler, in *DNAzymes* (Eds.: G. Steger, H. Rosenbach, I. Span), Springer US, New York, NY, **2022**, pp. 301–309.

[36]   A. V. Vlassov, B. H. Johnston, L. F. Landweber, S. A. Kazakov, *Nucleic Acids Res.* **2004**, *32*, 2966–2974.

[37]   D. A. Lafontaine, D. Beaudry, P. Marquis, J.-P. Perreault, *Virology* **1995**, *212*, 705–709.

[38]   A. Flynn-Charlebois, Y. Wang, T. K. Prior, I. Rashid, K. A. Hoadley, K. A. Hoadley, R. L. Coppins, and A. C. Wolf, S. K. Silverman, *J. Am. Chem. Soc.* **2003**, *125*, 2444–2454.

[39]   A. V. Lutay, E. L. Chernolovskaya, M. A. Zenkova, V. V. Vlasov, *Dokl. Biochem. Biophys.* **2005**, *401*, 163–166.

[40]   A. V. Lutay, E. L. Chernolovskaya, E. L. Chernolovskaya, M. A. Zenkova, M. A. Zenkova, V. V. Vlassov, V. V. Vlassov, *Biogeosciences Discuss.* **2006**, *3*, 243–249.

[41]   A. V. Lutay, I. V. Grigoriev, M. A. Zenkova, E. L. Chernolovskaya, V. V. Vlassov, *Russ. Chem. Bull.* **2007**, *56*, 2499–2505.

[42]   Y. Staroseletz, S. Nechaev, E. Bichenkova, R. A. Bryce, C. Watson, V. Vlassov, M. Zenkova, *Biochim. Biophys. Acta - Gen. Subj.* **2018**, *1862*, 705–725.

[43]   A. V. Lutay, M. A. Zenkova, V. V. Vlassov, *Chem. Biodivers.* **2007**, *4*, 762–767.

[44]   M. Renz, R. Lohrmann, L. E. Orgel, *Biochim. Biophys. Acta BBA - Nucleic Acids Protein Synth.* **1971**, *240*, 463–471.

[45]   M. Kinjo, R. Rigler, *Nucleic Acids Res.* **1995**, *23*, 1795–1799.

[46]   J. G. Wetmur, N. Davidson, *J. Mol. Biol.* **1968**, *31*, 349–370.

[47]   I. I. Cisse, H. Kim, T. Ha, *Nat. Struct. Mol. Biol.* **2012**, *19*, 623–627.

[48]   D. A. Usher, D. Yee, *J. Mol. Evol.* **1979**, *13*, 287–293.

[49]   J. Sheng, L. Li, A. E. Engelhart, J. H. Gan, J. W. Wang, J. W. Szostak, *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 3050–3055.

[50]   H. Mutschler, A. I. Taylor, B. T. Porebski, A. Lightowlers, G. Houlihan, M. Abramov, P. Herdewijn, P. Holliger, *eLife* **2018**, *7*, DOI 10.7554/eLife.43022.

[51]   K. Leu, B. Obermayer, S. Rajamani, U. Gerland, I. A. Chen, *Nucleic Acids Res.* **2011**, *39*, 8135–8147.

[52]   L. Jaeger, Luc Jaeger, M. C. Wright, G. F. Joyce, *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 14712–14717.

[53]   D. P. Horning, G. F. Joyce, *Proc. Natl. Acad. Sci.* **2016**, *113*, 9786–9791.

[54]   D. M. Shechner, R. A. Grant, S. C. Bagby, Y. Koldobskaya, J. A. Piccirilli, D. P. Bartel, *Science* **2009**, *326*, 1271–1275.

[55]   P. Jarvinen, M. Oivanen, H. Lonnberg, *J. Org. Chem.* **1991**, *56*, 5396–5401.

[56]   J. A. Doudna, Jennifer A. Doudna, Jennifer A. Doudna, N. Usman, J. W. Szostak, *Biochemistry* **1993**, *32*, 2111–2115.

[57]   J. A. Doudna, J A Doudna, Jennifer A. Doudna, S. Couture, J. W. Szostak, *Science* **1991**, *251*, 1605–1608.