REPLICATION AND KINETIC TRAPPING OF NUCLEIC ACIDS IN ALTERNATIVE ENVIRONMENTS

A Dissertation
Presented to
The Academic Faculty

by

Adriana Lozoya Colinas

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Chemistry and Biochemistry

Georgia Institute of Technology
December 2020

REPLICATION AND KINETIC TRAPPING OF NUCLEIC ACIDS IN ALTERNATIVE ENVIRONMENTS

Approved by:

Dr. Nicholas V. Hud, Advisor
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Amanda Stockton
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Martha A. Grover
School of Chemical & Biomolecular Engineering
Georgia Institute of Technology

Dr. Adegboyega (Yomi) Oyelere
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Loren Williams
School of Chemistry and Biochemistry
Georgia Institute of Technology

Date Approved: October 16, 2020

ACKNOWLEDGEMENTS

I would like to thank my mom and dad for all their support and setting up an example for me to follow. I really appreciate everything you have done to encourage me to succeed and follow my dreams.

I want to thank Mario for always supporting me. I know it hasn't always been easy being far away, thank you for being patient and supportive with me. Thank you for the time and adventures we lived together.

To all my Latino family at Georgia Tech, thank you for making me feel closer to home. We spent a lot of time together, learned a lot from each other and shared our culture, all of which have made my PhD experience more enjoyable.

I would also like to acknowledge my advisor, Nick Hud, for being supportive and sharing his passion for science with me. I have learned a lot from you, as a mentor, as a leader, and as a scientist. And to Martha Grover, you were always encouraging and empathetic. Thank you for helping me improve my communication skills and help me thrive inside and outside the lab.

I want to thank Christine Conwell for always being there for us and push us to always give a little more. You have always believed in us; you have been a great mentor and friend. To all the CCE members, for leading interesting discussions and providing good feedback. I am really thankful for the community and friendship we formed within the center.

To all the lab members, thank you for being great mentors and friends. I want to specially thank Christine He and Isaac Gallego for teaching me so many things in the lab. To Chiamaka for always being there for me, for being such a great friend and for all the fruitful discussions we had. To Gary Newnam, thank you for always being there to solve any problem we had, for your advice and willingness to help us succeed, and to Bryce Clifton for helping me push forward the last project in which we collaborated. Thank you Brian, David, Denise, Chelsea, Martin, Suneesh, Bradley, Moran, Sreejith, Tyler, Irene, Kim, Aiko, Cesar; it was a great pleasure to share the lab with you and be able to discuss a very wide variety of topics with all of you.

Lastly, I would like to acknowledge the Consejo Nacional de Ciencia y Tecnología (CONACYT) for their funding, scholarship 382818.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

bp	Base pairs
CD	Circular dichroism
Cy3, Cy5	Cyanine dyes
DNA	Deoxyribonucleic acid
ds	Double stranded
FAM	Carboxyfluorescein
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
MES	2-(N-morpholino) ethanesulfonic acid
nt	Nucleotides
PAGE	Polyacrylamide gel electrophoresis
RNA	Ribonucleic acid
ss	Single stranded
TAE	Tris/acetate/EDTA
TBE	Tris/borate/EDTA
Tris	Tris (hydroxymethyl) aminomethane
U	Units
UAcW	Urea/acetamide/water

SUMMARY

Many hypotheses in origins of life research state that informational polymers likely appeared before coded protein enzymes. Replication of these polymers is a process essential for life to exist. However, a robust enzyme-free replication system has not yet been demonstrated. Among the challenges associated with prebiotic replication of nucleic acid duplexes is strand inhibition. Specifically, to copy information from a nucleic acid duplex, the two strands need to be separated to serve as templates. This separation can be achieved through an increase the temperature; however, upon cooling, the two template strands will reanneal before oligomers can bind to the template to form new copies. Another challenge involves structured sequences, which are associated with catalytic activity (ribozymes), but due to their structures, they are less likely to serve as templates. A third challenge is the transition between the duplex state and the active fold of a functional sequence. Finally, having sustained replication cycles through simulated prebiotic environments has not yet been demonstrated.

In this thesis I explain how the use of alternative solvents such as glycholine and UAcW can provide a way to circumvent the aforementioned challenges to promote replication of DNA and RNA duplexes. Additionally, I demonstrate how viscous solvents can provide the conditions for trapping of sequences in their single stranded, active conformation despite the presence of their complementary strand.

CHAPTER 1. INTRODUCTION

1.1 Overview

Nucleic acids are essential for life. They are in charge of a lot more than just storing information. For example, mRNA can copy information from a DNA strand and transfer that information to direct the synthesis of proteins, which takes place in the ribosome. The ribosome is a molecular machine comprised of RNA and proteins, and the RNA is in charge of doing the catalysis for the peptide bond formation. tRNA is in charge of delivering the amino acids to the ribosome so a new protein can be synthesized[1]. Furthermore, other RNA sequences capable of catalyzing chemical reactions have been found in nature, for example, the hammerhead or the hairpin ribozymes are capable of self-cleavage at a specific site[2-4]. The discovery of RNA catalysis was so important that Altman and Cech were awarded the Nobel Prize in Chemistry in 1989[5].

All of these discoveries led many researchers to support the RNA World hypothesis, which states that RNA could have been one of the first informational polymers at the origins of life[6]. It was proposed that RNA in addition of being in charge of information storage, was also fulfilling some of the roles that enzymes play nowadays as catalysts. Given this scenario, a lot of research has been done to explore ways in which RNA could catalyze its self-replication. Polymerase ribozymes that can self-replicate or replicate a template strand have been selected in vitro[5,7,8], and non-enzymatic ways of template-directed primer extension have been proposed by using activated substrates[9,10]. Nonetheless, all of this research has shown a lot of new challenges and more questions to be answered.

1.2 Challenges to prebiotic replication of nucleic acids

Information storage and transfer is essential for life. However, the replication of informational polymers is still an unanswered question from the origins of life perspective. There are many challenges associated with this process, which have prevented efficient information transfer in a prebiotic way. Some of these challenges, which will be addressed in this work include the strand inhibition problem, the replicator-catalyst paradox and the transition between the duplex and the active fold of a functional sequence.

First, when information is transferred from one DNA or RNA strand, the product formed will result in a duplex. This duplex prevents the two strands from serving as templates in further replication rounds. The strands can be separated by increasing the temperature, however, as the temperature is lowered, the two strands will reanneal before monomers or oligomers can bind to the template. This is known as the strand inhibition problem[11-14]. Previous studies of non-enzymatic replication have used a simplified model system with a single strand that directs the synthesis of its complement[10,15,16], however, multiple replication cycles are not possible in these models. Another approach that has been taken to overcome the strand inhibition problem is the lowering of the melting temperature of a duplex, for example, by shortening the oligomers used or by using G:U wobble pairs[17]. Previous work in the Hud lab showed that this problem can be overcome by using a viscous environment, which slows down the reformation of the duplex and allows shorter oligonucleotides to diffuse faster and bind to the template allowing information transfer from the template strand[11].

Related to this strand inhibition challenge is also the fact that after replication, an active sequence needs to transition from the duplex state to its active folded structure to perform catalysis despite the presence of its complementary sequence. This process is not favored in aqueous solution and even though there are some reports where ribozymes were obtained through template directed synthesis, they were not proved to be active in the presence of their complementary strand[8,10,18,19].

Another challenge associated with replication in the RNA world scenario is that if the templates possess catalytic activity, such as ribozymes or DNAzymes, then they are required to fold in very specific ways to be active. Such stable structures prevent them from being good templates for replication in aqueous solutions, also known as the replicator-catalyst paradox[12,20]. Even in cases where an RNA polymerase ribozyme was used to replicate structured templates, degradation of RNA outcompeted template synthesis of new sequences[18]. In most cases, sequences that have been replicated are too short to have complex structures or are specifically designed to not possess internal structure[21].

Finally, the replication cycles that have been demonstrated are through PCR-like reactions, using a regular polymerase[22], or a polymerase ribozyme[8], in which individual nucleotides are added one a time, an enzyme is present and fast temperature changes are required. Non-enzymatic replication cycles were reported using activating agents to ligate two oligomers assembled on a template[21], but the sequence was designed to not have any internal structure and continuous feeding of substrate was required.

To overcome these problems, we propose the use of alternative environments with no water or low water activity. Previously, information transfer from one strand of a DNA

and an RNA duplex was demonstrated in glycholine, a solvent composed of glycerol and choline chloride in a 4:1 molar ratio. Glycholine lowered the Tm of nucleic acids, and slowed down the reannealing of the template strands, allowing oligomers to bind to the template. The assembled oligomers were ligated, demonstrating information transfer from one of the template strands. This study demonstrates a possible solution to the strand inhibition problem, one of the main challenges in nucleic acid replication[11]. A similar system is reported here to promote replication of both strands of a duplex, including an active sequence, thus overcoming the strand inhibition problem.

Furthermore, glycholine promotes the folding of 2D DNA origami under isothermal conditions[23]. Therefore, it could potentially be used to promote the folding of active sequences once they are separated from their complementary strand, targeting the challenge of transitioning between duplex and active folds of a functional sequence. We validated this hypothesis by kinetically trapping a DNAzyme and a ribozyme in their active conformation in glycholine.

Finally, multiple replication rounds of a DNA duplex were achieved through wet-dry cycling in a urea and acetamide based solvent, something that had never been reported before. This simulates a more natural system, since the components of the mixture were likely abundant on the prebiotic Earth. It also simulates temperature and water content oscillations that can drive replication forward.

1.3 Nucleic acids in alternative solvents

Nucleic acids are not just soluble in water, they are also soluble in other non-aqueous solvents. Previous studies have shown that DNA stability can be affected and it can acquire different conformations in the presence of additives or in non-aqueous environments[24-27]. The solvents that have been explored include deep eutectic solvents (DES). DESs can be defined as "systems formed from a eutectic mixture of Lewis or Brønsted acids and bases which can contain a variety of anionic and/or cationic species"[28]. Most DES are composed of a quaternary ammonium salt and either a metal salt or a hydrogen bond donor. The interaction between these compounds causes the decrease in the melting point of the mixtures as compared to the individual components[28].

DNA can be present in different conformations in these solvents. For example, DNA triplex structures and G-quadruplexes have been identified in reline (a eutectic formed of urea and choline chloride in a 2:1 molar ratio), and the Tm of DNA and RNA sequences in this solvent is lower compared to the aqueous solutions containing the individual components of the eutectic[24]. In this solvent, the human telomer sequence (HTS) DNA adopts a parallel-stranded G-quadruplex. HTS becomes kinetically trapped after quick cooling. For this particular case, the solvent composition and viscosity determine the folding pathway and kinetics of folding [25]. This type of study can be extended to understand how other DNA and RNA sequences fold in alternative solvents and the dependence on the solvent compositions.

Another non-aqueous solvent that has been used to explore DNA folding is glycholine, composed of glycerol and choline chloride in a 4:1 molar ratio. In this solvent,

DNA acquires the B-form helical structure, and folding of 2D DNA origami under isothermal conditions has been demonstrated[23]. Furthermore, information transfer from one strand of a DNA and an RNA duplex was possible with the aid of this solvent. Glycholine slows down the reannealing of the template strands, allowing oligomers to diffuse faster and bind to the template; these assembled oligomers can then be ligated to form a copy of one of the template strands. This study demonstrates a possible solution to the strand inhibition problem, one of the main challenges in nucleic acid replication[11], which was used in this dissertation for replication of an RNA duplex.

Finally, SELEX was recently used to obtain DNA aptamers that target gluten proteins in ethaline, a deep-eutectic solvent composed of choline chloride and 1,2-ethanediol in 1:2 molar ratio. This study provides a method for selection of aptamers that can work with molecules insoluble in water[29], demonstrating that DNA can be used in alternative solvents and provide solutions to current problems, such as low solubility of a target molecule.

Overall, these studies demonstrate that the environment provided by the solvent is important to promote nucleic acid folding and the conformation of the sequences will depend on their surroundings. Therefore, this thesis focuses on how the environment provided by the solvent can affect the trapping of functional sequences and replication of nucleic acids under different prebiotic model scenarios.

The solvents that have been reported so far to evaluate the behavior of nucleic acids in viscous environments were not likely present in the early earth. However, they serve as good models to demonstrate how the environment can significantly alter the folding and

diffusion of nucleic acids. Under prebiotic conditions, solvents with low water content could have formed through water evaporation of a pool, leading to the increased concentration of solutes and resulting in a different environment (high viscosity, different dielectric constant, etc).

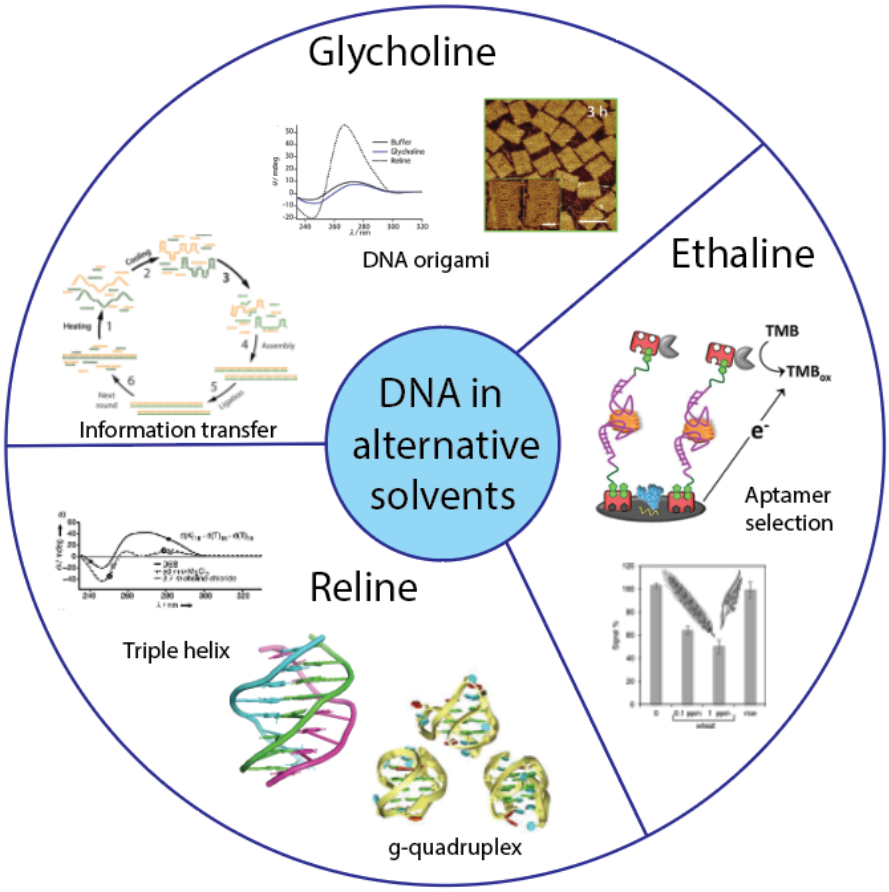


Figure 1. The environment can promote different DNA conformations and processes. Glycholine can enable information transfer and promotes B-form DNA[11,23]. G-quadruplexes and triple helices have been reported in reline[24,25]. Aptamers specific for gluten have been developed in ethaline[29].

CHAPTER 2. KINETIC TRAPPING OF A DNAZYME IN GLYCHOLINE

2.1 Introduction

Control over configurations of nucleic acids that can alternate between different states is crucial for the creation of switchable, functional nucleic acids. Switchable nucleic acids can have applications that range from development of nanotechnology[30,31] to the selection of functional sequences which could have implications in the origins of life[32]. As mentioned in the previous chapter, switching between the duplex and the active state of a functional sequence would be crucial to have replication and activity in a prebiotic environment.

It has been shown that molecular machines can alternate between active and inactive conformations through changes between environmental conditions. For example, variations in the chemical environment such as pH or salt concentration[33-35] can drive a switch towards the kinetic or thermodynamic conformation, nanomachines require continuous feeding of fuel strands[36,37], and some nano-switches require cooling rates of over 100 °C/ms[35]. Nonetheless, such drastic changes were unlikely to happen under prebiotic conditions, pH swings and addition of fuel strands can result in molecular crowding or high dilution after multiple switching cycles since they require constant addition of material to promote the switching, irreversibly changing solution conditions. Therefore, we explored other ways that could lead to the kinetic trapping of a structured

strand under mild conditions, without the need of continuous pH swings or extreme temperature changes (~100 °C/s).

Eutectic solvents have been demonstrated to affect the conformation of DNA[24,25]. In particular, glycholine was used to promote the folding of DNA origami[23] and slow down the diffusion of long strands due to its high viscosity. These reports suggest that glycholine is a solvent that could be used to slow down the reannealing of complementary strands and thus favor the folding of intramolecular structures under mild conditions.

Herein, we demonstrate that by varying the properties of the solvent, we can control the conversion between the thermodynamic (DNA duplex) and the kinetic (intramolecular fold) states of a DNA active sequence (DNAzyme). A self-cleaving DNAzyme (I-R3)[38] was selected for this study as a model sequence to evaluate the degree of intramolecular folding and self-cleaving activity depending on the conditions of the system. Based upon its predicted secondary structure, some variations were proposed to evaluate their influence in the trapping of the single strand.

2.2 Selection of DNAzyme sequence

There are many nucleic acid sequences that can catalyze reactions. Some of them consist of RNA (ribozymes) and are found in nature. For example, the ribosome, which performs the catalysis in protein synthesis, or the hairpin[39,40] and the hammerhead ribozymes[41,42], which hydrolyze specific RNA sequence targets. There are also DNA sequences that can perform catalysis, known as deoxyribozymes or DNAzymes[43]. Unlike ribozymes which

are found in nature, all deoxyribozymes that have been reported were artificially evolved in a lab[44]. For this project, a DNAzyme was chosen instead of a ribozyme because of its DNA is less prone to degradation than RNA and its secondary structure can be easily predicted. I-R3 is a self-cleaving DNAzyme. It belongs to class I deoxyribozymes, which have a catalytic core of 15 nucleotides and are active in the presence of Zn^{2+} (Figure 2). This particular sequence was chosen because it has a relatively simple architecture and has been well characterized[38].

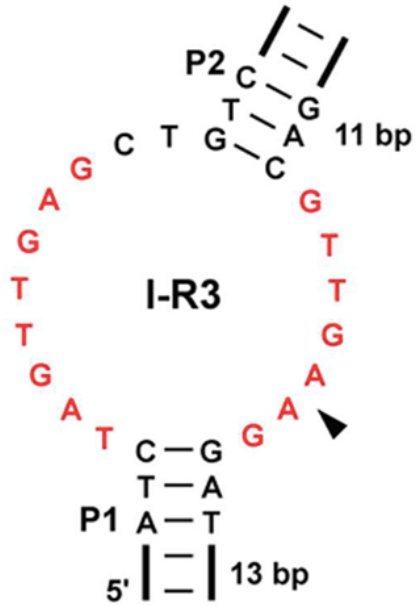


Figure 2. I-R3 DNAzyme structure. The conserved catalytic sequence is marked in red, and the arrow indicates the cleavage site. Image taken from Figure 3 of Gu, H., Furukawa, K., Weinberg, Z., Berenson, D. F. & Breaker, R. R. Small, highly active DNAs that hydrolyze DNA. *Journal of the American Chemical Society* 135, 9121-9129, doi:10.1021/ja403585e (2013).

2.3 Kinetic trapping of DNA strands in glycholine

The trapping of DNA structures in viscous solvents such as reline and glycholine has been previously reported[11,23,25]. However, the tunable properties of the system have not been exploited to control the degree of trapping and activity of a functional sequence in the presence of its complementary strand. We hypothesized that by changing environmental conditions such as cooling rates or viscosity, the system can be driven towards the intramolecular active conformation or the inactive duplex.

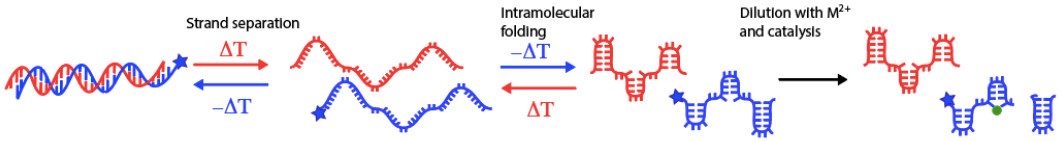


Figure 3. The DNA duplex is kinetically trapped by forming intramolecular structures after heat cycling. After aqueous dilution with Zn2+, the DNAzyme cleaves itself, demonstrating that the trapped structure folded in its active conformation.

The first aim was to demonstrate the formation of intramolecular folds of DNA sequences in the presence of their complementary strands (first two steps in Figure 3). To this end, a mixture containing the DNA duplex in glycholine was heated to 80 °C to promote denaturation of the DNA duplex (or 90 °C in aqueous solution), and then cooled back down to 20 °C, with a cooling rate of -48 °C/min (Figure 3). This test was done with three DNA strands of different lengths that form different structures, all of them with a fluorescein tag at the 5' end to quantify the degree of trapping. The analysis was done through non-denaturing polyacrylamide gel electrophoresis. The structures analyzed were:

one hairpin (51 bases long that contains the active part of I-R3 in its loop), and the I-R3 DNAzyme with two different stem lengths (49 and 80 bases long, Figure 4A).

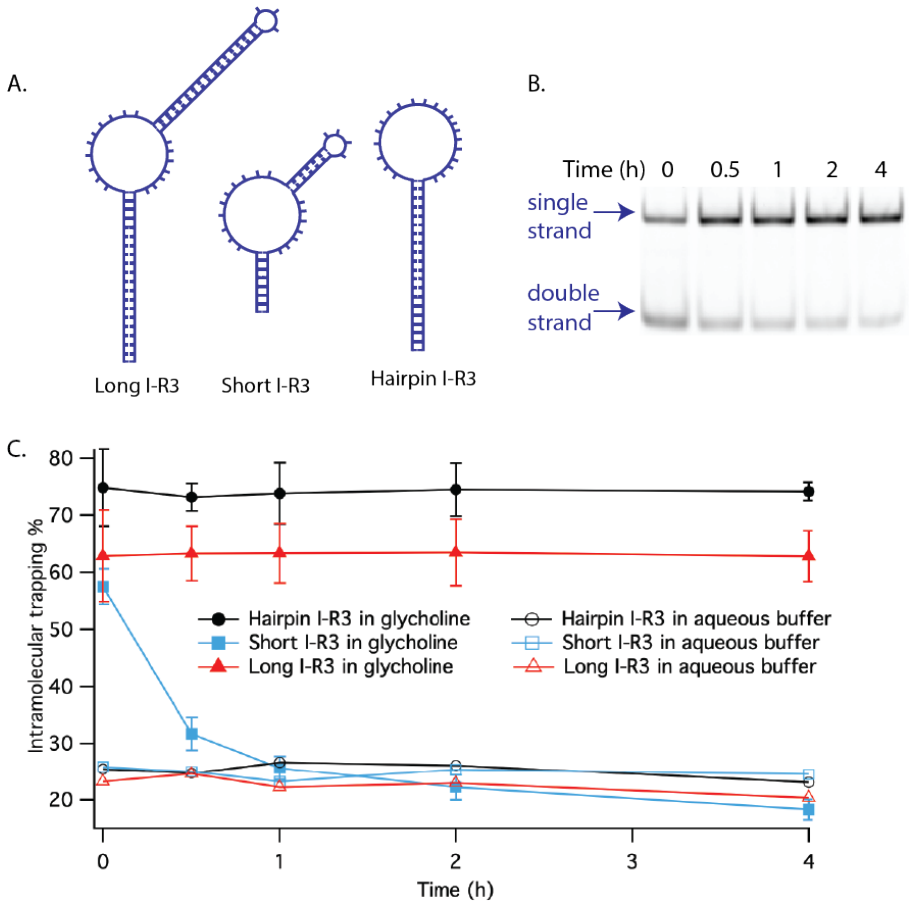


Figure 4. A) Secondary structures predicted by mfold[45] of the sequences used. B) Gel displaying trapping over time for the short-stem I-R3. C) Degree of intramolecular trapping over time of the different secondary structures after incubation for 5 minutes at 80 °C in glycholine, followed by cooling to room temperature at -48 °C/min.

Overall, a higher percentage (>60%) of intramolecular structures were formed in glycholine, whereas in water only up to 20% of the strands were kinetically trapped (Figure 4, time 0).

Following the heating and cooling cycle, samples were left at room temperature for up to 4 hours in glycholine to evaluate if the duplex reformed within this period of time. The hairpin and the long stem I-R3 stay kinetically trapped in their intramolecular folds. And even if some photobleaching of the fluorophore is observed for long times, the fraction of kinetically trapped strands remains the same. On the other hand, the short stem DNAzyme reanneals with its complement after one hour which is a more favorable thermodynamic state. This could be related with a lower melting temperature for the short stem I-R3. The single stranded long stem hairpin has a Tm of 50 °C in glycholine, whereas the short stem I-R3 has a Tm of 27 °C (Figure 5). Therefore, incubation at 20 °C can lead to the duplex reformation since a fraction of the sequences will not be completely folded.

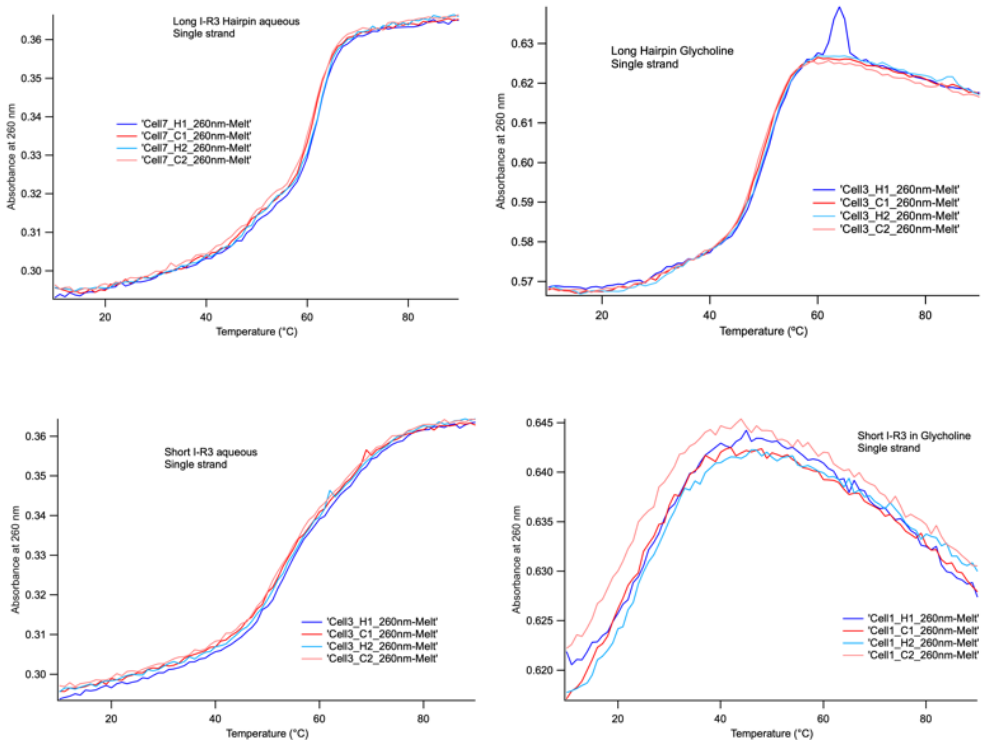


Figure 5. UV-Vis melts of the single stranded hairpin and short stem I-R3 in aqueous solution and in glycholine.

In addition to analyzing the duplex reformation kinetics in glycholine for up to 4 hours, samples were diluted 6x in aqueous buffer after trapping the sequences to see if upon water addition the duplex reformed. Buffer with and without Zn^{2+} was added immediately after the heat cycle. There appears to be no effect in the trapping of the hairpin and the long-stem I-R3 after the dilution; these sequences remained trapped for the length of the experiment. The short-stem I-R3 duplex reformation slows down when the dilution buffer contains Zn^{2+} (Figure 6). This is likely because the active conformation requires Zn^{2+} for cleavage, so it is likely that the divalent cation favors the folding of the active site, slowing down reannealing with its complementary sequence.

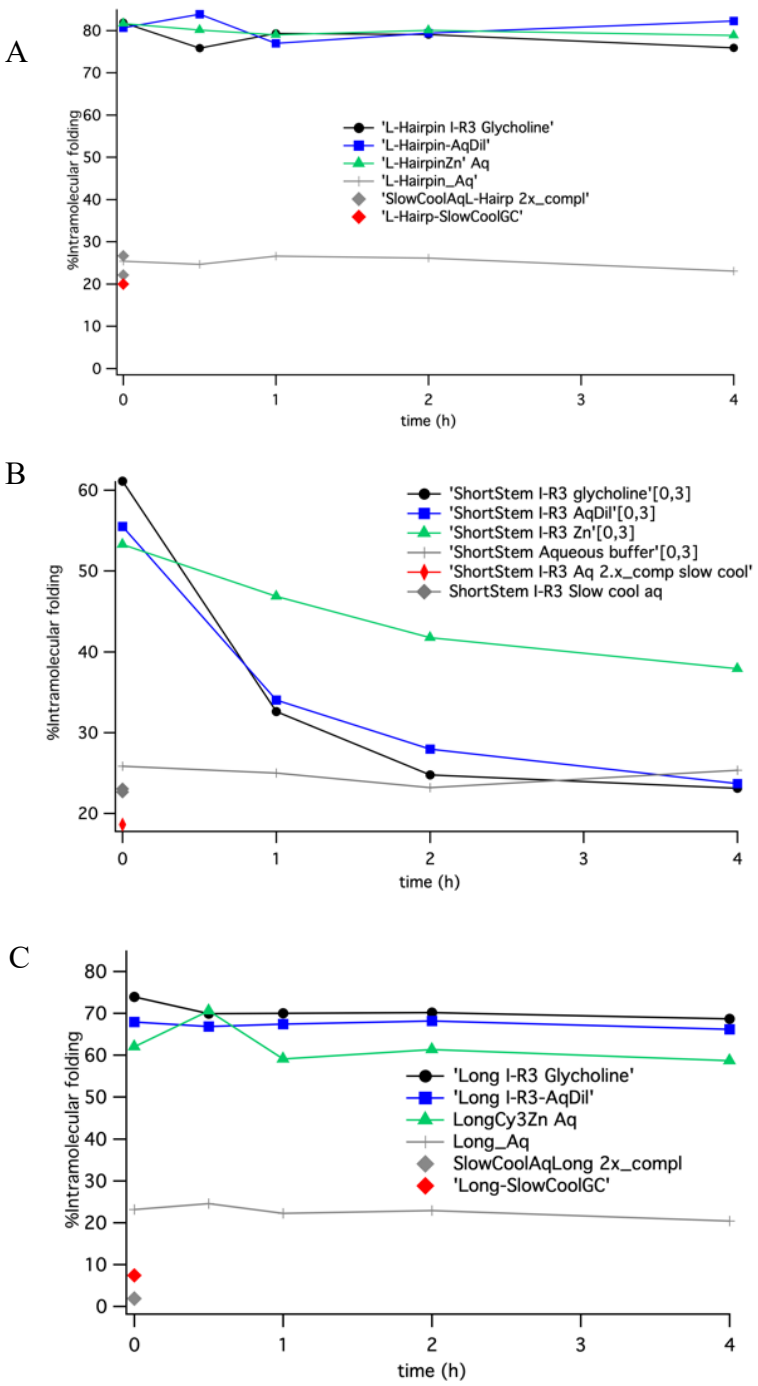


Figure 6. Kinetics of duplex reformation after aqueous dilution for A) hairpin, B) short-stem I-R3 and C) long-stem I-R3. Black markers: glycholine (no dilution), blue markers: aqueous dilution after heat cycle with buffer, green markers: aqueous dilution with Zn^{2+} and buffer after heat cycle, gray line: heat cycle in aqueous buffer with slow cooling rate, red markers: slow cooling rate in glycholine, gray markers: slow cooling in aqueous solution with 2x of complementary strand.

2.4 Conditions that affect kinetic trapping

There are various factors that could affect the trapping of intramolecular folds. In nanotechnology development, switchable nucleic acids have been widely explored in which the change between their conformations can be triggered by different environmental factors. Some of the most common factors that have been reported include variations in pH[33,34] and ionic strength[35]. Such variations require adding other reagents or coupled reactions to promote the change in the environment, which result in variations of the concentration and complicates reversibility to the initial state. Furthermore, variations in viscosity of the solvent to favor the change between conformations has not been reported. Other factors that were analyzed to see whether they affect the trapping of the single strands include the cooling rate and the concentration of the strands.

Assays in which the cooling rate was varied between -48 °C/min and -3.75 °C/min were performed to evaluate how the cooling rate affects the trapping of intramolecular folds and duplex reannealing. The two sequences used for this series of experiments were the I-R3 DNAzyme and the hairpin which proved to have a higher degree of trapping in glycholine. Overall, with slower cooling rates, the formation of intramolecular folds decreases, favoring the formation of the duplex with its complementary strand (Figure 7A and B, cooling rate axis).

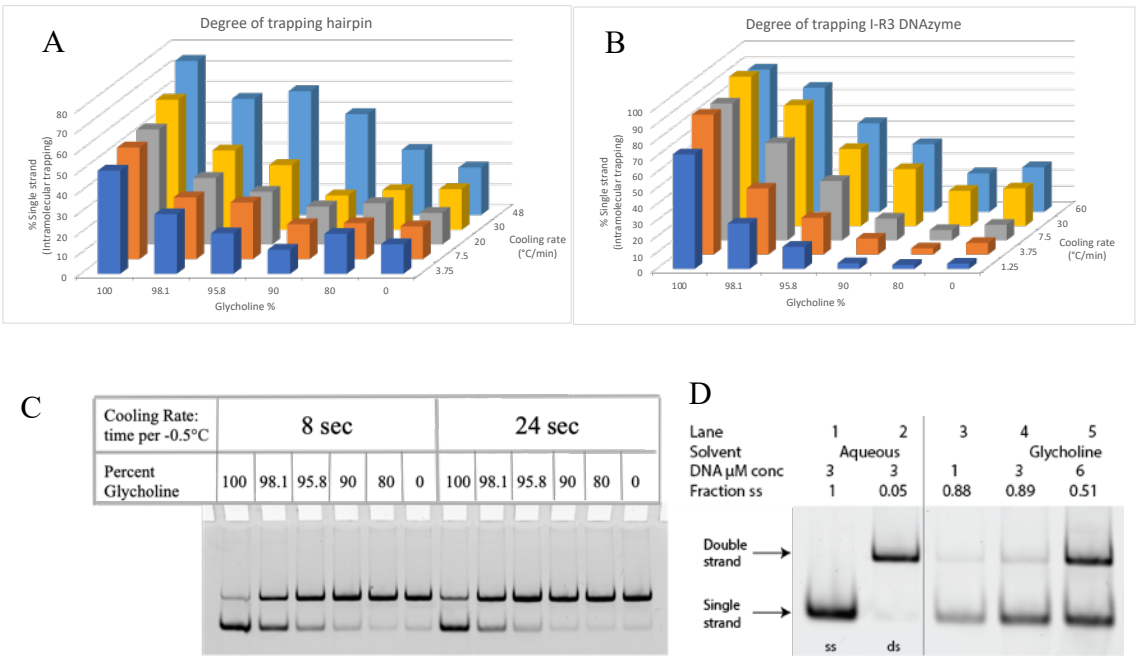


Figure 7. Degree of intramolecular trapping under different cooling rates and water content for A) Hairpin (initial DNA concentration: 3 µM), and B) I-R3 DNAzyme (initial DNA concentration: 1 µM). C) Sample polyacrylamide gel that shows the single and double strands of the I-R3 DNAzyme in aqueous solution and glycholine with different DNA concentrations (all gels can be found in the Appendix). D) Concentration dependence on trapping: at 1 and 3 µM, the degree of trapping is consistent, whereas at 6 µM the duplex formation starts to be more favourable due to the higher concentration of the strands.

Additionally, glycholine has the advantage of being miscible with water, which enables variations in solvent viscosity depending on the water content[11]. Therefore, trapping of these sequences was evaluated as a function of solvent viscosity. It was observed that with higher viscosity of the solvent (higher glycholine %), there is a higher percentage of kinetically trapped sequences (Figure 7A and B, Glycholine % axis).

The last variable evaluated was concentration of the duplexes, since lower concentration of the strands in the mixture promote more intramolecular trapping (Figure

7D). At 1 and 3 μM concentration the fraction of trapping of the single strand for the I-R3 DNAzyme is approximately the same, 0.88 ± 0.02 and 0.84 ± 0.04 respectively; while at a higher concentration (6 μM), the fraction for trapping for the single strand is reduced to 0.62 ± 0.12.

From these results, we can conclude that by tuning the environmental conditions, such as viscosity, cooling rate and concentration, the kinetic trapping of intramolecular folds can be controlled as desired.

We hypothesize that such fluctuations in the environment, which lead to the formation of either the kinetic (intramolecular folds) or thermodynamic (double stranded duplex) states of a DNA strand, could have led to the selection of functional nucleic acids with particular secondary or tertiary structures in the prebiotic Earth.

2.5 DNAzyme is active after folding in glycholine

It is important to demonstrate that the trapped sequences can perform the chemical reaction they were designed to do after being trapped as their single strands under adequate conditions. DNA duplexes have been shown to conserve their B-form helix in glycholine[23], consequently, we anticipated that the trapped sequences of the DNAzyme should be active given the proper conditions after aqueous dilution. It is important to note that I-R3 requires Zn^{2+} to cleave itself, nonetheless, nucleic acids tend to precipitate in glycholine in the presence of divalent cations. It is therefore necessary to make an aqueous dilution when Zn^{2+} is added. To demonstrate that the trapped strand can perform catalysis, the same

trapping experiment was done, and then the sample was diluted in aqueous buffer at pH 7.4 with 4 mM $ZnCl_2$ as proposed in Figure 3. The yield of self-cleaving activity from the trapped short-stem I-R3 sequence in glycholine in the presence of its complementary strand (black markers in Figure 8) is in agreement with the maximum activity achieved for the single stranded control in aqueous solution (green markers); whereas there is a very low yield of cleaved product after heating and cooling the duplex in aqueous buffer, consistent with the low degree of trapping under these conditions. This data is consistent with the higher degree of trapping shown in the previous section, suggesting that maximum trapping of the single strands in their active conformation can be achieved in a viscous solvent such as glycholine. It is important to note that even with the single strand control, 100% cleavage by I-R3 is hard to achieve. This DNAzyme is very sensitive to pH and $ZnCl_2$ concentration[38]. It has been reported that Zn^{2+} can form ZnO nanoparticles that can promote DNA adsorption and inhibit the cleavage reaction[46].

Overall, these experiments demonstrate that tuning of environmental conditions can lead to the trapping of structured folds and promote the folding of a DNAzyme into its active conformation. Therefore, the challenge for switching between a duplex state and the active single stranded conformation of a functional sequence can be solved by using a viscous solvent and varying some of the environmental conditions such as solvent viscosity and cooling rates.

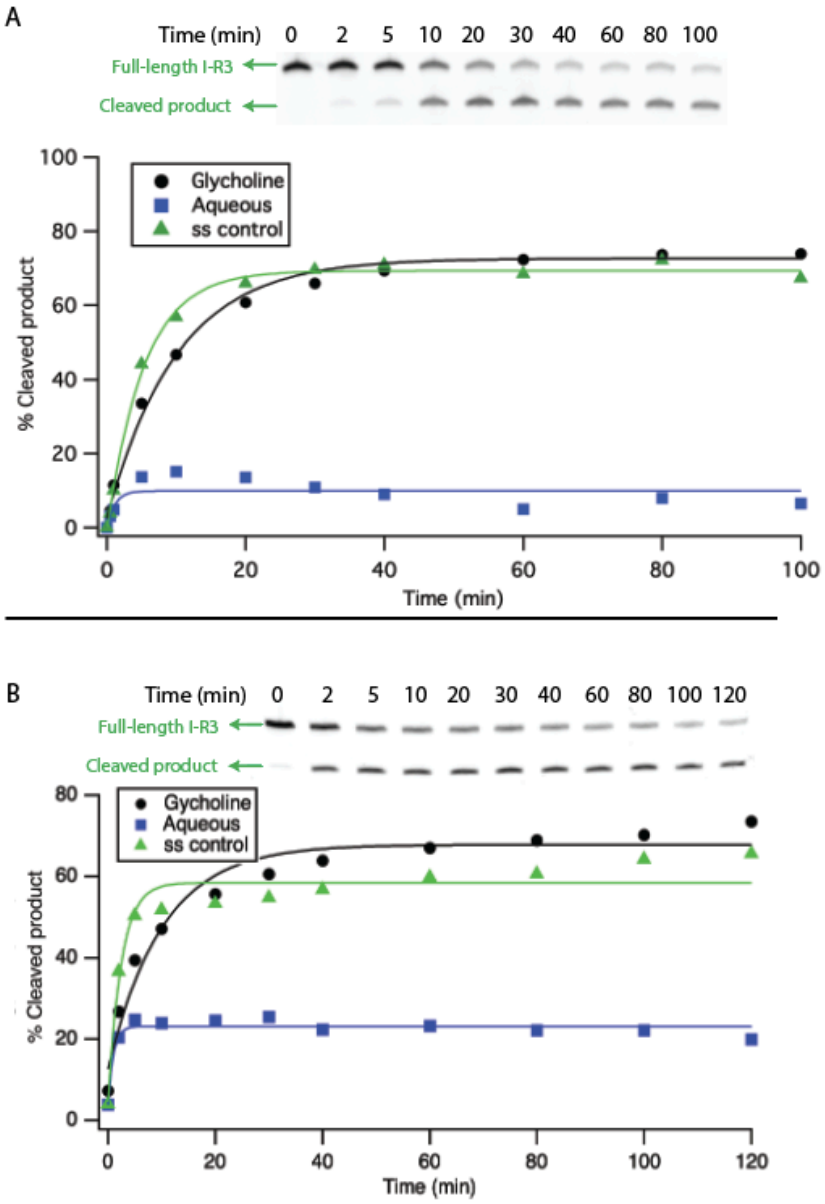


Figure 8. Plots of cleavage yield with a sample gel image for self-cleaving DNAzyme in presence of its complementary sequence after folding in glycholine (black markers) and in aqueous buffer (blue markers). A) Short stem I-R3. B) Long stem I-R3. Markers represent experimental data, and the lines are from exponential fits. Gels can be found in the Appendix.

One of the advantages of this system is that it does not require a change in pH or continuous feeding of strands as most molecular switches and machines do. Within the same system, just by changing the heating and cooling rate, one can drive the system towards the trapping of the single strand or the formation of a duplex. We have shown that this is possible with two sequences that possess different stability. We predict that this system should work with any sequence that has a defined secondary structure.

2.6 Sequence and structure modifications affect degree of trapping

All experiments presented so far were done with a highly structured sequence. Even though in the predicted secondary structure shown in Figure 4 there appears to be a loop in the strands, this part is likely to be structured since the conserved active and cleaving sites are in this region. To test how slight and major variations in the structure can affect the trapping of intramolecular structures, some modifications were introduced in the IR-3 sequence, both in the stem and the loop: i) mismatches were introduced in the middle of the stem, ii) the active fragment of the DNAzyme (located in the loop) was substituted by an unstructured sequence, iii) a toehold was added at the 3'end of one stem, iv) an unstructured sequence was used for comparison. The same length of 80 nucleotides was conserved for all of these strands. The secondary structures predicted from mfold[45] for these sequences are depicted in Figure 9 with the variations highlighted in light blue, and the sequences with their thermodynamic parameters can be found in Table 2.

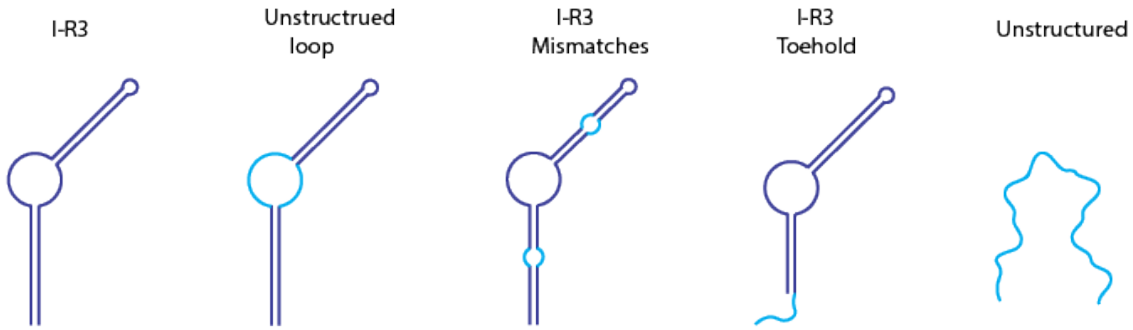


Figure 9. Secondary structure prediction of I-R3 and the modified sequences used to evaluate their effect on trapping: I-R3, Unstructured loop, I-R3 with mismatches, I-R3 with toehold, Unstructured 80mer.

2.6.1 Kinetic trapping of the single strands

The same protocol for trapping in glycholine was followed. Briefly, the DNA duplexes were heated to 80 °C and cooled to room temperature at -48 °C/min. These sequences were not FAM labeled, instead the gels were stained with SYBRgold and EtBr after running them. EtBr was used for densitometry analysis, and SYBRgold was used to verify whether there were any unstructured single strands that remained trapped.

As seen from Figure 10, I-R3 is the sequence with highest degree of internal structure of all the sequences that were evaluated. Following the original I-R3, the sequence that had the most trapping was the one designed to have an unstructured loop. This sequence still has the two fully complementary stems, which explains the high degree of trapping, followed by sequence that had mismatches within the stems. These mismatches lower the internal structure stability and thus favour rehybridization with its complementary strand. The sequence with a toehold showed almost no trapping, indicating that if there are enough free bases to which the complement can bind, then the duplex can reform through strand

displacement. Finally, there was no trapping observed for the unstructured sequence as expected. This doesn't mean that the two strands did not remained separate from each other in glycholine, just that since there are no intramolecular folds, the two strands cannot remain trapped apart from each other given the experimental conditions; if the strands were trapped in glycholine, they are likely to reanneal during the dilution process required to load the samples in the gel.

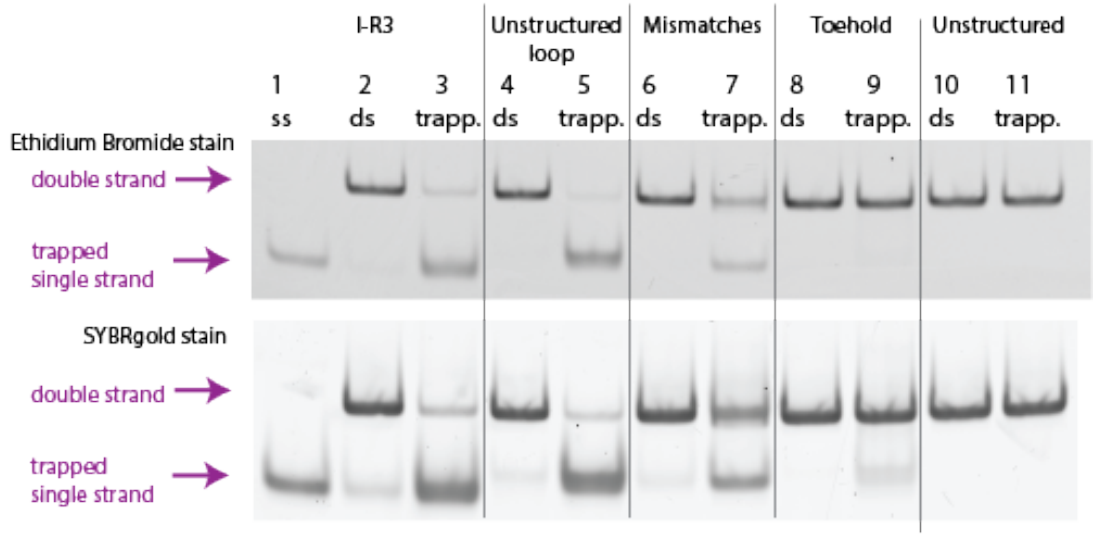


Figure 10. Comparison of trapping of single strands in glycholine for sequences with modified structures.

2.6.2 Binding of oligomers to kinetically trapped strands

Binding of short oligomers to a long template in glycholine was previously reported by He, Gallego et al[11]. However, in that case, the size difference between the template and the oligomers was of two orders of magnitude (the template was 3kb and the

oligonucleotides were 32 bases each). We hypothesize that such a difference in size length might not be necessary for the same process to occur, and that shorter oligomers and shorter templates might display similar results.

A similar experiment was then performed with these 80 bp strands with different degree of internal structure in the presence of oligomers that are 4 times shorter and complementary to one of the duplex strands. The fraction of bound oligomers would provide insight on whether these sequences remain apart in glycholine for long enough to allow shorter oligomers to diffuse and bind to them. It would also provide information on how the internal structure could affect this binding of the oligomers.

Briefly, oligomers that are 4 times shorter (20 nt) and complementary to one of the duplex strands were added to the mixture, with the 5' oligomer containing a fluorescein tag at the 5' end for ease of binding quantification. Followed by heat cycling in glycholine, the fraction of bound oligomers was analyzed. Similar degree of binding was observed for the unstructured sequence, the sequence that had mismatches in the middle of both stems, and the one containing the toehold, while less binding was observed for the sequences that had previously displayed higher degree of trapping (Figure 11). These results confirm that first, short oligomers diffuse faster than the template's complementary strand as shown by their binding; and second, that intramolecular folding is not necessary for binding of shorter strands to occur in glycholine. In fact, very high internal structure can hinder the binding of oligomers to a template strand, but some degree of structure does not affect this process. Nonetheless, having a sequence with such a high degree of structure is unlikely to have evolved naturally under prebiotic conditions, since naturally occurring ribozymes such as the hairpin ribozyme have less hybridized bases in each stem[47].

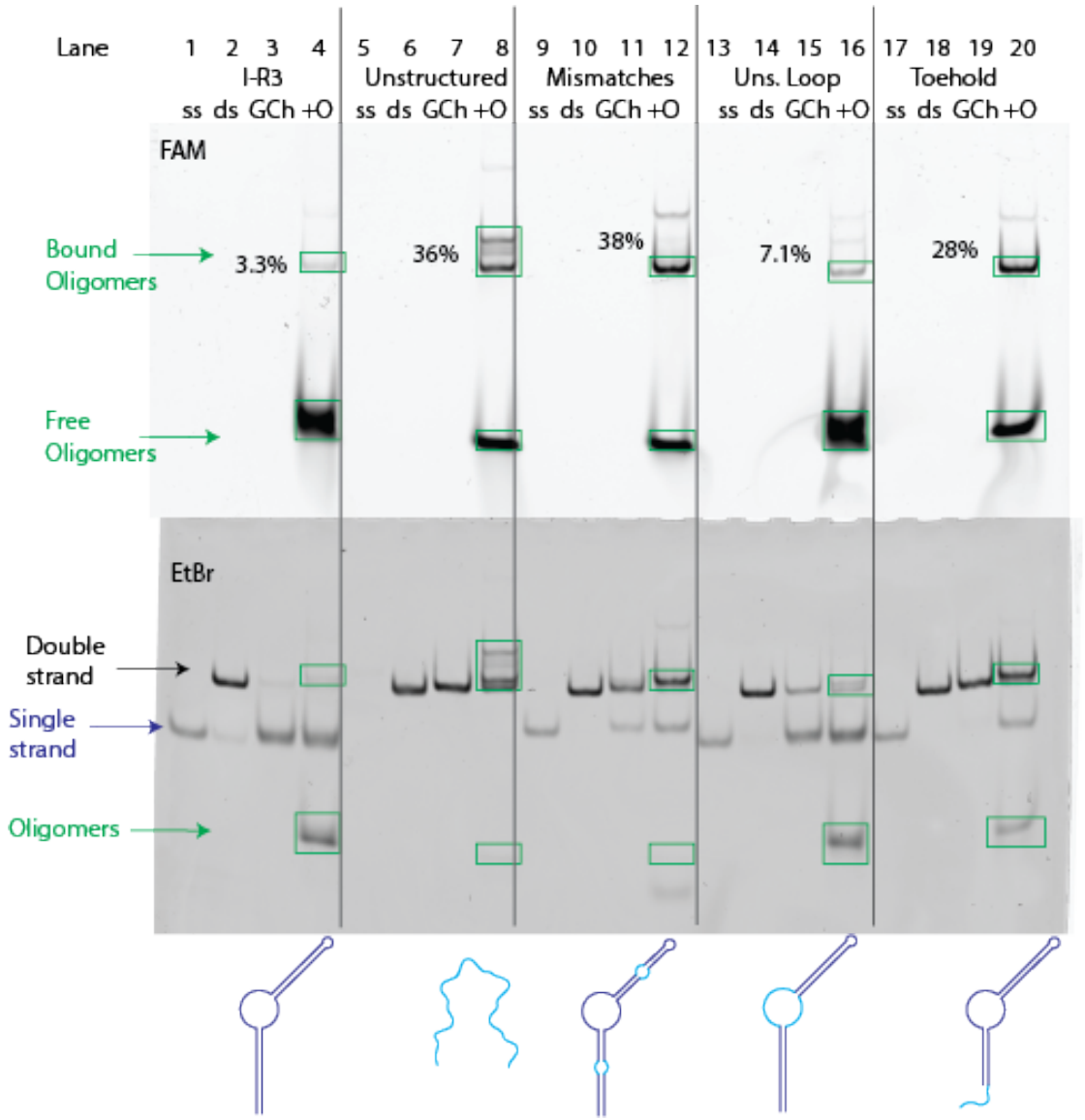


Figure 11. Binding of oligomers to one of the template strands after heat cycling in glycholine. Higher degree of structure results in less binding.

2.6.3 Information transfer from trapped strands

Since there is binding of oligomers to these strands, then information transfer should also be possible as reported by He et al.[11] Therefore, the same protocol from trapping and binding of oligomers was used, followed by an aqueous dilution with addition

of a ligase (oligomers were 5' phosphorylated) to ligate the oligomers that were assembled on the template. Three out the five strands were selected with high, medium, and no internal structure for ease of comparison (I-R3, I-R3 with mismatches and unstructured 80mer), and different cooling rates were also evaluated that ranged from -0.5 °C/s to -0.25 °C/min. Information transfer was observed for all strands, and as seen with the binding, a higher yield of full-length products was observed for the unstructured sequence and I-R3 with mismatches, with higher yields for faster cooling rates (Figure 12). Nonetheless, the main product obtained for I-R3 is the full-length product, with a maximum ligation yield of approximately 1/3 of the theoretical maximum (25% - oligomers are in 4x excess compared to the template) for the fastest cooling rate. When the heat cycling was done in aqueous buffer, very low yields of product were obtained for the fastest cooling rates, corroborating that the viscous solvent is indeed slowing down the diffusion of the templates.

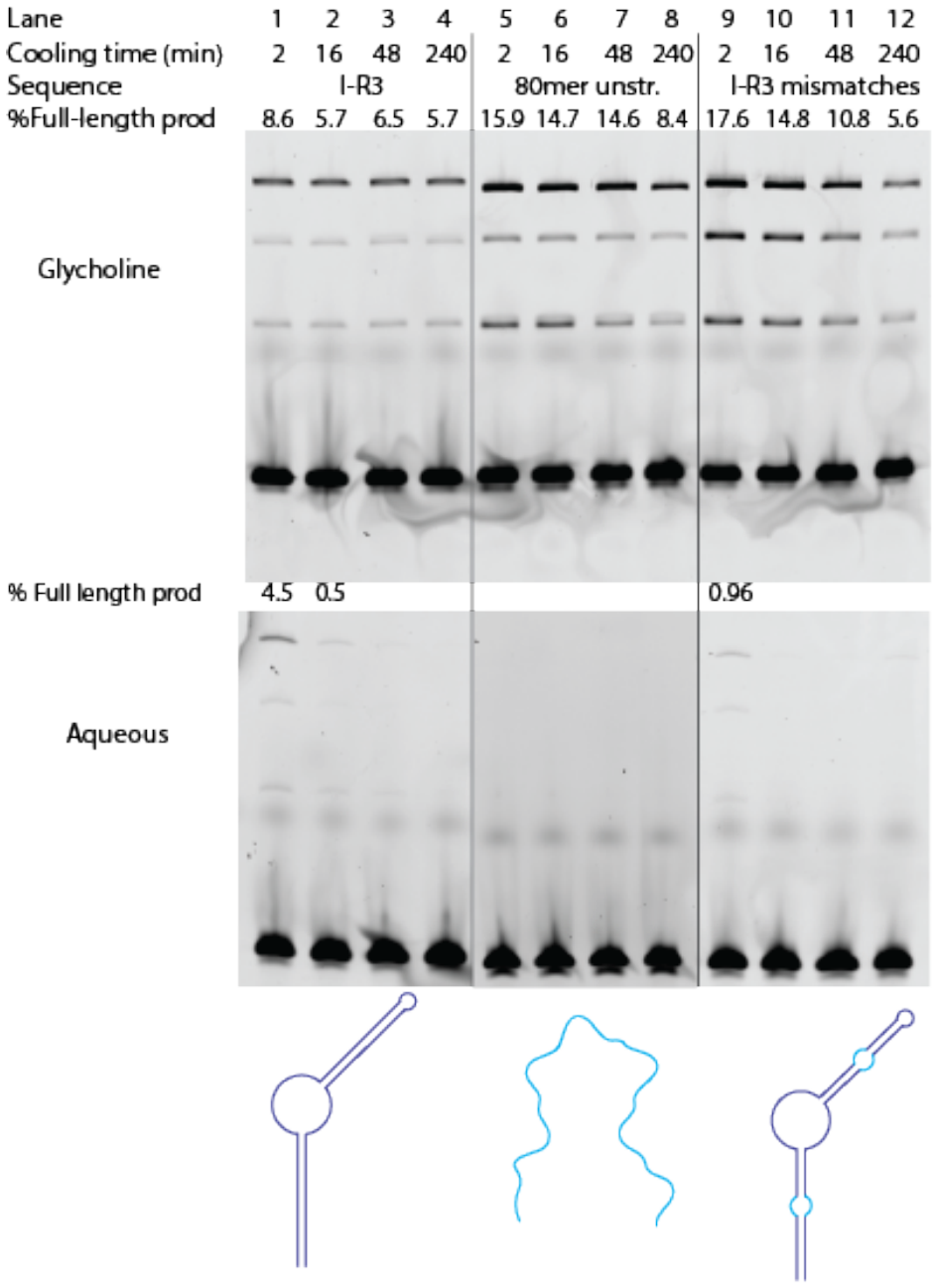


Figure 12. Information transfer from one of the strands from the duplex after heat cycling the duplex in glycholine (top gel) or in aqueous buffer (bottom gel) with different cooling rates. Information transfer takes place in glycholine for all three sequences, with higher yield for the unstructured strand and I-R3 with mistmatches. There is no information transfer of the strands in aqueous buffer unless a very fast cooling rate is used.

2.7 Trapping and catalysis from a hammerhead ribozyme

In the previous sections, it was shown that a self-cleaving DNAzyme can be trapped into its functional conformation and display cleavage activity after dilution in the presence of Zn^{2+}. In addition, if there are complementary strands that are shorter than the template sequences, they can bind to the longer strands and promote information transfer. This suggests that having a trans- cleaving system (strand that cleaves a substrate that is not part of its sequence) should be possible if the substrate is shorter and at least partially complementary to the strand. To further prove that this is possible with more than only the DNAzyme, a cleavage experiment was done with a hammerhead ribozyme after trapping in the presence of its complement and its substrate in glycholine. The RNA used was a 43 nt *trans*-acting hammerhead ribozyme sequence from *Schistosoma mansoni*[48], which has a 21 nt substrate as shown in Figure 13. This ribozyme was chosen because it has been widely studied and characterized[41,48].

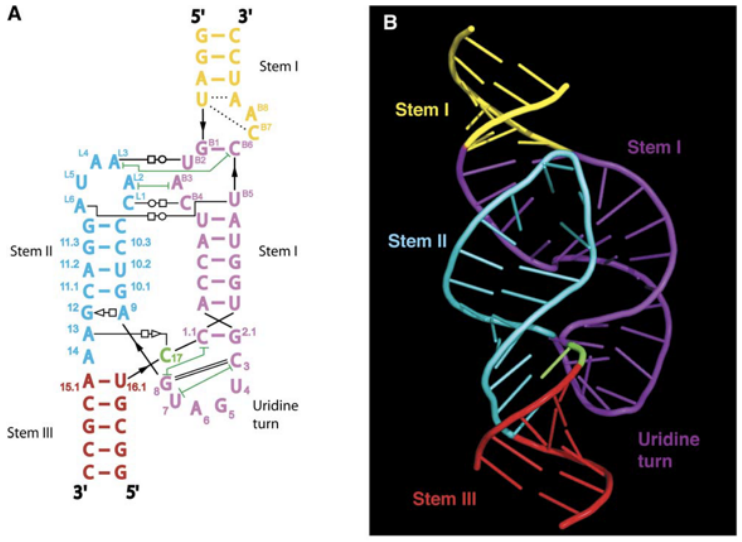


Figure 13. Structure of the full-length hammerhead ribozyme. A) Base pairing interactions. B) Three-dimensional representation. Color matching between A and B. Images were taken from Figure 1 of Martick and Scott, "Tertiary Contacts Distant from the Active Site Prime a Ribozyme for Catalysis" Cell (2006) 126, 309–320[48].

Since this particular ribozyme is much shorter than the DNA strands used in the previous sections and the substrate has less base paring interactions with the ribozyme, it was not possible to show trapping of the hammerhead ribozyme bound to its substrate in a non-denaturing gel. However, an indirect way of showing trapping is by testing the cleavage of the substrate after the heat cycling, the same way it was shown in section 2.5, in which the trapping results were consistent with the cleavage yields. Therefore, the hammerhead ribozyme was heat cycled in glycholine (cooling rate of -0.5°C/s) in presence and absence of its complementary sequence, with equimolar and 4x excess of substrate; after 10x dilution with MES buffer pH 6 and addition of $MgCl_2$ to a final concentration of 10 mM, the kinetics of cleavage were followed by gel electrophoresis (the substrate was labelled with fluorescein for easy quantification of cleavage yield).

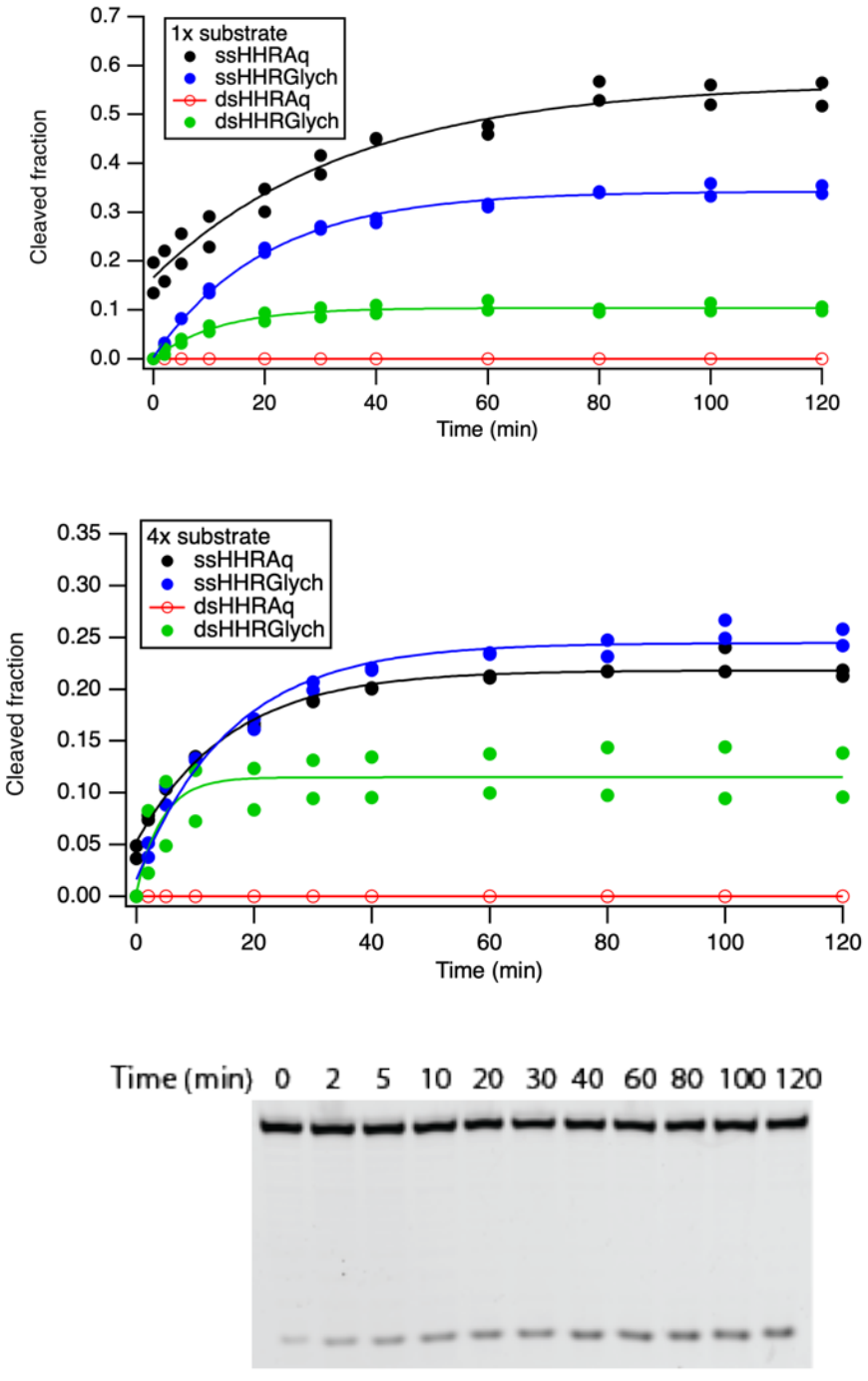


Figure 14. Cleavage kinetics of the hammerhead ribozyme substrate after trapping in glycholine with 1:1 and 4:1 substrate to ribozyme ratio, and sample gel image. The gels from where these data was obtained are in the Appendix.

Overall, the kinetics are not affected in the presence of 10% glycholine after dilution for the single stranded HHR in presence of excess substrate, but it is slightly reduced with glycholine when present in equimolar amounts. Furthermore, there is trapping of the active sequence with its substrate after heat cycling in glycholine in the presence of its complementary sequence. When equimolar amount of the substrate is present, only 1/3 of the sequences were trapped, but if the substrate concentration is increased to 4x the concentration of the HH ribozyme, approximately half of the sequences got trapped (Figure 14). These results demonstrate that shorter active sequences can also be kinetically trapped in glycholine in the presence of their substrate.

2.8 Conclusions

Changes in the environmental conditions are capable of driving a denatured strand towards intramolecular trapping or reannealing with its complement. There are different factors that can affect the degree of trapping. In this case, the viscosity of the medium and fast cooling rates will derive into a higher degree of kinetic trapping of a particular structure, but there might be other factors such as a decrease in the Tm that were not fully explored for this system. Furthermore, these structures can remain trapped for long periods of time even after aqueous dilution, which can then lead them to cleave or perform catalysis.

Moreover, the degree of internal structure of a sequence can significantly affect the degree of trapping of the single strands. Higher structure leads to more trapping, and if there are free bases within the structure, such as toeholds or mismatches, then the trapped

strand can reanneal with its complement through strand displacement to the more thermodynamically favourable duplex.

The degree of internal structure can also affect binding of oligomers and information transfer as has been suggested in previous reports[20,49], nonetheless, information transfer was observed for all three tested sequences with low, medium and high level of intramolecular structure. Furthermore, information transfer can take place with relatively short strands, in a 1:4 size ratio of template to oligomer length.

Finally, the trapping of a sequence in its functional conformation is not unique to a particular sequence, as it was shown not only for the I-R3 DNAzyme, but also for the hammerhead ribozyme. The relevance of trapping strands in their functional conformation is relevant and could provide an easy alternative to the methods currently used to isolate single stranded DNA in artificial selection of nucleic acids[50].

2.9 Materials and methods

2.9.1 Materials

Glycholine was prepared by mixing glycerol (from Sigma, 99%), and choline chloride (Acros Organics, 99%) in a 4:1 molar ratio. Before mixing, choline chloride was recrystallized in ethanol, and both, glycerol and choline chloride were dried under vacuum for >16 hours. The mixture was then heated up to 85 °C until a homogeneous liquid was formed.

Glycholine was kept dry under vacuum centrifugation for at least 12 hours before any experiments were done.

All DNA sequences were purchased from IDT and purified through polyacrylamide gel electrophoresis. Some of the sequences of interest were FAM labeled on the 5' end in order to quantify the trapping by densitometry analysis of the gels.

2.9.2 Trapping of intramolecular structures

Samples were prepared with $6x10^{-9}$ moles of the DNA in 6 mg of glycholine for a final concentration of 1 or 3 μm. The samples were dried under vacuum centrifugation for more than 8 hours. Then they were heated to 80 °C for 5 min and cooled down to room temperature at different rates that vary between -48 and -3.75 °C/min. Samples were then diluted in Tris buffer 20mM (pH 7) to a final concentration of 0.6 μM and loaded in a non-denaturing 12% polyacrylamide gel. The gels were run for 1.5 hours at 150 V. Gels were then imaged on a Typhoon FLA 9500 laser scanner (GE Healthcare) at a resolution of 50 μm. FAM fluorescence was measured using an excitation wavelength of 473 nm and the long pass blue (LPB) emission filter.

2.9.3 Cleaving activity of DNAzyme

The same process used for trapping was followed: Samples were prepared with 6 nano moles of the DNA in 6 mg of glycholine for a final concentration of 1 μm; after drying under vacuum centrifugation for at least 8 hours, they were heated to 80 °C for 5 min and cooled down to room temperature at -48 and °C/min. To provide the best cleaving conditions, samples were diluted 6x in aqueous Tris buffer 20mM (pH 7) with final

concentration of 2 mM $ZnCl_2$. Samples were left to react for 30 min, and then 6 μL of sample were mixed with 6 μL of RNA loading buffer (95% formamide, 0.01% SDS, 0.5 mM EDTA, pH 8). The samples were then heated to 90 °C for two minutes and placed on ice before loading in a 12% denaturing polyacrylamide gel. Gels were pre-run for at least 30 minutes at 14 W and 45 V/cm before loading the samples. After sample loading, gels were run in these conditions for one hour. The gels were then imaged for FAM fluorescence as indicated above.

2.9.4 Cleaving activity of hammerhead ribozyme

Samples were prepared with 6 nano moles of the hammerhead ribozyme and its complement in 6 mg of glycholine for a final concentration of 1 μm, and 4x excess of the substrate. After drying under vacuum centrifugation for at least 8 hours, they were heated to 80 °C for 2 min and cooled down to room temperature at -0.5 and °C/s. To provide the best cleaving conditions, samples were diluted 10x in aqueous MES buffer 50 mM (pH 6) with final concentration of 10 mM $MgCl_2$. Samples were left to react for up to 2 hours, with 5 μL aliquots taken over time. The 5 μL of sample were mixed with 1μL of 0.5 M EDTA and 6 μL of RNA loading buffer (95% formamide, 0.01% SDS, 0.5 mM EDTA, pH 8). The samples were then heated to 90 °C for two minutes and placed on ice before loading in a 22% denaturing polyacrylamide gel. Gels were pre-run for at least 30 minutes at 14 W and 45 V/cm before loading the samples. After sample loading, gels were run in these conditions for 1.5 hours. The gels were then imaged for FAM fluorescence as indicated above.

2.9.5 Sequences

Table 1. Sequences used for comparison of trapping in glycholine.

Name	Sequence (80 nt)
Long I-R3	5'-AGT CAG AAA GAT AAT C-TA GTT GAG CT-G TCT GCA TCA CGA GGA / ACT CGT GAT GCA GAC- GTT G**AA** G-GA TTA TCT TTC TGA CT-3'
L I-R3 complement	5'-AGT CAG AAA GAT AAT CCT TCA ACG TCT GCA TCA CGA GTT / CCT CGT GAT GCA GAC AGC TCA ACT AGA TTA TCT TTC TGA CT -3'
Oligomers	I-R3_Oligo1: /56-FAM/ AGT CAG AAA GAT AAT CTA GT I-R3_Oligo2: /5Phos/ TGA GCT GTC TGC ATC ACG AG I-R3_Oligo3: /5Phos/ GAA CTC GTG ATG CAG ACG TT I-R3_Oligo4: /5Phos/ GAA GGA TTA TCT TTC TGA CT
80mer_Unstructured	5'-CGT GTC AGT AAG CCA TTC GTG TTC CTC ATA GTC GTC TCA CCA GAT CCT CAT AGT CGT CTC ACC AAA TCC TCA TAG TCG TC-3'
80mer_Comp_Unst	5'-GAC GAC TAT GAG GAT TTG GTG AGA CGA CTA TGA GGA TCT GGT GAG ACG ACT ATG AGG AAC ACG AAT GGC TTA CTG ACA CG-3'
Oligomers	80mer_S1: /56-FAM/ CGT GTC AGT AAG CCA TTC GT 80mer_S2: /5Phos/ GTT CCT CAT AGT CGT CTC AC 80mer_S3: /5Phos/ CAG ATC CTC ATA GTC GTC TC 80mer_S4: /5Phos/ ACC AAA TCC TCA TAG TCG TC
Long I-R3 with mismatches	5'-AGT CAG AAA GAT AAT C-TA GTT GAG CT-G TCT GCA TCA CGA GGA / ACT CGT AGGA CA GAC- GTT G**AA** G-GA TTA TGA CAC TGA CT-3'
Complement to L I-R3 with mismatches	5'- AGT CAG TGT CAT AAT CCT TCA ACG TCT GTC CTA CGA GTT CCT CGT GAT GCA GAC AGC TCA ACT AGA TTA TCT TTC TGA CT -3'
Oligomers	I-R3_Oligo1: /56-FAM/ AGT CAG AAA GAT AAT CTA GT I-R3_Oligo2: /5Phos/ TGA GCT GTC TGC ATC ACG AG I-R3_Mod_Oligo3: /5Phos/GAA CTC GTA GGA CAG ACG TT I-R3_Mod_Oligo4: /5Phos/GAA GGA TTA TG CAC TGA CT

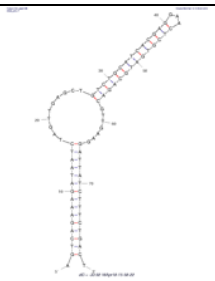
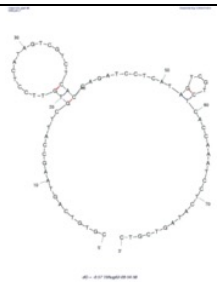
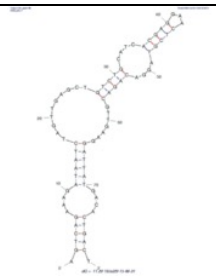
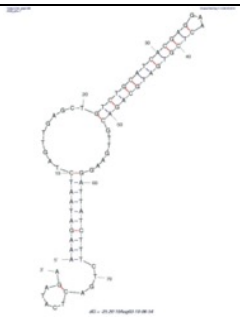
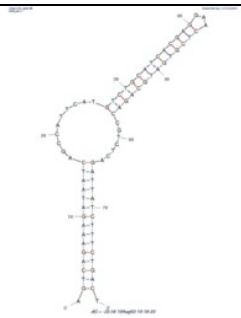
Table 1 (continued). Sequences used for comparison of trapping in glycholine.

Long I-R3 with toehold	5'- AAA GAT AAT C-TA GTT GAG CT-G TCT GCA TCA CGA GGA / ACT CGT GAT GCA GAC- GTT G**AA** G-GA TTA TCT TTC TGA CT CATAGA-3'
Complement to L I-R3 with toehold	5'- TCT ATG AGT CAG AAA GAT AAT CCT TCA ACG TCT GCA TCA CGA GTT CCT CGT GAT GCA GAC AGC TCA ACT AGA TTA TCT TT -3'
Oligomers	TH1: /56-FAM/AAA GAT AAT CTA GTT GAG CT TH2: /5Phos/GTC TGC ATC ACG AGG AAC TC TH3: /5Phos/GTG ATG CAG ACG TTG AAG GA TH4: /5Phos/TTA TCT TTC TGA CTC ATA GA
Structured sequence with unstructured loop	5'-AGT CAG AAA GAT AAT C- AGCCATTCAT-G TCT GCA TCA CGA GGA / ACT CGT GAT GCA GAC- CGTCTCA-GA TTA TCT TTC TGA CT-3'
Complement to Structured sequence with unstructured loop	5'- AGT CAG AAA GAT AAT CTG AGA CGG TCT GCA TCA CGA GTT CCT CGT GAT GCA GAC ATG AAT GGC TGA TTA TCT TTC TGA CT -3'
Oligomers	SS-UL_1: /56-FAM/ AGT CAG AAA GAT AAT CAG CC SS-UL_2: /5Phos/ATT CAT GTC TGC ATC ACG AG SS-UL_3: /5Phos/GAA CTC GTG ATG CAG ACC GT SS-UL_4: /5Phos/CTC AGA TTA TCT TTC TGA CT

2.9.6 Calculation of thermodynamic values and secondary structure of the sequences used.

Thermodynamic parameters for the different designed structures were estimated through mfold[45], with 0.1 M NaCl and at 20°C. Secondary structures and values are indicated in Table 2.

Table 2. Secondary structure and Thermodynamic parameters at 20 °C, 0.1 M NaCl, calculated through mfold

Name	Long I-R3	80mer_Unstructured	Long I-R3 with mismatches
Predicted secondary structure			
Thermodynamic parameters, assuming a 2-state model	ΔG = -33.92 kcal/mol ΔH = -240.40 kcal/mol ΔS = -704.3 cal/(K·mol) Tm = 68.1 °C	ΔG = -0.57 kcal/mol ΔH = -48.20 kcal/mol ΔS = -162.4 cal/(K·mol) Tm = 23.5 °C	ΔG = -11.20 kcal/mol ΔH = -171.10 kcal/mol ΔS = -545.4 cal/(K·mol) Tm = 40.5 °C
Name	**Long I-R3 with toehold**	**Structured sequence with unstructured loop**	
Predicted secondary structure			
Thermodynamic parameters, assuming a 2-state model	ΔG = -25.20 kcal/mol ΔH = -202.60 kcal/mol ΔS = -605.1 cal/(K·mol) Tm = 61.6 °C	ΔG = -33.54 kcal/mol ΔH = -236.80 kcal/mol ΔS = -693.3 cal/(K·mol) Tm = 68.3 °C	

CHAPTER 3. MODEL REPLICATION CYCLE OF AN RNA DUPLEX CONTAINING A RIBOZYME MOTIF IN A VISCOUS SOLVENT*

3.1 Introduction

There are some challenges associated with RNA replication under prebiotic conditions. One of these is the strand inhibition problem, where the reannealing of the two template strands is favored and prevents binding of mono- or oligonucleotides to generate copies of the template [12]. Second, is the "replicator-catalyst" paradox, which states that the structure needed for RNA catalysis, prevents the sequence to serve as a template for replication in water [15,20]. Third, also related to strand inhibition, is the transition between the replication duplex and the active state of a ribozyme[49].

Previous work by He, Gallego et al. [11] proposed a way to overcome the strand inhibition problem using glycholine. They demonstrated information transfer from one strand of a 3 kb and a 545 DNA duplex was possible with this system. Building upon this work, we decided to explore the possibility of simultaneously copying both strands of a 613 bp RNA duplex to have a full replication cycle (steps 1-5 as shown in Figure 15).

* This project was started by Christine He, and finished by me. It resulted in a joint first author publication: He, C., Lozoya, A., Gallego, I., Grover, M.A. and Hud, N. (2019) Solvent viscosity facilitates replication and ribozymecatalysis from an RNA duplex in a model prebiotic process. Nucleic Acids Research, 47, 6569–6577.

In a similar fashion, oligomers 32 nt long were used, which are within length range reported for oligomers synthesized on mineral surfaces[51]. Furthermore, oligomers could have promoted unfolding[19] of templates and could have assembled on the templates in a variety of conditions. In this new system, however, there was the risk of having a variation of the strand inhibition problem because the oligomers that would bind to the templates to form the new strands would also be complementary to each other. Therefore, a competition between the oligomers binding to each other and binding to the template could emerge. Furthermore, based upon the results shown in the previous chapter, where active strands were kinetically trapped and proved to be active; the hammerhead ribozyme sequence was inserted in the middle of one of the template strands to demonstrate its activity after kinetic trapping in glycholine.

We propose 6 steps of a replication cycle enabled by a viscous solvent as displayed in Figure 15: 1) separation of the two strands of the duplex by high temperatures; 2) trapping of the template single strands by formation of intramolecular structures in a viscous environment; 3) assembly oligonucleotides to the template strand can start to unfold the intramolecular structure; 4) complete coating of the template by oligonucleotides; 5) ligation of the assembled oligomers to form a new copy of the template strand, 6) start of a new replication round. Step 3* leads to the trapping of the active form of the ribozyme in a similar way it was demonstrated in the previous chapter.

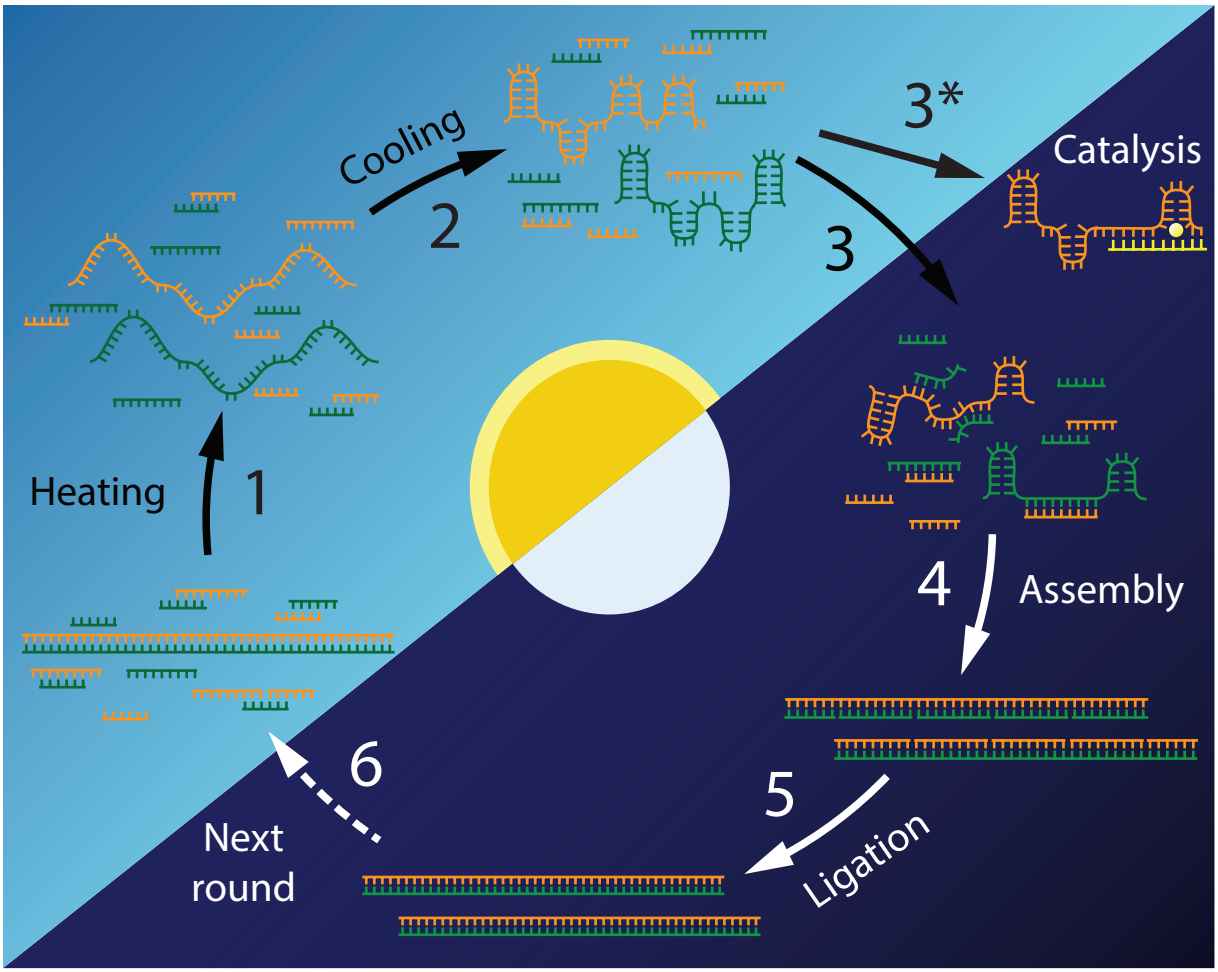


Figure 15. Cycle for viscosity-mediated replication of duplex RNA. In a viscous solvent thermal cycling of an RNA duplex (Steps 1-2) leads to kinetic trapping of the RNA in a folded, single stranded state. Shorter, more mobile oligonucleotides assemble on the template strands (Steps 3 and 4). Ligation of the assembled substrates (Step 5) completes the replication cycle, and another round of replication can begin (Step 6). The kinetic trapping of intramolecular structure also allows catalytic RNA sequences to access their active forms (Step 3*), which are inaccessible in the duplex state that follows replication. (Figure design by Christine He, taken with authors permission from He, C.; Lozoya-Colinas, A. et al., 2019 [52])

3.2 Assembly of oligomers to the template in glycholine

First, we demonstrated that glycholine enables the binding of the oligomer substrates to the single strands of the RNA template, which corresponds to steps 1-4 in Figure 15. To this end, a 613 bp ds RNA duplex was heated to 80 °C for 2 min and cooled to room temperature at a rate of -3.75 °C/min in presence of its complementary oligomers in glycholine. A total of 10 oligomers for each strand were used (S1 to S10, and S1' to S10'), each was 32 nucleotides long, covering a 320 bp fragment from the template. One of the oligomers from each strand was labeled with a fluorophore (^{Cy5}S4 and ^{Cy3}S4') to monitor the binding of the oligomers to their templates through an agarose gel mobility shift assay (Figure 16). The EtBr image indicates that the templates were trapped as single strands after heating and cooling in glycholine, and the binding of the oligomers can be followed by the Cy3 and Cy5 image. Quantification of the observed bands suggests that there is a higher percentage of the oligomers bound to the antisense template (Figure 16), consistent with the sense template having a more stable intramolecular structure. This demonstrates that even though there is a stable intramolecular structure on both templates, binding of oligomers still occurs in glycholine. There is no trapping or binding observed in aqueous solution or without heating and cooling in glycholine.

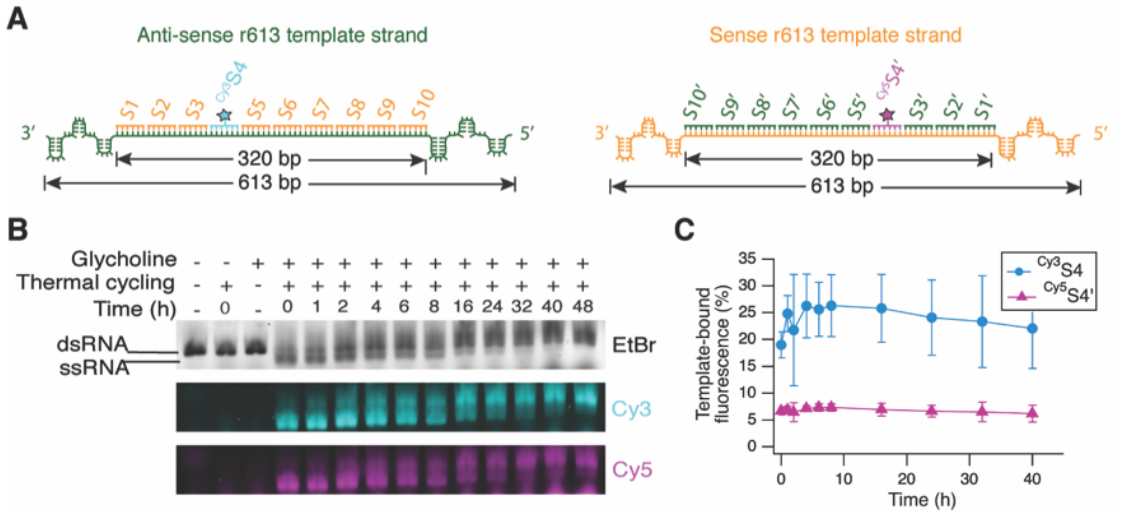


Figure 16. Solvent viscosity enables oligonucleotide assembly on both strands of an RNA duplex. (A) RNA oligonucleotides S1-S10 and S1'-S10' were designed to coat a 320 bp region of the r613 template duplex. Oligonucleotides ^{Cy3}S4 and ^{Cy5}S4' are fluorescent markers for template binding. (B) Agarose gel showing the kinetics of oligonucleotide binding to the RNA template. The Cy3 and Cy5 fluorescence images indicate that ^{Cy3}S4 and ^{Cy5}S4' bind to both the ss and dsRNA templates after heating and cooling in glycholine, but not in water. (C) Kinetics of Cy3S4 and Cy5S4' binding to the r613 template.

3.3 Information transfer from both strands of the r613 template

3.3.1 Information transfer from both strands of the r613 template with 32 nt oligomer substrates

After demonstrating the assembly of the oligomers to the template strand, we proceeded to covalently link the oligomers that were bound to the templates, corresponding to step 5 in Figure 15. T4 RNA ligase 2 was used as a robust and efficient way of ligation. It is important to note that the enzyme was not added until the last step of the cycle. The same template and oligomers used to see the trapping and assembly were used, with ^{Cy3}S4 as a marker for synthesis of the sense strand, and ^{Cy5}S4' as a marker for the synthesis of the

anti-sense strand. All possible products are shown in lanes 5-9 of Figure 17A, showing that both strands can be simultaneously copied. These products are not observed if there is no heating and cooling, or if the same procedure is followed in aqueous solution. The intensity of the product bands is lower when both sets of oligomers are present, this is consistent with the oligomers being self-complementary and binding to each other in addition to binding to the template, resulting in a lower effective concentration of free oligomers which are competing between binding with themselves and binding to the template.

Furthermore, we carried out the same experiment with shorter templates that range between 96 and 192 bp, which are closer in length to known ribozymes. All possible products are again observed in Figure 17B, proving that full templates with a wide range of lengths can be replicated by this system. These results also demonstrate that it is not necessary to have overhangs at the end of the template for replication to take place, and the size ratio of template to oligomer can go down to 1/3 and still undergo replication.

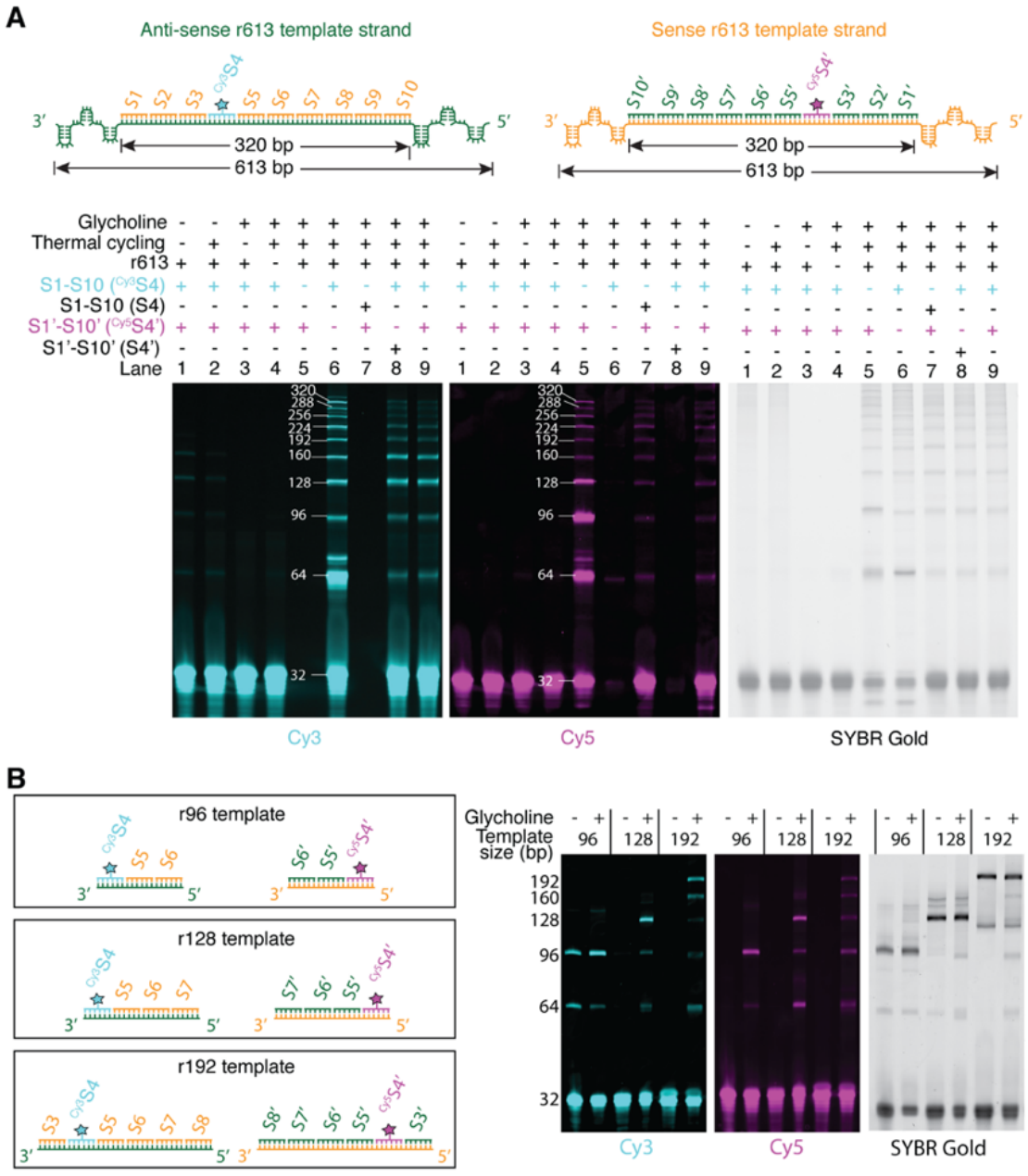


Figure 17. Replication of RNA duplex templates in glycholine. (A) Samples containing the r613 template duplex and oligonucleotides S1-S10 and S1'-S10' were thermal cycled and then ligated. The Cy3 and Cy5 images track synthesis of the RNA sense and anti-sense strands, respectively. In glycholine, copying from both strands of the r613 duplex is observed. (B) Glycholine enables replication from a range of shorter RNA template duplexes (96-192 bp).

3.3.2 Information transfer from both strands of the r613 template in the presence of random oligomers

Simultaneous information transfer from both strands (replication) of an RNA duplex is possible with the aid of a viscous solvent. We used 2 sets of 10 oligomer substrates (32 nt each) that were fully complementary to the template. However, under prebiotic conditions, more sequences were likely present in addition to the oligomers that are complementary to the template. Therefore, we also performed a replication experiment in the presence of random oligomers with increasing concentrations in presence and absence of the complementary oligomers (Figure 18).

Replication is observed in the presence of random oligomers up to 64x the concentration of the template (lanes 2-8), while there is no ligation of the random pool of oligomers if the complementary oligomers are not present (lane 9). This further demonstrates that there is template directed synthesis of the new strands, and that the presence of random oligomer substrates does not interfere with the assembly and ligation process needed for replication.

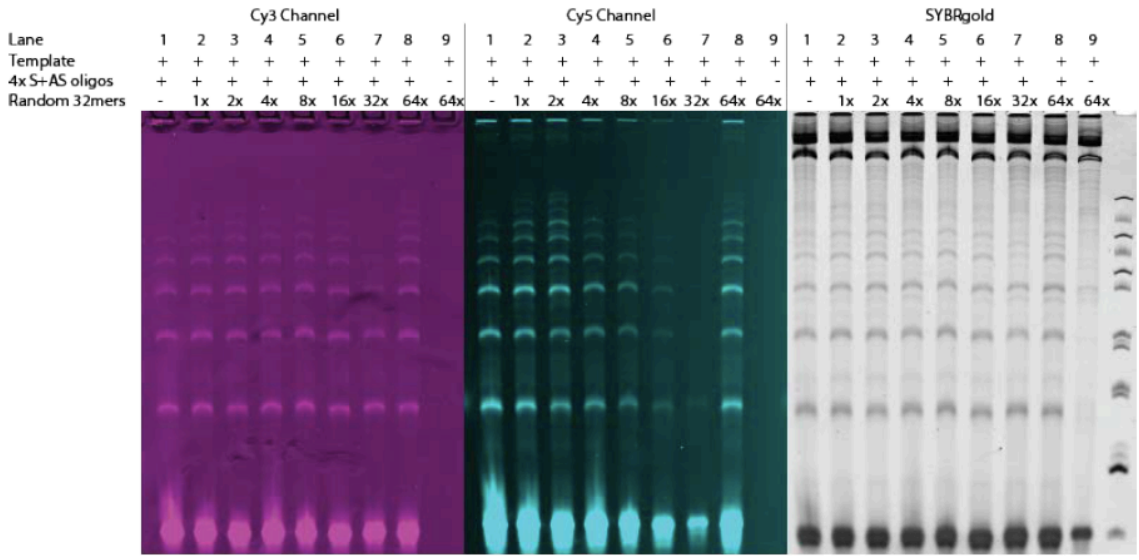


Figure 18. Replication of an RNA duplex in the presence of random oligomer substrates in addition to the complementary S1-S10 and S1'-S10' substrates.

3.3.3 Replication enabled by other viscous solvents

He and Gallego demonstrated that information transfer was dependent on the viscosity of the solvent: higher viscosity led to higher yields, with highest yield achieved in pure glycholine which has a viscosity of 437 cP[11], To further demonstrate that the replication is indeed dependent on the solvent viscosity and not the nature of the solvent itself, replication following the same protocol was followed in different viscous solvents: glycerol, which has a viscosity of approximately 1400 cP[53,54]; and reline, a mix of urea and choline chloride in a 2:1 molar ratio, which has a viscosity of above 1200 cP at room temperature[55,56].

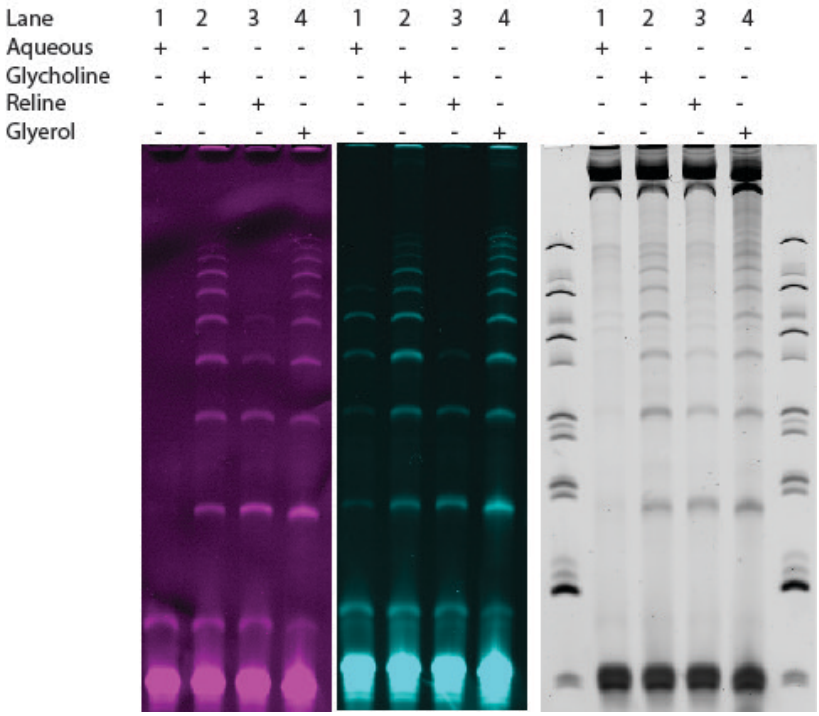


Figure 19. Replication of the r613 template duplex in glycholine, reline (2:1 mixture of urea and choline chloride) and glycerol.

As shown in Figure 19, rreplication was observed in all three viscous solvents that were tested. The yield in reline is lower compared to the glycerol containing solvents. This could be due to the high urea concentration, a known denaturing agent for nucleic acids and proteins[57,58]. Urea might affect the hybridization of the oligomer substrates to the template and could also affect the activity of the ligase used. Nonetheless, ligation products from both strands were observed in all three cases, suggesting that this process is transferable to different solvents that could be formed of prebiotic compounds.

3.4 Substrate cleavage by HH ribozyme in r613 template and newly synthesized strands

Lastly, we evaluated the catalytic activity of the hammerhead ribozyme contained within the template and in the ligated products obtained after one replication round. The same heating and cooling cycle protocol previously used for replication was followed with the 613bp RNA template (or the ligated products) and the HH ribozyme substrate present in 4x excess as compared to the HH ribozyme-containing sequence.

3.4.1 Cleavage by the r613 template containing the ribozyme sequence

As previously mentioned, the HH ribozyme sequence was inserted in the middle of the template used for replication. Since the strand is much larger than the ribozyme sequence (613 bp as compared to 43 bp for the HH ribozyme), it was important to verify that the template could still fold in its active conformation and be active within the longer strand. The strands were heat cycled in glycholine, followed by a 1/10 aqueous dilution and incubation at 20 °C for 1 hour in 2 mM $MgCl_2$. The cleaved product is observed with a 20% yield as compared to the minimal hammerhead ribozyme used as control (Figure 20). This reduced yield is probably due to multiple possible secondary and tertiary structures formed within the 613bp template that prevent binding of the substrate to the active site. No cleaving activity was observed in aqueous buffer, without thermal cycling or in absence of magnesium.

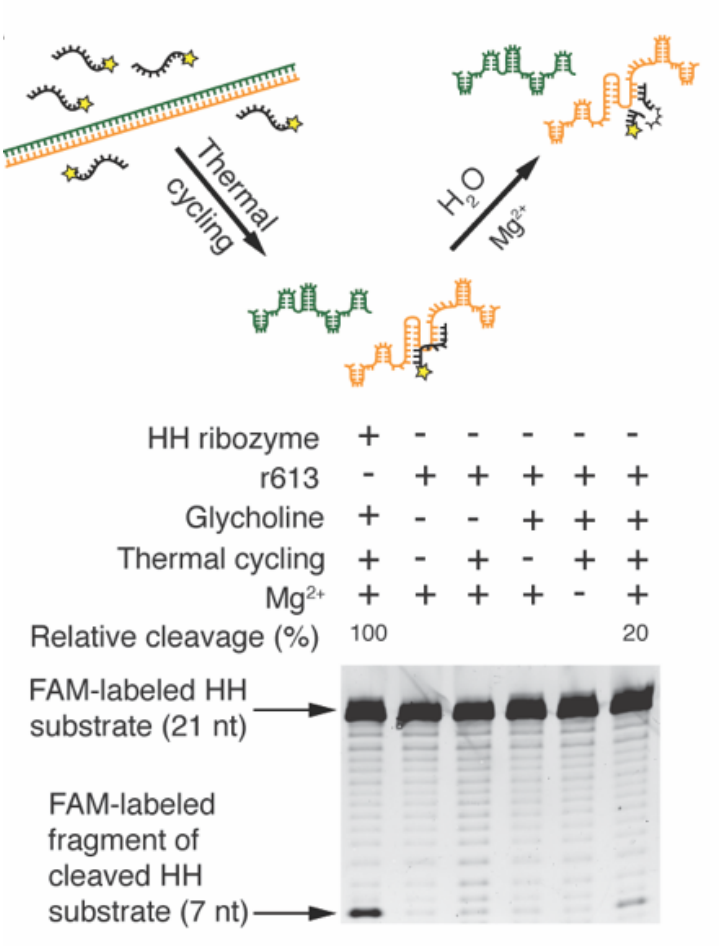


Figure 20. HH substrate cleavage from a kinetically trapped RNA duplex. Schematic illustrating cleavage of the HH substrate (black) by the RNA template strand containing the HH ribozyme (orange) after thermal cycling in glycholine. A denaturing polyacrylamide gel shows that an RNA duplex containing the HH ribozyme can cleave its substrate.

3.4.2 Cleavage by the newly synthesized strands containing the HH ribozyme sequence

Finally, after showing that ligation of the oligomers can produce full-length products, is was important to prove that these newly formed strands were properly synthesized and contained the HH ribozyme sequence. This could be done by demonstrating substrate cleavage after the replication process. As a first approach, only the

antisense RNA strand was used as template to produce the strand containing the HH ribozyme using the sense oligomers. As seen in Figure 21, the sense strand was formed through template-directed synthesis from the antisense strand, and the newly formed product was capable of cleaving the HH substrate, corroborating that the sequence formed contained the active ribozyme.

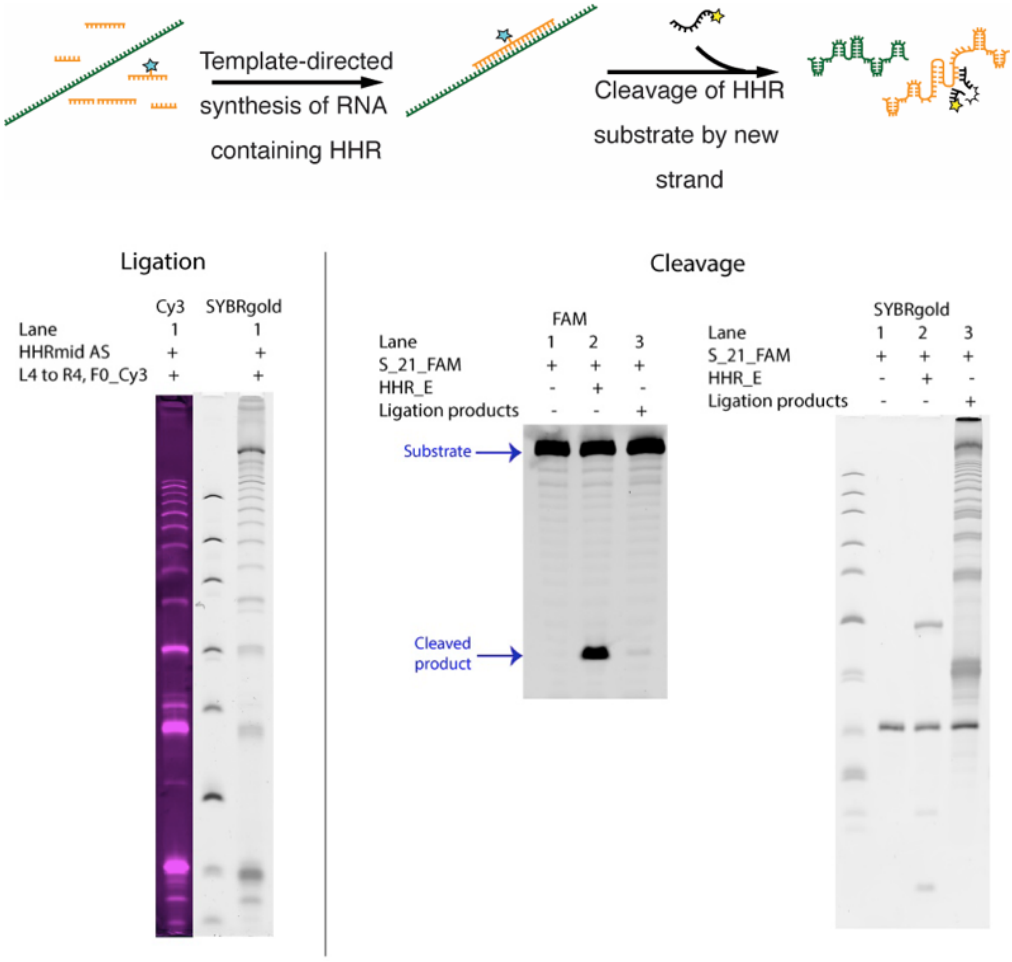


Figure 21. Hammerhead cleavage by the sense r613 strand synthesized from the anti-sense r613 strand using oligonucleotide substrates S1-S10 in glycholine. PAGE analysis confirms that the newly synthesized sense strand exhibits hammerhead cleavage.

Then, a full replication round was done (both sense and antisense strands synthesized from the template duplex), in which the ligated products ranging from 192 to 320 bases of 90 ligation reactions were PAGE purified and their cleaving activity was evaluated in the same conditions as the template. The product band observed in Figure 22 corroborates that the ligated products were properly synthesized and can also be trapped into their active conformation, giving a cleavage yield of 9% as compared to the minimal HH ribozyme. It is important to note that in this case, the concentration of the strands containing the ribozyme is just an estimate since there are strands of different sizes, and the ligation yield was not the same for the sense and the antisense strands.

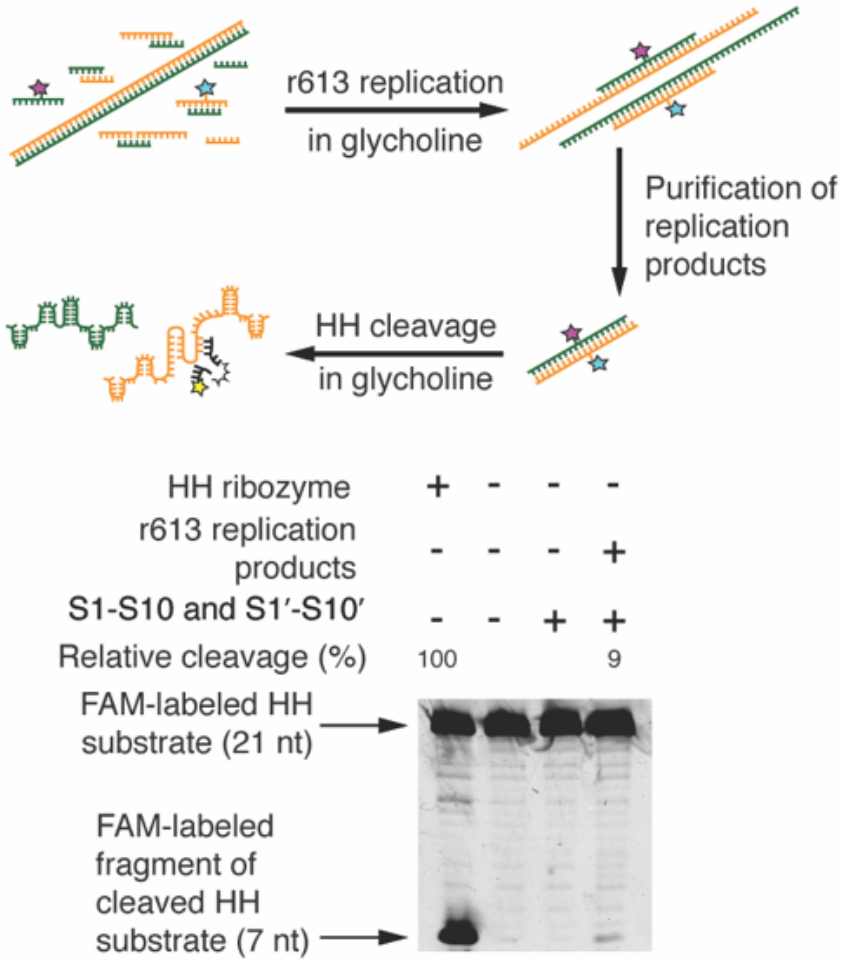


Figure 22. Cleavage of the HH substrate is observed from products of viscosity-mediated replication with r613 and oligonucleotides S1-S10 and S1'-S10'.

3.5 Conclusions

All these results demonstrate that a viscous solvent such as glycholine can favor the replication of templates that vary in size (100-600 bp) and intramolecular structure. The proposed system not only overcomes the strand-inhibition problem, but it can also be used to trap the templates and the ligated products in their active form, promoting catalysis. This

process likely occurs by slowing the diffusion of long templates, enabling the short oligomers to diffuse faster and bind to the templates, even in the presence of fully complementary oligomers.

We have verified that long and highly structured sequences can be replicated at the same time as their complement. Thus, the use of such a viscous solvent and oligomers instead of monomers could help to solve the "replicator-catalyst paradox" that has prevented the replication of highly folded structures[20].

Glycholine was used as a model solvent due to its high viscosity and miscibility with water, and even though glycerol could be prebiotically formed[59], there is no record of a prebiotic synthesis for choline chloride. Nonetheless, as replication was demonstrated with different solvents, this process is likely not unique to glycholine, and might have occurred under prebiotic conditions. There are a wide variety of prebiotic organic compounds and salts which could have formed mixtures that provided the adequate viscosity and environmental conditions through constant water evaporation and hydration in a pond or puddle for replication to take place.

Overall, one full round of replication from an RNA duplex was achieved by using glycholine as a model solvent. This environment also allowed for trapping of the strands into their active conformation as shown by the cleavage of the HH substrate.

3.6 Materials and methods

3.6.1 Materials

Glycholine was produced as stated in section 2.9.1.

DNA was ordered from IDT, RNA was ordered from IDT or Dharmacon, and the sequences were purified through PAGE unless they had been HPLC purified when ordered and looked clean in a polyacrylamide gel.

3.6.2 Production of r613HHR template

DNA templates were ordered from IDT as gBlocks, the sequences were taken from the plasmid pBluescript II SK(-), a phagemid excised from lambda ZAPII (GenBank accession: X52330.1). This plasmid was chosen as a model of a mixed, arbitrary sequence (i.e., not designed to have a particular sequence or base pair composition). A 43 nt *trans*-acting hammerhead ribozyme enzyme sequence from *Schistosoma mansoni*[48] was placed into the middle of the RNA template sequences. This is the same ribozyme used in the previous chapter for where trapping and cleavage was evaluated.

PCR of the DNA templates was performed using NEB Q5 High Fidelity DNA polymerase using two sets of primers (IVT_S_FWD and IVT_S_REV, and IVT_A_FWD and IVT_A_REV) to produce separate dsDNA templates for in vitro transcription of sense and anti-sense RNA template strands. After PCR, amplicons were purified by PAGE in native conditions (0.5x TBE, 7.5 % polyacrylamide).

Double stranded RNA templates were then produced by co-transcription of the sense and anti-sense DNA strands in the same one-pot reaction. 750 ng of each, sense and anti-sense strands, were mixed and transcribed with the HiScribe T7 High Yield RNA Synthesis Kit from New England Biolabs at 37 °C overnight. Then the DNA templates were digested with Turbo DNase (ThermoFisher Scientific) at 20 °C for 15 min. The produced ds RNA templates were purified by PAGE in native conditions (0.5x TBE, 7.5 % polyacrylamide).

3.6.3 DNA and RNA sequences used

Table 3. gBlock (ordered from IDT)

IVT_gBlock	GATCGATCTCGCCCGCGAAATAATACGACTCACTATAGGGGAATAGTGTATG CGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCAC ATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAA ACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTG CACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGAGTTGCTCT TGCAGGTACATCCAGCTGATGAGTCCCAAATAGGACGAAATGCGCGTCCGAG CTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGG GCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAG CATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGA AAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGA CGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGC GTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCC TTCCTTTCTCGCCACGTTCGCG

Table 4. Primers used for PCR reactions (ordered from IDT)

IVT_r96_S_FWD	TAATACGACTCACTATAGGCTTCAGCATCTTTTACTTTC
IVT_r96_S_REV	AGCTCGGACGCGCATTTCGT
IVT_r96_A_FWD	TAATACGACTCACTATAGGAGCTCGGACGCGCATTTCGT
IVT_r96_A_REV	CTTCAGCATCTTTTACTTTC
IVT_r128_S_FWD	TAATACGACTCACTATAGGCTTCAGCATCTTTTACTTTC
IVT_r128_S_REV	GCGGCATTTTGCCTTCCTGT
IVT_r128_A_FWD	TAATACGACTCACTATAGGGCGGCATTTTGCCTTCCTGT
IVT_r128_A_REV	CTTCAGCATCTTTTACTTTC
IVT_r192_S_FWD	TAATACGACTCACTATAGGTTCGATGTAACCCACTCGTG
IVT_r192_S_REV	AACATTTCCGTGTCGCCCTTA
IVT_r192_A_FWD	TAATACGACTCACTATAGGAACATTTCCGTGTCGCCCT
IVT_r192_A_REV	TTCGATGTAACCCACTCGTG
IVT_r613_S_FWD	GATCGATCTCGCCCGCGAAA
IVT_r613_S_REV	CGCGAACGTGGCGAGAAAGGA
IVT_r613_A_FWD	GATCGATCTCGCCCGCGAAA
IVT_r613_A_REV	CGAATAGTGTATGCGGCGACC

Table 5. Oligomers used for binding and ligation assays (ordered from Dharmacon)

S1	/5Phos/ UCAUCAUUGGAAAACGUUCUUCGGGGCGAAAA
S2	/5Phos/ CUCUCAAGGAUCUUACCGCUGUUGAGAUCCAG
S3	/5Phos/ UUCGAUGUAACCCACUCGUGCACCCAACUGAU
^{Cy3}S4	/5Phos/ CUUCAGCAUCU[Cy3]UUUACUUUCACCAGCGUUUCU
S5	/5Phos/ GAGUUGCUCUUGCAGGUACAUCCAGCUGAUG
S6	/5Phos/ GUCCCAAAUAGGACGAAAUGCGCGUCCGAGCU
S7	/5Phos/ GGGUGAGCAAAAACAGGAAGGCAAAAUGCCGC
S8	/5Phos/ AAAAAAGGGAAUAAGGGCGACACGGAAAUGUU

Table 5 (Continued). Oligomers used for binding and ligation assays.

S9	/5Phos/ GAAUACUCAUACUCUUCCUUUUUCAAUAUUAU
S10	/5Phos/ UGAAGCAUUUAUCAGGGUUAUUGUCUCAUGAG
S1’	/5Phos/ UUUUCGCCCCGAAGAACGUUUUCCAAUGAUGA
S2’	/5Phos/ CUGGAUCUCAACAGCGGUAAGAUCCUUGAGAG
S3’	/5Phos/ AUCAGUUGGGUGCACGAGUGGGUUACAUCGAA
^{Cy5}S4’	/5Phos/ AGAAACGCUGGUGAAAGU[Cy5]AAAAGAUGCUGAAG
S5’	/5Phos/ UCAUCAGCUGGAUGUACCUGCAAGAGCAACUC
S6’	/5Phos/ AGCUCGGACGCGCAUUUCGUCCUAUUUGGGAC
S7’	/5Phos/ GCGGCAUUUUGCCUUCCUGUUUUUGCUCACCC
S8’	/5Phos/ AACAUUUCCGUGUCGCCCUUAUUCCCUUUUUU
S9’	/5Phos/ AUAAUAUUGAAAAAGGAAGAGUAUGAGUAUUC
S10’	/5Phos/ CUCAUGAGACAAUAACCCUGAUAAAUGCUUCA

Sequence of hammerhead ribozyme, taken from Martick and Scott[48].

Table 6. Ribozyme and substrate used (ordered from IDT)

HH ribozyme enzyme	GCAGGUACAUCCAGCUGAUGAGUCCCAAAUAGGACGAAAUGCC
HH ribozyme substrate	/56-FAM/GGCAUCCUGGAUUCCACUGC

RNA template sequences used for this project.

Table 7. RNA template sequences (HHR sequence is underlined)

r96 template (sense strand)	CUUCA GCAUC UUUUA CUUUC ACCAG CGUUU CUGAG UUGCU CUUGC AGGUA CAUCC AGCUG AUGAG UCCCA AAUAG GACGA AAUGC GCGUC CGAGC U
r128 template (sense strand)	CUUCA GCAUC UUUUA CUUUC ACCAG CGUUU CUGAG UUGCU CUUGC AGGUA CAUCC AGCUG AUGAG UCCCA AAUAG GACGA AAUGC GCGUC CGAGC UGGGU GAGCA AAAAC AGGAA GGCAA AAUGC CGC
r192 template (sense strand)	UUCGA UGUAA CCCAC UCGUG CACCC AACUG AUCUU CAGCA UCUUU UACUU UCACC AGCGU UUCUG AGUUG CUCUU GCAGG UACAU CCAGC UGAUG AGUCC CAAAU AGGAC GAAAU GCGCG UCCGA GCUGG GUGAG CAAAA ACAGG AAGGC AAAAU GCCGC AAAAA AGGGA AUAAG GGCGA CACGG AAAUG UU
r613 template (sense strand)	GGGGA AUAGU GUAUG CGGCG ACCGA GUUGC UCUUG CCCGG CGUCA AUACG GGAUA AUACC GCGCC ACAUA GCAGA ACUUU AAAAG UGCUC AUCAU UGGAA AACGU UCUUC GGGGC GAAAA CUCUC AAGGA UCUUA CCGCU GUUGA GAUCC AGUUC GAUGU AACCC ACUCG UGCAC CCAAC UGAUC UUCAG CAUCU UUUAC UUUCA CCAGC GUUUC UGAGU UGCUC UUGCA GGUAC AUCCA GCUGA UGAGU CCCAA AUAGG ACGAA AUGCG CGUCC GAGCU GGGUG AGCAA AAACA GGAAG GCAAA AUGCC GCAAA AAAGG GAAUA AGGGC GACAC GGAAA UGUUG AAUAC UCAUA CUCUU CCUUU UUCAA UAUUA UUGAA GCAUU UAUCA GGGUU AUUGU CUCAU GAGCG GAUAC AUAUU UGAAU GUAUU UAGAA AAAUA AACAA AUAGG GGUUC CGCGC ACAUU UCCCC GAAAA GUGCC ACCUG ACGCG CCCUG UAGCG GCGCA UUAAG CGCGG CGGGU GUGGU GGUUA CGCGC AGCGU GACCG CUACA CUUGC CAGCG CCCUA GCGCC CGCUC CUUUC GCUUU CUUCC CUUCC UUUCU CGCCA CGUUC GCG

3.6.4 Assay of ^{Cy3}S4 and ^{Cy5}S4' binding to r613 template

The binding of the oligomers to the template was monitored through electrophoretic mobility in an agarose gel. The mixtures of the oligonucleotide substrates and the r613 template were mixed in a 20:1 molar ratio respectively, the template duplex was at a concentration of 2.7×10^{-8} molal in 6 mg of glycholine. Samples were dried down under vacuum centrifugation overnight. They were heated to temperatures above their Tm (heated to 95 °C for aqueous buffer and 80 °C for glycholine) for two minutes and then cooled down to room temperature at a rate of -0.5 °C / 8 s. Samples were left at room temperature for the indicated amount of time (between 0 and 48 hours), they were then mixed with 6 μL of water and loaded in a 2 % agarose gel and subjected for electrophoresis (30 V for 10 min, and 70 V for 1 h 15 min). The agarose gel was then imaged on a Typhoon FLA 9500 laser scanner (GE Healthcare) at a resolution of 50 μm. Cy3 fluorescence was measured using an excitation wavelength of 532 nm and the bandpass green (BPG1) emission filter. Cy5 fluorescence was measured using an excitation wavelength of 635 nm and the long pass red (LPR) emission filter. The gel was then stained with ethidium bromide in 1x TAE buffer with gentle shaking for 15 min and imaged using an excitation wavelength of 532 nm and the long pass green (LPG) emission filter.

3.6.5 Ligation of assembled oligonucleotide substrates

Samples were prepared in 2 mg of glycholine with the RNA template (r613) at a concentration of 4×10^{-7} molal (300 ng of template used per sample), and each of the 5'

phosphorylated oligonucleotide substrates present in a 4:1 molar ratio to the template. Samples were dried overnight; they were heated to 80 °C (glycholine) or 90 °C (aqueous buffer) for two minutes and then cooled down to room temperature at a rate of -0.5 °C / 8 s. Samples were diluted with water, 10x T4 RNA ligase 2 buffer (to a concentration of 1x), and 10 U of T4 RNA ligase 2 (New England Biolabs) were added, for a final volume of 20 μL. After incubating at 37 °C for one hour, 6 μL of sample were combined with 6 μL of 2x RNA loading dye (95% formamide, 0.01% SDS, 0.5 mM EDTA, pH 8). The samples were then heated to 95 °C for two minutes and placed on ice before loading in a 10% denaturing polyacrylamide gel, 1x TBE. Gels were pre-run for at least 30 minutes at 14 W and 45 V/cm before loading the samples. After sample loading, gels were run in these conditions for one hour. The gels were then imaged for Cy3 and Cy5 fluorescence (as described above) before staining with SYBR gold, and then imaged using an excitation wavelength of 495 and the long pass blue (LPB) emission filter.

3.6.6 Cleavage of substrate by r613 duplex

Samples were prepared in 2 mg of glycholine with the RNA template (r613) at a concentration of $4x10^{-7}$ molal, and the substrate (FAM-labeled in the 5' end) in a 4:1 molar ratio to the template. Oligonucleotide substrates were added in a 1:1 molar ratio to the template to a control sample to verify they had no effect on the cleavage activity from the template containing the HH ribozyme. Similar to the ligation assay, the samples were dried overnight; heated to 80 °C (glycholine) or 90 °C (aqueous buffer) for two minutes and then cooled down to room temperature at a rate of -0.5 °C / 8 s. Samples were diluted with water

and 10x T4 RNA ligase 2 buffer (to a concentration of 1x), for a final volume of 20 μL. After incubating at 20 °C for one hour, 6 μL of sample were combined with 6 μL of 2x RNA loading dye (95% formamide, 0.01% SDS, 0.5 mM EDTA, pH 8). The samples were then heated to 90 °C for two minutes and placed on ice before loading in a 20% denaturing polyacrylamide gel, 1x TBE. Gels were pre-run for at least 30 minutes at 14 W and 45 V/cm before loading the samples. After sample loading, gels were run in these conditions for one hour. The gels were then imaged for FAM fluorescence, measured using an excitation wavelength of 473 nm and the long pass blue (LPB) emission filter, before staining with SYBR gold (as described above).

3.6.7 Purification of ligated products and cleavage assay

In order to demonstrate that the newly ligated products containing the HH ribozyme sequence were active, a set of 90 ligation reactions were carried out in parallel, each with 5 mg of solvent, with 750 ng of template ($4x10^{-7}$ molal) and each of the oligonucleotide substrate in a 4:1 molar ratio to the template. The same protocol for ligation as described above was followed. The solutions were then mixed together and concentrated using Zymo Oligo Clean & Concentrator columns. The products between 192 and 320 base pairs were purified by denaturing PAGE (1x TBE, 10% polyacrylamide, 8 M urea). The purified products were then used for a cleavage assay using the same protocol described in section 3.6.6.

CHAPTER 4. MULTIPLE REPLICATION ROUNDS IN A PREBIOTIC SOLVENT

4.1 Introduction

In the previous chapters, some of the challenges associated with nucleic acid replication were addressed: the strand inhibition problem, the paradox of structured sequences that can serve as templates, and the transition between the duplex and the active state of a ribozyme. However, multiple replication cycles were not possible due to an irreversible change in solvent conditions: added Mg^{2+} required for ligation could cause DNA and RNA precipitation upon water removal in glycholine.

Multiple replication rounds under prebiotic conditions have been reported for PCR-like reactions, in which they use a polymerase (RNA-based or protein enzyme) to add individual nucleotides to a primer and use heat-cool cycles to promote replication[8,22]. However, the sequences that have been replicated by this method are relatively short and fast temperature changes are required to overcome strand inhibition. A different approach that has been taken for continuous replication, is the use of oligomers and non-enzymatic ligation by using activating agents such as EDC[21], but the sequence was designed to be unstructured and the system was only composed of two oligomers that were ligated after assembly. Again, the main challenge for multiple replication rounds to occur is strand inhibition, and long sequences of more than 80 bp have not been reported to undergo continuous replication because they are also likely to form intramolecular structures.

Therefore, based upon the promising results we had observed with the glycholine replication system, our next goal was to achieve multiple replication cycles of a long duplex with the aid of a different environment provided by the solvent.

As has been mentioned, glycholine prevented multiple replication rounds due to precipitation of nucleic acids in the presence of divalent cations after water evaporation. However, if this process took place in a pond or a puddle, then changes in the environment could have promoted different steps of the process through wet-dry and hot-cold cycles. These environmental changes, have been reported to favor polypeptide formation from amino acid and hydroxy acid building blocks[60], as well as synthesis of building blocks such as nucleosides[61].

Therefore, we sought to explore conditions that could bring the system closer to a prebiotic environment and allow multiple replication rounds through wet-dry cylcing. There is a wide variety of compounds that were likely abundant in the early earth, such as urea, acetamide, inorganic salts, amino acids, sugars, organic acids, etc.[62] Consequently, mixtures of these compounds could have provided an environment favorable for replication of nucleic acids. Thus, our goal was to analyze the behavior of DNA and RNA under different mixtures and conditions containing some of these chemical compounds and investigate how these mixtures can provide an environment that promotes multiple replication cycles of these informational polymers.

We propose a model very similar to the one proposed for RNA replication, highlighting the wetting and drying, as well as the temperature changes of the system to promote replication of a long duplex (Figure 23).

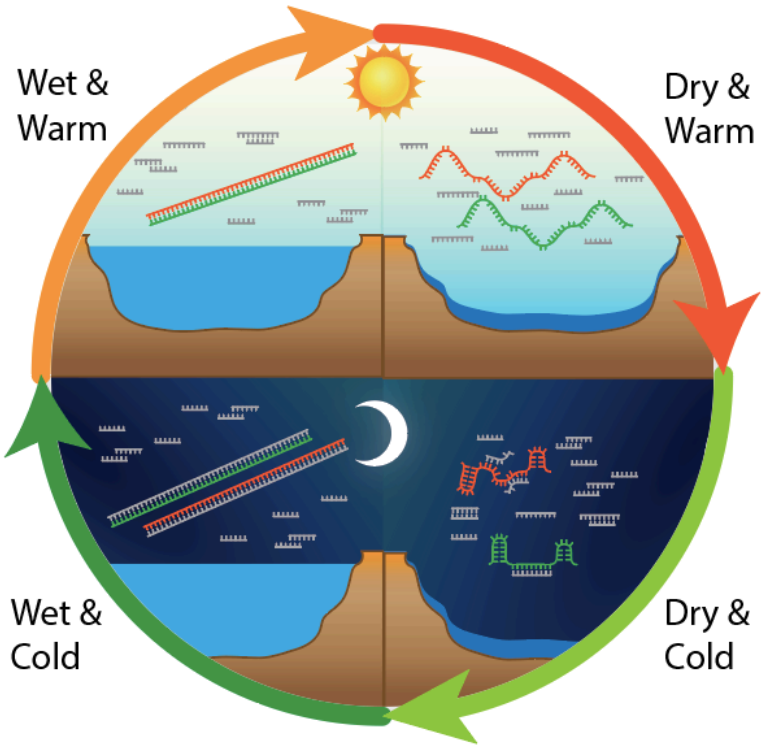


Figure 23. Representation of how an environmental cycle can promote replication. Water evaporates as the environmental temperature increases. The reduced water activity and high temperature promotes the denaturation of a template duplex. Upon cooling, the single stranded template can be kinetically trapped, allowing complementary oligomers to bind to the templates. The assembled oligomers can then be ligated upon rehydration of the system to form new copies of the template strands.

4.2 Solvent selection

Selection of a good solvent is crucial for this system, as mentioned, nucleic acids tend to precipitate in non-aqueous solvents in presence of divalent cations due to the change of the dielectric constant. We were also looking for solvents in which DNA could form duplexes and melt under mild conditions, and preferably consisting of prebiotic compounds. Based on previous experience with glycholine and reline[11,23,24,63] and what was presented in the previous chapters related to viscosity dependence, we were also

looking for a viscous solvent that could promote the trapping of long sequences and slow their diffusion compared to the oligomers. There were many different mixtures that were evaluated, some of them based on different eutectics reported in the literature[64-69], and other variations from those reported (changing the molar ratios or the hydrogen bond donor or acceptor). Some of the mixtures that were explored contained urea, acetamide, ammonium nitrate, magnesium chloride, zinc chloride, sugars, and organic acids. Among the mixtures that were evaluated, the ones that proved to work best contained urea mixed with acetamide and water in different proportions (abbreviated as UAcW).

Some of the mixtures that were tested are displayed in Table 8. Among the challenges involved with these solvents was that they became solids upon cooling to room temperature, the ones based on organic acids had a very low pH, DNA was denatured due to the high concentrations of urea, or the high concentration of divalent cations could degrade RNA[70].

After running preliminary tests with the solvents that looked more promising and testing solubility of DNA in the presence of $MgCl_2$, we decided to further explore the urea/acetamide/water solvent (UAcW) even though its viscosity was very low compared to glycholine.

Table 8. Solvent mixtures tested for stability and DNA solubilization

Components	Molar ratio	Observations
Urea/$ZnCl_2$[64]	3.1/1	Very viscous, there is precipitate formation upon addition of water (probably $Zn(OH)_2$, DNA appears to be denatured in it
Urea/$ZnCl_2$/NH_4Cl	22.7/7.6/1 and other ratios	Formation of crystals, not very reproducible
Acetamide/$ZnCl_2$[64]	1/4	Not very viscous, too much zinc can affect the DNA
Urea/$Zn(OAc)_2.2H_2O$[65]	4/1 and other ratios	Not very reproducible, formation of crystals over time
($EtNH_3Cl$)/Urea[66]	1/1.5	Not stable over time, forms a solid
Lactic acid/NH_4OAc[67]	3/1	Not viscous enough, pH is too low to work with nucleic acids
Urea/$FeCl_2$	1/1 to 4/1	None of them formed a eutectic, iron was oxidized with the heating and upon exposure to air
Urea/$MgCl_2.6H_2O$	4.5/1	Not very viscous, CD signal of DNA looks ok and melt is reversible
Urea/$MgCl_2.6H_2O$/NH_4Cl	4.5/1/0.1 4.5/1/0.2	Slightly more viscous than the mix without ammonium chloride, hard to incorporate the ammonium chloride in the mix
ChCl/$MgCl_2.6H_2O$[68]	1/1	Very viscous, but not really prebiotic (trying to stay away from choline chloride)
Acetamide/$MgCl_2.6H_2O$	2/1 and 1/1	Crystals never disappeared completely, solid after cooling
Acetamide/urea/MgTriflate[69]	11.4/7.6/1	Became a solid overnight at room temperature
Acetamide/urea/$MgCl_2.6H_2O$	11.4/7.6/1 8/5.28/1	Not very viscous, water and acetamide removal (on speedvac or in oven) increased viscosity but final composition can vary on time and conditions
Acetamide/urea/$Mg(OAc)_2.4H_2O$	11.4/7.6/1 8/5.28/1	Not very viscous, acetate can affect pH when water is added
Urea/Acetamide/Water	1/2/4	Recommended by Bradley Burcar, not very viscous but DNA is soluble with $MgCl_2$

4.3 UAcW characterization

First, we proceeded to evaluate the solvent stability and behavior under different conditions. Starting with a mixture containing urea, acetamide and water in a 1/2/4 molar ratio, water was removed either under vacuum or by heating at 65 °C. The composition of

the solvent varies slightly as demonstrated by NMR analysis of the mixture (Figure 24). Some acetamide is lost in the drying process under vacuum, which could potentially limit the number of replication cycles. This loss of acetamide was verified by drying pure acetamide on the speed vac and after one day, all of it was gone, verifying its volatilization under these conditions. Nonetheless, acetamide could easily be added during the rehydration phase in each cycle. This is a very stable solvent, since there is no degradation of any of the compounds and no new peaks appear over time as can be seen by NMR analysis (Figure 24).

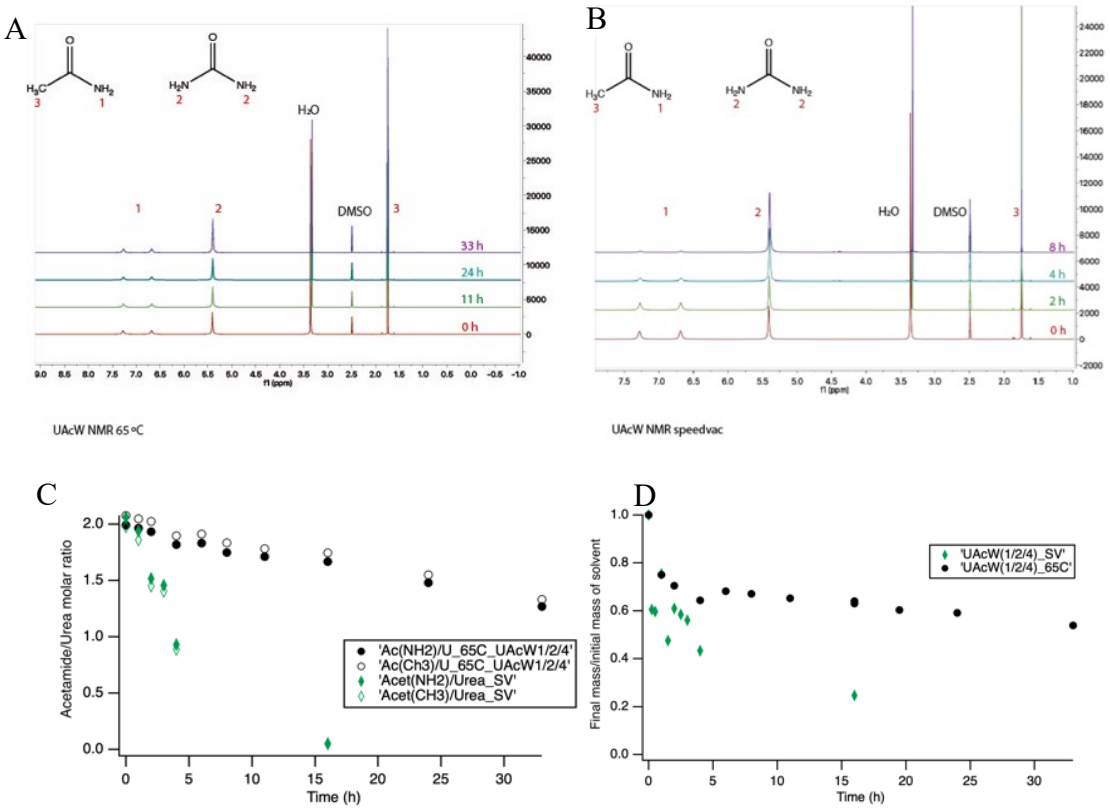


Figure 24. (A) NMR spectra of UAcW drying at 65 °C or (B) under vacuum centrifugation. No new peaks appear over time, demonstrating there is no degradation of the solvent components. (C) Molar ratio of urea/acetamide determined from the NMR spectra, there is loss of acetamide, mainly after drying under vacuum. (D) Mass change over time represented as fraction of the initial mass for the same drying conditions.

To further explore different possible scenarios for driving out the water, samples were dried under different conditions[71,72], and the mass was monitored over time. Initial volume of UAcW (1/2/4) was 2 µL, and water was added to a final volume of 25 µL. The rate at which water is evaporated varies depending on the conditions, but in most cases, except under high relative humidity (75%), it reaches equilibrium and a final mass below 2 mg (Figure 25).

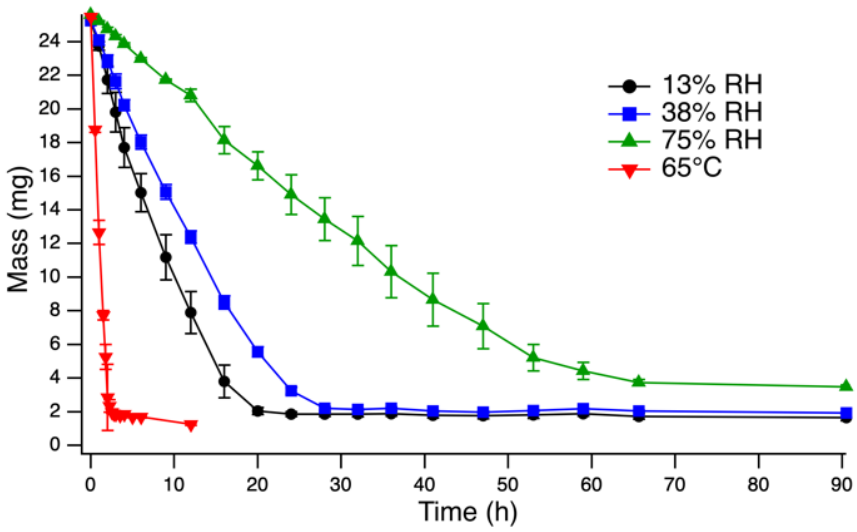


Figure 25. Mass loss over time of hydrated UAcW under different drying conditions: 13% RH, 38% RH, 75% RH at room temperature, and 65 °C.

4.4 DNA behavior in UAcW

DNA and RNA are soluble in UAcW, even in the presence of divalent cations, such as Mg^{2+}. A CD melt was done with a 32 bp duplex, and its Tm decreases as the amount of

water in the mixture is reduced (Figure 26). Additionally, the CD trace of the double stranded DNA in the UAcW solvent is very similar to the CD trace in aqueous buffer, suggesting that the double helix structure is not greatly affected in the solvent.

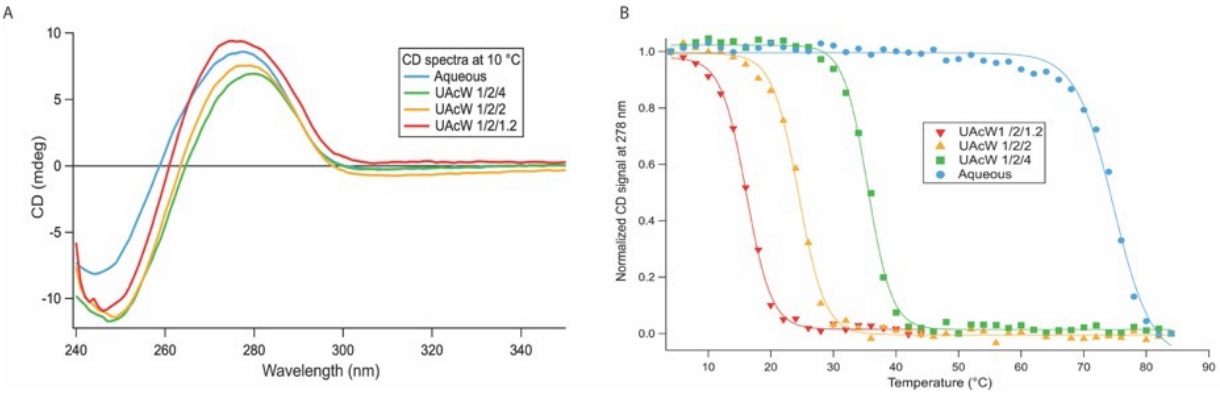


Figure 26. (A) CD spectra of a 32 bp DNA in aqueous solution and in UAcW with different amounts of water. The B-form of the DNA does not appear to be affected in the UAcW solvent with low water activity. (B) Normalized CD signal from the heating traces of the melts at 278 nm with different amounts of water (markers represent the experimental data and the line corresponds to a sigmoidal fit to estimate Tm values).

Based on the CD melts for the 32 bp sequence, the Tm decreases from 74 °C in aqueous buffer to 16 °C in the UAcW (1/2/1.2). A 641 bp DNA duplex was selected for the replication experiments, to demonstrate that long templates with different possible internal structures could be copied. As part of its characterization, melts were done in the UV-Vis spectrometer because a lower concentration is needed with this instrument. The Tm was estimated from absorbance at 260 and 270 nm since the amides from the solvent start to absorb at around 260 nm. Nonetheless, melting curves look clean enough to estimate the Tm of the DNA. For this sequence, the Tm decreases from 83 °C in aqueous

buffer to 28 °C in the UAcW (1/2/1) solvent (Figure 27). Given the decrease in the Tm of the DNA sequences analyzed in this prebiotic model mixture, extreme conditions such as high temperatures are not needed to denature the strands, thus decreasing the temperature ranges needed for denaturation of the templates and annealing of the substrates to the template strands.

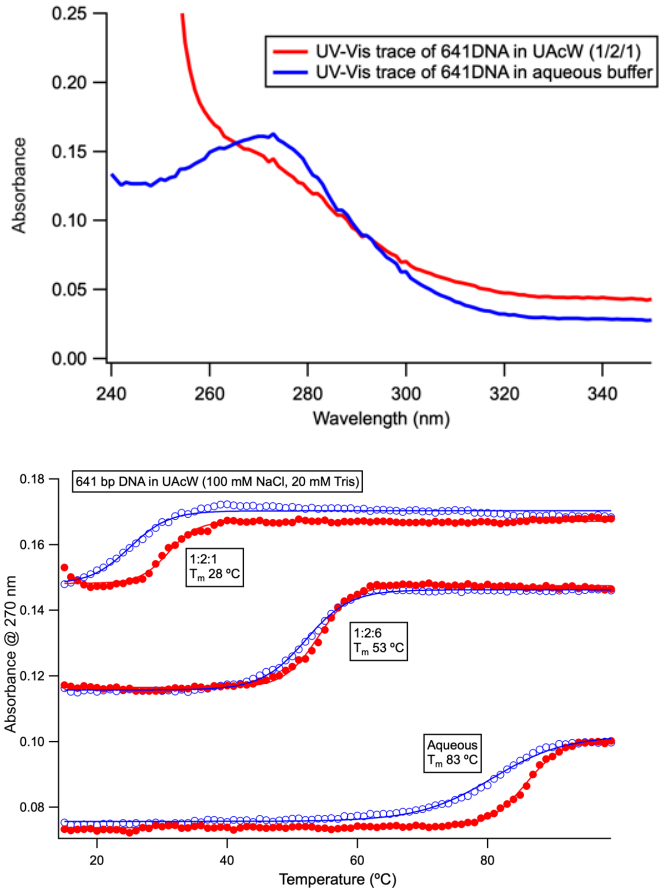


Figure 27. UV-Vis spectra in UAcW (1/2/1) and in aqueous buffer; and melting curves for the 641 bp template (100 mM NaCl, 20 mM Tris). The melting curves were provided by Bryce Clifton.

4.5 Assembly of oligomers on trapped template

To verify whether there was trapping of the template in a similar way it was observed in glycholine, a heat cycle was done with the DNA template alone (641 bp long), and in the presence of the complementary oligomers. As can be seen in Figure 29, two bands appear after heating and cooling of the template in UAcW, representing the trapped and the double-stranded template. Furthermore, oligomers assemble onto the trapped template and remain bound for approximately 2 hours (Figure 29). The kinetics of duplex reformation in UAcW appears to be much faster with this system as compared to glycholine. This might be due to the lower viscosity of the solvent.

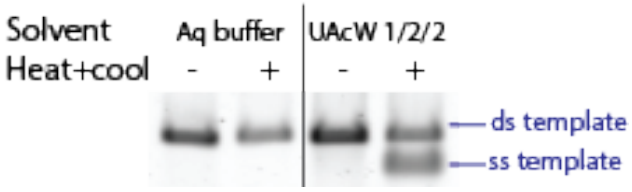


Figure 28. Trapping of the template after heating and cooling is observed in UAcW, but not in aqueous buffer in a 2% agarose gel. Staining with EtBr confirms the formation of intramolecular folds within the trapped single stranded template

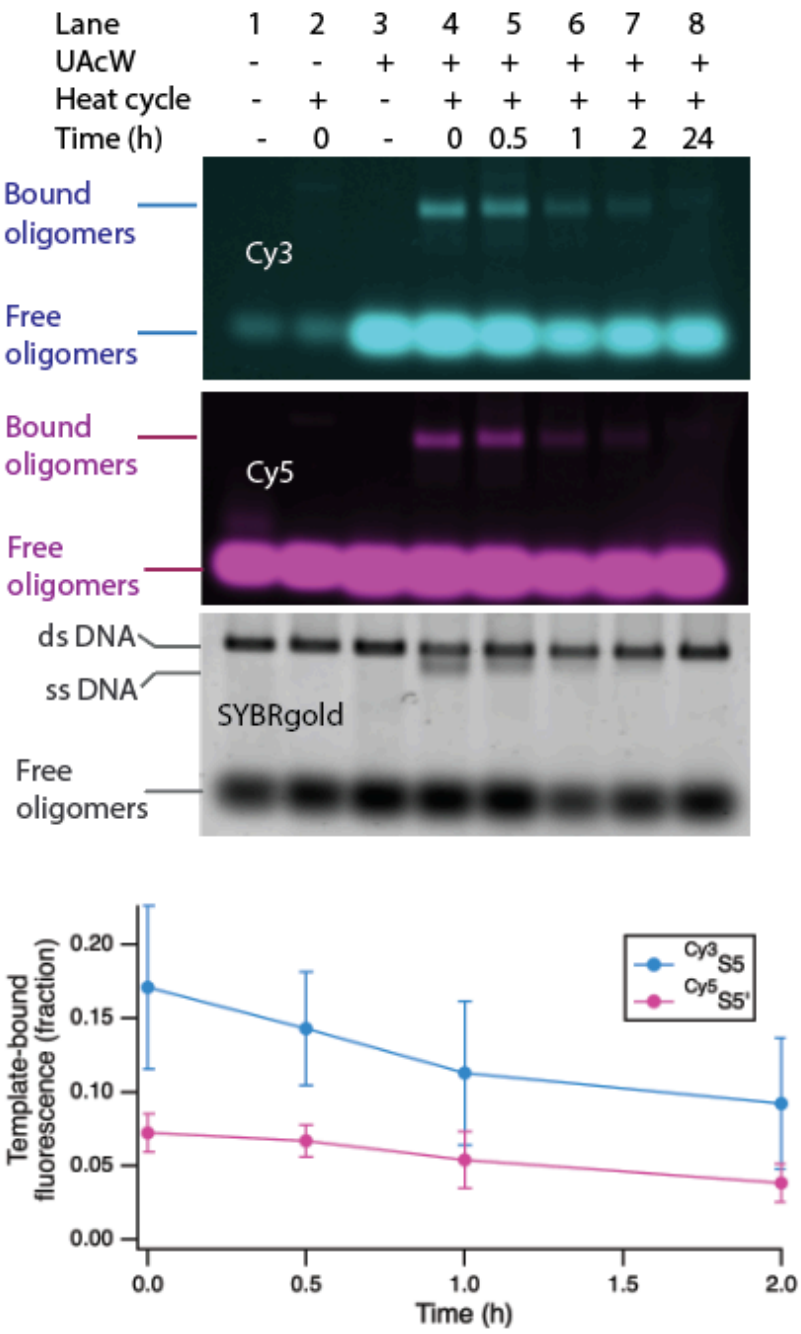


Figure 29. Trapping of template and assembly of oligomers on trapped template after heating and cooling in UAcW. No trapping is observed in aqueous buffer or without the heat cycle. Graphic representation of the bound oligomers based on densitometry analysis from the gels.

4.6 Replication and selection of a ligase

For this project, we define replication as a simultaneous template directed synthesis of a copy of both strands from a template. To achieve this, we use oligomers that are complementary to the template, the same way we did in Chapter 3. Once the oligomers are assembled on the template, we use a ligase as a robust method of ligating them. A ligase capable of ligating nicks in double stranded template is needed. It is important that it doesn't ligate blunt ends, since the oligomers are complementary to each other, they will likely pair and this could lead to non-templated ligation of the oligomers.

A 641 bp template was used, derived from the template originally used by He and Gallego[11]. Ten oligomers that are 32 nucleotides long complementary to each of the duplex strands were used for the replication assays. Heat cycling was done in 1/2/4 UAcW, followed by a 10-fold dilution with aqueous buffer.

Three ligases that are capable of ligating nicks in double stranded DNA were evaluated: T4 DNA ligase, T7 DNA ligase and 9°N ligase (Figure 30). From these enzymes, T4 DNA ligase displayed better yields, but it is capable of ligating blunt ends, resulting in products without template as shown in lanes 1 and 2 from Figure 30. T7 DNA ligase and 9°N DNA ligase resulted in similar ligation yields, and none of them is heat deactivated. In absence of the template, up to trimers can be observed with the T7 ligase. On the other hand, 9°N ligase requires Triton X-100 in the buffer, a detergent which could cause problems for the ultimate goal of multiple replication rounds with the wet-dry cycling. Since these last two ligases seemed to be more adequate for our purpose and did

not exhibit a lot of blunt end ligation, we proceeded to evaluate and optimize their working conditions.

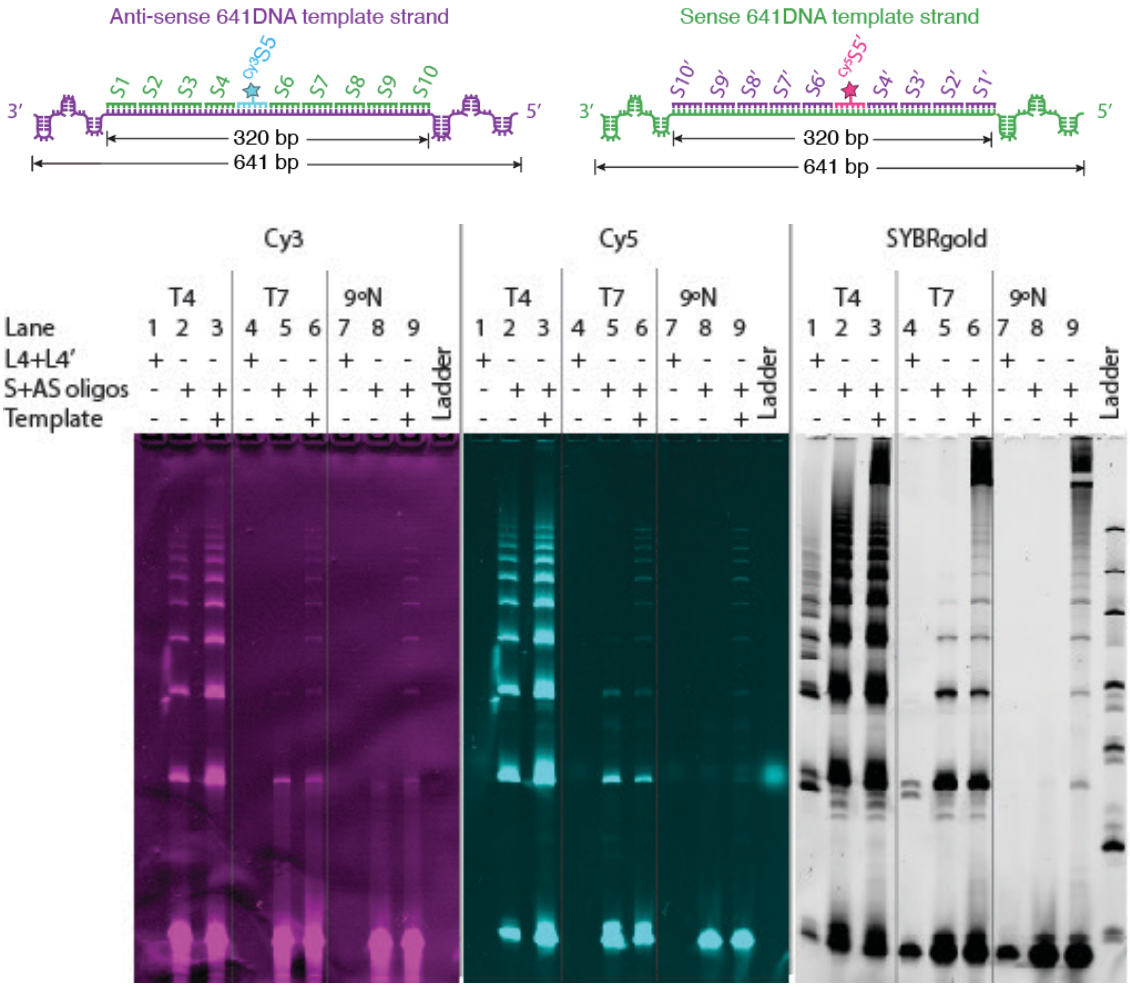


Figure 30. Replication of a 320 bp fragment from a 641 bp template with different ligases: T4 DNA ligase, T7 DNA ligase and 9°N DNA ligase. Experiments were done in presence and absence of template to evaluate if there was non-templated ligation of free oligomers.

To further evaluate possible conditions for the use of the ligases, different buffers and temperatures were tested with the T7 DNA ligase and the 9°N ligase. For the T7 DNA ligase, incubation temperatures of 15, 25 and 37 °C, with two different buffers: the T7 and

T4 DNA ligase buffer. For the 9°N ligase, temperatures of 40 and 50 °C were tested (Figure 31). Overall, the trimers disappeared when the T4 DNA ligase buffer was used with the T7 DNA ligase in absence of the template. This implies that there is less non-specific ligation and products that are longer than the dimers can be quantified. The yield with this ligase was better than the yield for the 9°N ligase. Therefore, the enzyme selected for further replication was T7 DNA ligase, with T4 DNA ligase buffer (at 1x: 50 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM ATP, 10 mM DTT, pH 7.5 @ 25°C).

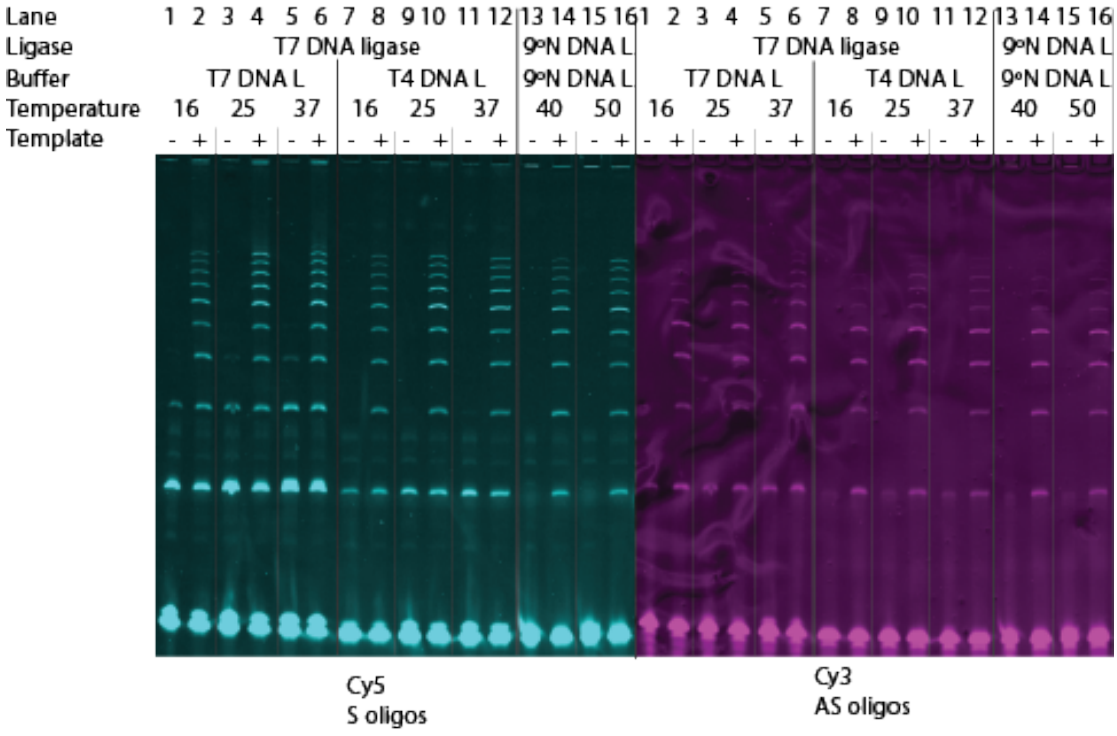


Figure 31. Ligation of oligomers with T7 DNA ligase and 9°N ligase in presence and absence of template to verify there is no blunt-end ligation of free oligomers. Experiments were done under different temperature and buffer conditions to optimize replication conditions without non-specific ligation of the oligomers.

4.7 Variables that affect replication from a DNA template in UAcW

With this new system, there are multiple variables that can affect information transfer from the strands. These variables include the amount of water present in the system, the heating and cooling temperatures and rates, as well as the overall composition of the solvent. Therefore, individual experiments to evaluate how these conditions can affect replication were explored.

4.7.1 Water content in UAcW

To evaluate how the amount of water present in the system can affect information transfer from the templates, replication experiments with known amounts of water were performed. In addition, different compositions of all three components were explored.

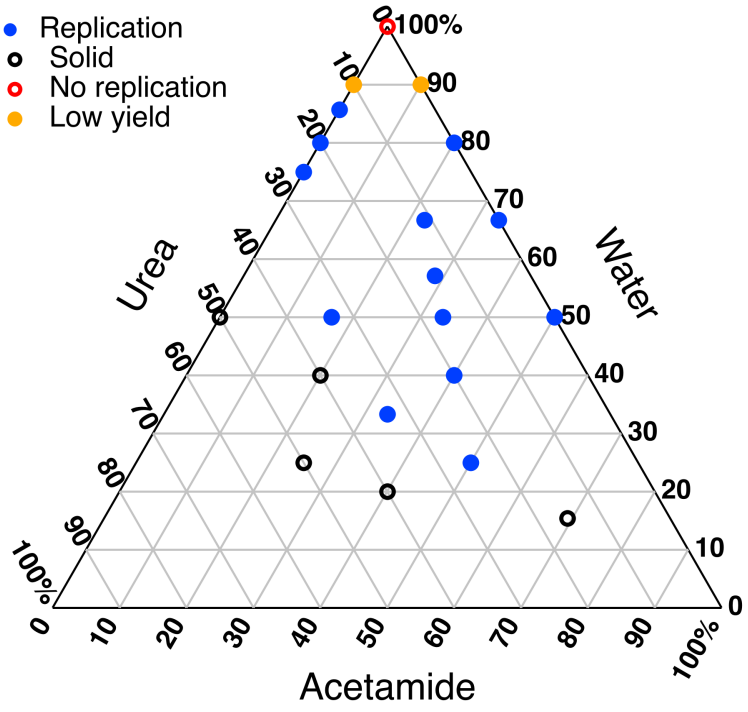


Figure 32. Ternary diagram that indicates the different molar fraction of each component of UAcW that were prepared and evaluated for replication. Blue dots represent compositions under which replication occurred, red circles indicate compositions under which the solvent was a solid at room temperature and were not used for replication, and red circles indicate compositions at which there was no replication observed.

The ternary diagram in Figure 32 shows the different UAcW ratios that were evaluated for replication. Open circles indicate that the mixture remained solid despite heating or that it became solid upon cooling to room temperature, and therefore, these mixtures were not used in further experiments. As can be seen in Figure 33, there is a wide range of molar ratios that promote replication. It is important to note that the mixtures at 1/1/1 and 2/1/3 tend to go to a solid if they are left at room temperature for a few days, so they were made fresh for this experiment.

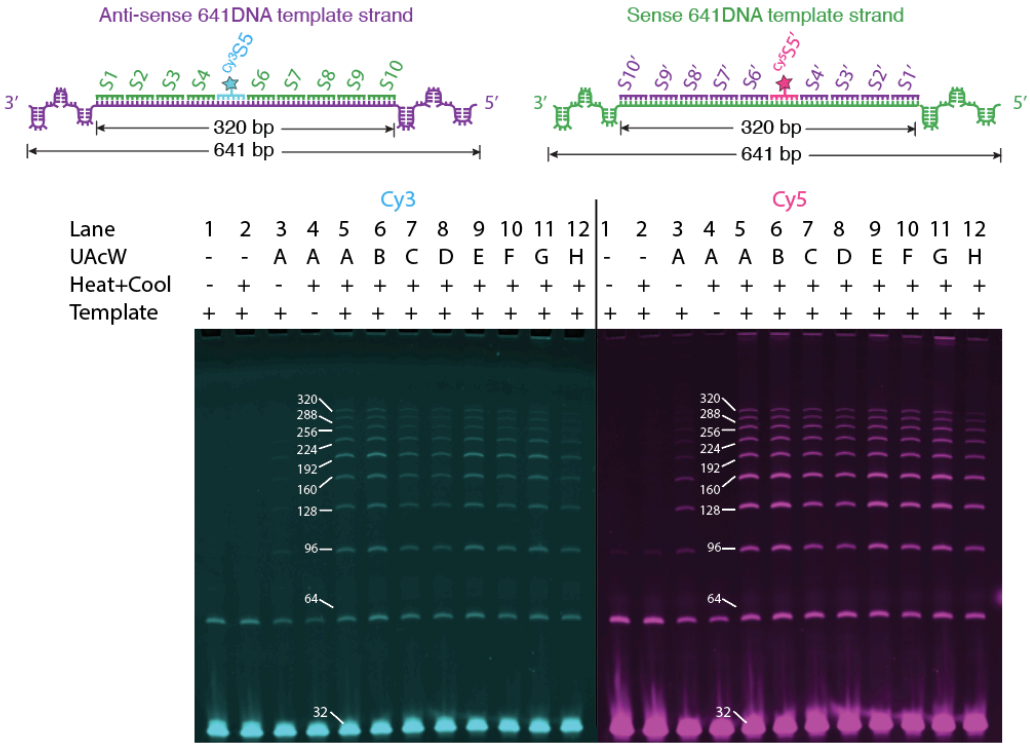


Figure 33. Denaturing polyacrylamide gel showing replication of a 320 nt fragment from a 641 bp template is favored under different ratios of urea/acetamide/water. Molar ratio compositions Urea/Acetamide/Water as follows: A. 1/2/2, B. 1/2/1, C. 1/2/4, D. 1/2/6, E. 1/1/1, F. 2/1/3, G. 0/1/1, H. 1/0/4.

There was no replication observed in aqueous solution (lanes 1, 2), when there is no heat cycling in either mixture (lanes 1, 3), or when the template is missing in UAcW (lane 4). Replication occurred when all three components (urea, acetamide and water) were present in the mixture and there was heat cycling. This corroborates there is no self-assembly of the oligomers without template, and products are result of template-directed synthesis. Furthermore, replication occurs in the binary mixtures of acetamide/water and urea/water. This suggests that in addition to the trapping of the templates, there are other factors involved that promote assembly of the oligomers on the template, leading to

replication. The ternary mixture as well as the solutions with its individual components (urea/water and acetamide/water) reduce the thermostability of nucleic acid duplexes, promoting the single-stranded state [73,74]. It has been reported that organic co-solutes that contain amides, such as formamide, increase the rate of strand invasion with short substrates to a hairpin [75]. These two phenomena might be related, resulting in the trapping of the single stranded templates and the assembly of the oligomers onto the trapped templates that we observe. However, a more detailed study would be needed to really understand how the acetamide and urea are favouring this process. Nonetheless, it was very unlikely to find a pool only containing water and urea or water and acetamide in the prebiotic Earth. These results are showing that replication occurs in a solvent composed of prebiotic molecules such as acetamide and urea with different water content. More complex mixtures were likely present 4 billion years ago and having different ratios of these three components enables replication.

4.7.2 Effect of salts in replication

To further explore different scenarios and evaluate the robustness of this system, salts with divalent cations were added to the solvent mixture (Figure 34). Even though the yield of ligated products decreases with the presence of salts, replication was observed in all cases, including the concentration that has been estimated was present in seawater 3.2 billion years ago (lane 8 from Figure 34).[76] This results corroborate the there is no one specific set of conditions or concentrations needed for replication to take place, it works in the presence and absence of salts and various solvent compositions.

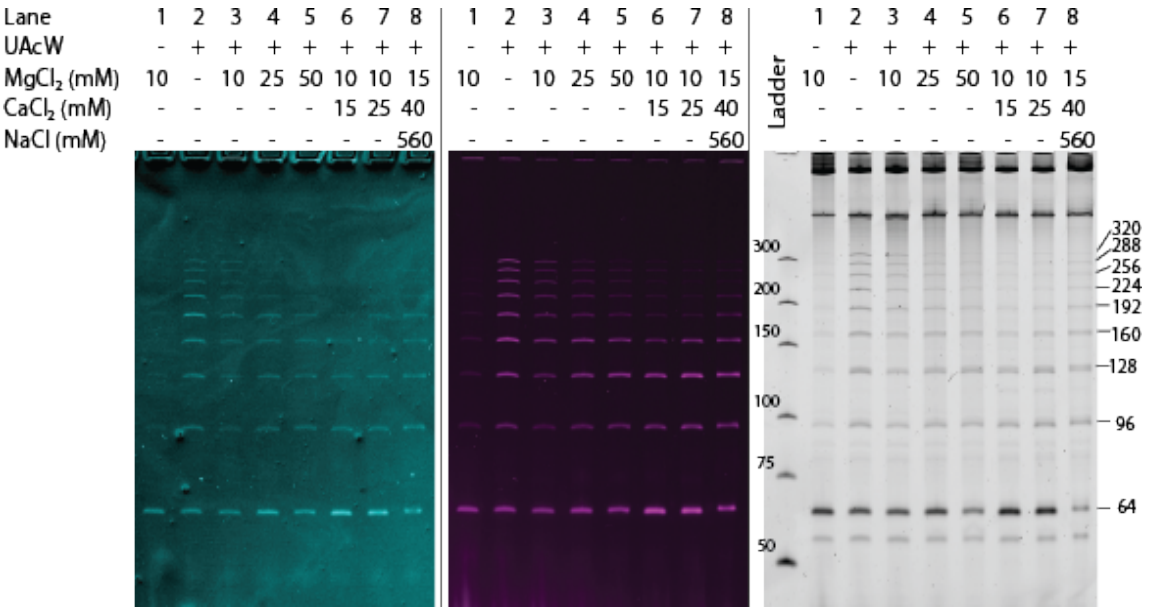


Figure 34. Denaturing polyacrylamide gel showing replication with addition of salts in various concentrations. The indicated salt concentrations correspond to the final concentration after 1/10 dilution during the rehydration process for ligation.

4.7.3 Drying conditions

Based on the characterization from the system and as a preliminary experiment before proceeding with multiple replication rounds, we explored replication after drying in different conditions for varying periods of time.

First, samples were prepared with an initial volume of 25 µL and dried at 65 °C for different amounts of time. Ligation yield was evaluated to find the conditions that would provide the highest yield (Figure 35).

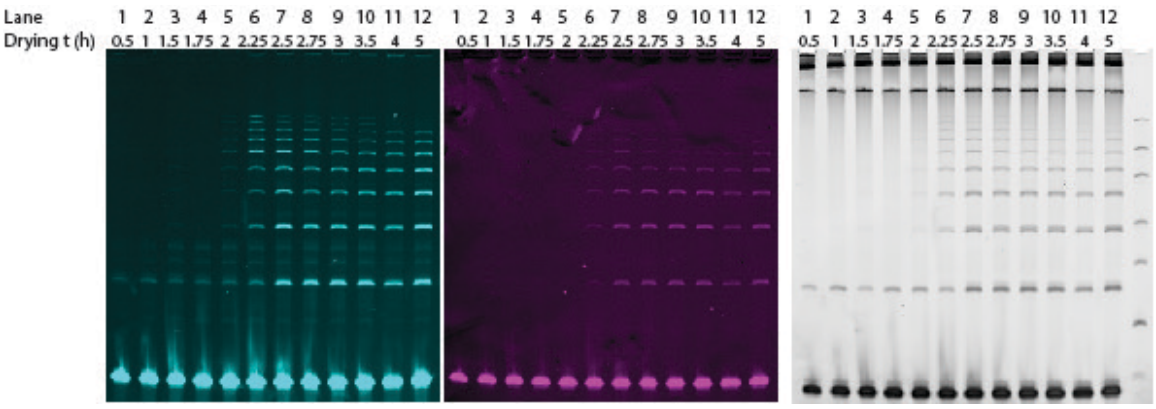


Figure 35. Replication of 320 bp fragment from a 641 bp template after drying at 65 °C for different times. Initial volume for all samples was 25 µL which contained 2 µL of UAcW (1/2/4) and 2µL of the 10x T4 DNA ligase buffer. The system requires drying for at least 2.25 hours for full replication to take place, and after 3 hours the replication yield starts to decrease.

As seen in Figure 35, drying for at least 2 hours is needed to remove enough water to promote replication. If the mixture is dried for too long (more than 3 hours), the ligation yield starts to decrease, indicating that some water is still needed for optimum ligation, and some of the acetamide loss might also be related. Based on the mass loss analysis from Figure 32, approximately 90% of the total weight is lost, which is attributed to water evaporation, resulting in a final composition within range of the blue data points in the ternary diagram.

Furthermore, we explored different ways of removing the water from our system. Drying at 65 °C could have been possible in a prebiotic Earth, nonetheless, as shown in Section 4.3, there are other ways in which water could be removed. Based on these results we decided to investigate if drying at lower temperatures with low relative humidity could provide similar results with longer drying times. To reach a similar water loss based on the

mass, up to 2 days of drying were required at 13 and 38 % RH at room temperature (see section 4.3, and Figure 25). Therefore, a ligation experiment after drying under these low relative humidity conditions for up to 48 hours was performed. Drying at 13% RH for 24-36 hours, and at 28% RH for 36-48 hours resulted in full-length products after rehydration and addition of the ligase.

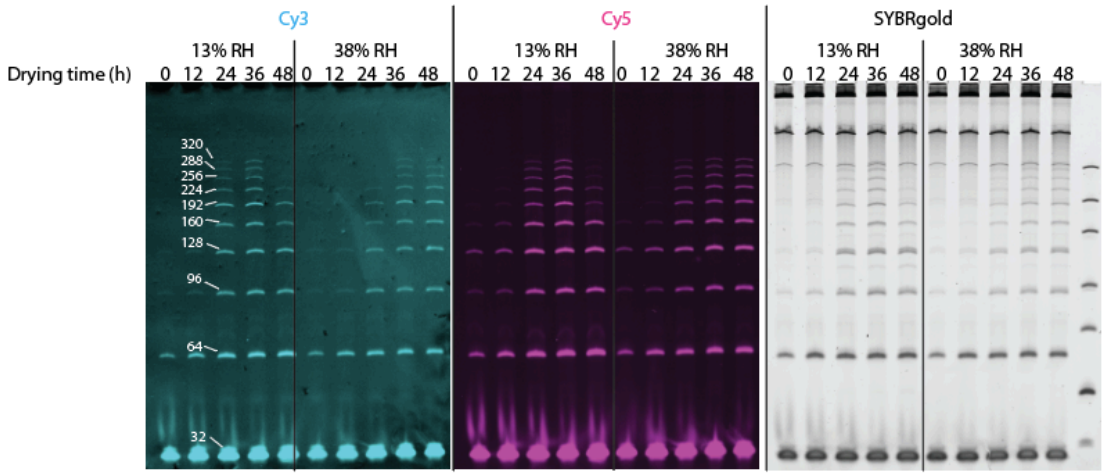


Figure 36. Ligation after drying at room temperature with low relative humidity. At least 24 hours of drying are needed to observe full replication.

Replication of the template is observed under very different conditions: various solvent compositions, presence and absence of salts, and different drying conditions and times. These results show that the proposed system is very robust and is not limited to a particular set of conditions. Without heat cycling and in aqueous solution, no ligation is observed, suggesting that temperature variations and low water content are important to favor separation of the template strands and assembly of the oligomers.

4.8 Multiple replication rounds of a DNA sequence

4.8.1 Information transfer from one of the template strands

After solvent characterization and evidence for replication within this system, the next step and ultimate goal for this project was to demonstrate that multiple replication rounds are possible. A preliminary experiment with information transfer from only one of the template strands from the 641 DNA duplex was done as a first approach to evaluate the conditions necessary for multiple replication cycles. Up to 5 drying/hydrating cycles were done, drying under vacuum for 1 hour to reduce the time of the experiment. No buffer (Tris and $MgCl_2$), oligomers or template were added between each round. T4 DNA ligase, ATP and DTT were added during each rehydration cycle since T4 DNA ligase is heat deactivated and some of the ATP and DTT might be lost or consumed during the drying and ligating process. As shown in Figure 37, there is an increase of the ligated products with each cycle.

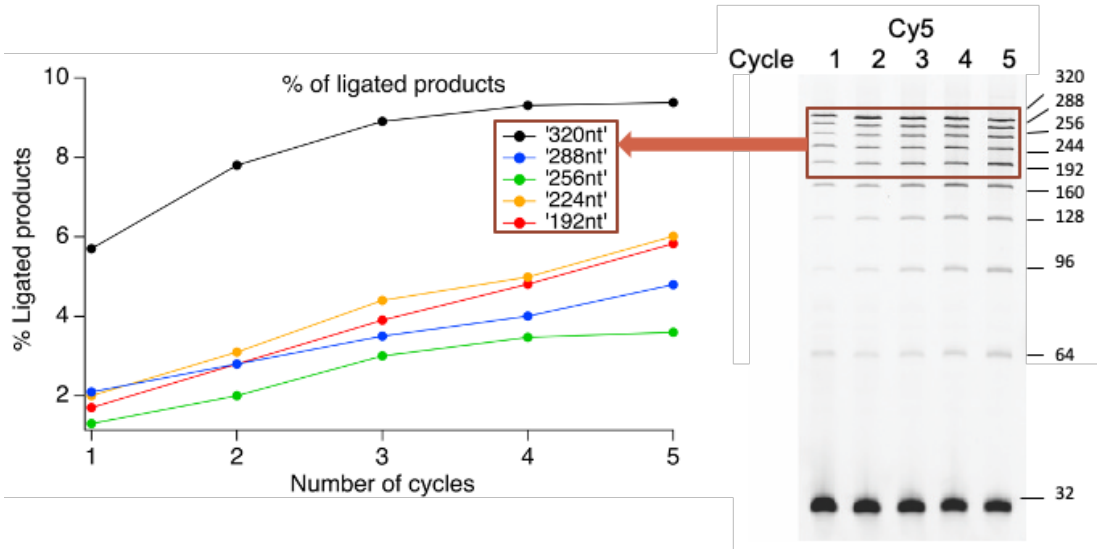


Figure 37. Information transfer from one of the template strands during 6 wet-dry cycles. Product yields of products with lengths between 192 and 320 nt are shown in the plot, calculated as mole percentage of ligated products based on densitometry analysis from the gel.

4.8.2 Replication rounds through wet-dry cycling

After demonstrating that multiple rounds of information transfer from one of the strands are possible, multiple replication rounds (copying both strands at the same time) were assayed. A similar protocol was followed, with the difference that T7 DNA ligase was used instead because T4 DNA ligase can ligate blunt ends, producing ligation of the oligomers even without a template (see section 4.6).

First, to evaluate whether it was necessary to add enzyme, ATP or DTT between each cycle, three wet-dry cycles were performed with and without addition of these components. Samples were dried for 30 min in a speed vacuum concentrator to save time. As seen in Figure 38, better yields were obtained when the ligase was not added after the

first cycle. This might be because with each addition of the ligase, glycerol (used for enzyme storage) is also added to the mix, which might be inhibiting the overall reaction. Moreover, addition of ATP and DTT within each cycle improved the yield. In this case, enough ATP and DTT were added to replenish approximately half of the initial concentrations. Some of the ATP will be consumed during the ligation reaction, and a small amount of both components might be removed/degraded during the drying process, so replenishing them helps with the ligation reaction.

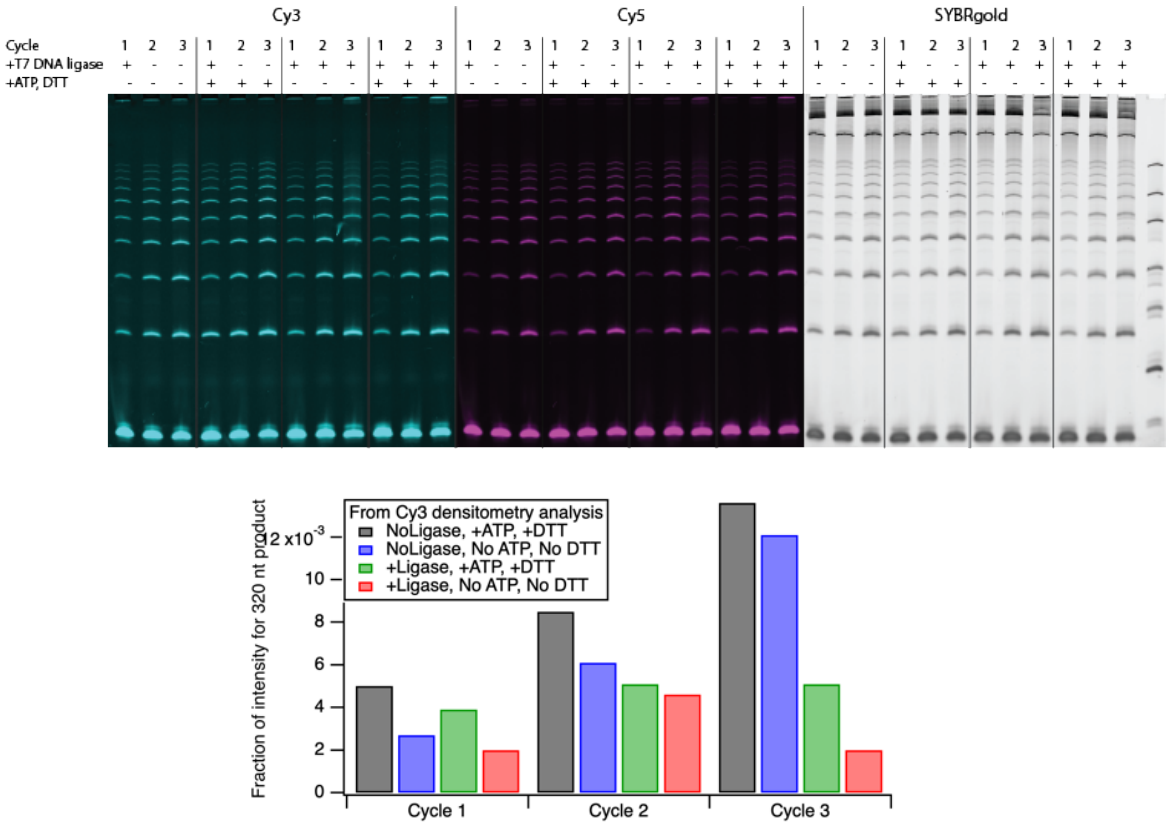


Figure 38. Denaturing gel and densitometry analysis from 3 replication cycles for the full ligation product using T7 DNA ligase, with and without addition of the ligase, ATP and DTT in each cycle for optimization of conditions for wet-dry cycling.

Once the conditions were optimized for the replication cycles drying under vacuum, up to five cycles were done in triplicate without addition of ligase, buffer or DNA in each cycle. Only a small amount of ATP and DTT were added in each rehydration step. As seen in Figure 39, each cycle results in a higher yield of the ligated products.

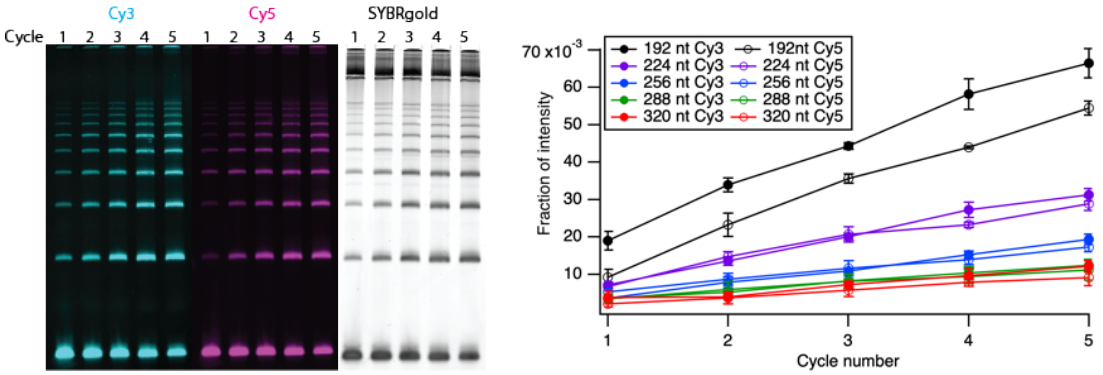


Figure 39. Five replication cycles were done after wet-dry cycling in UAcW. Drying was done on a speed vacuum concentrator, and the plot contains the fraction of intensity from densitometry analysis.

Next, to simulate a more prebiotic environment, and based on previous results obtained for replication, instead of drying under vacuum, samples were dried at 65 °C for 2.25 h. As with the previous experiment, T7 DNA ligase was added after the first hydration cycle, and it was not added in subsequent cycles (this enzyme does not get heat-deactivated), and only a small amount of ATP and DTT were added in each cycle.

There is an increase in replication products for the first three cycles. After the fourth cycle, products that were formed within the first three cycles start to degrade (Figure 40). This degradation is likely due to the constant heating of the samples, which results degradation of the ligase, template and substrates.

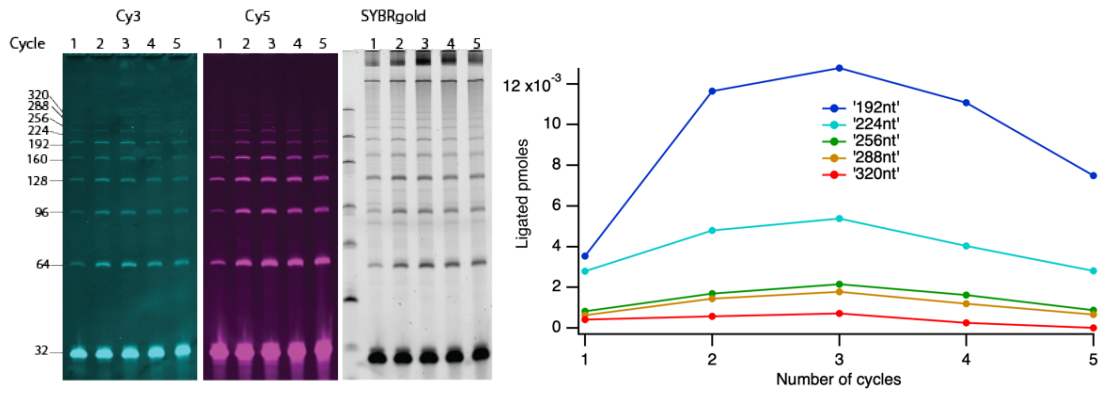


Figure 40. Wet-dry cycling promotes replication. Gel displays 5 replication cycles drying at 65 °C. No ligase was added after the first cycle. Degradation is observed after the third cycle.

Lastly, a similar test was done by drying under low relative humidity at room temperature, conditions that proved replication was also possible. The cycles with drying under low relative humidity at room temperature took approximately 2 weeks, since each drying cycle took approximately 36 hours. This resulted in even more degradation of the template and the substrates compared to the other two drying methods that were previously evaluated (Figure 41).

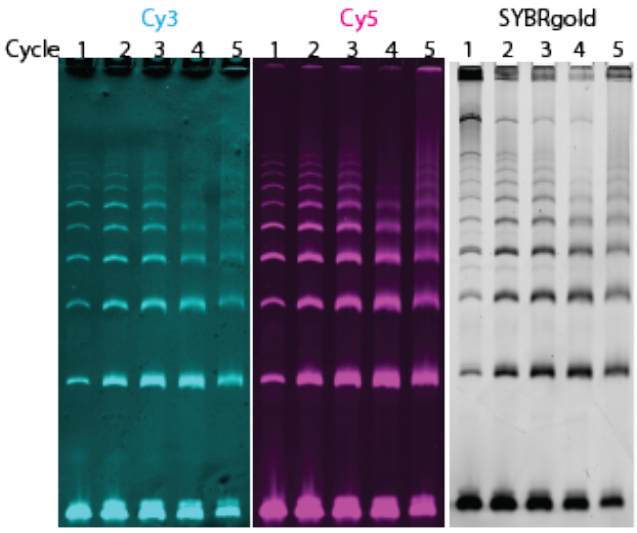


Figure 41. Replication cycles drying at room temperature under 13% relative humidity.

Overall, these results demonstrate that multiple rounds of replication are possible through wet-dry cycling. The drying conditions tend to affect the final yield after a few replication rounds, constant heating and long-term experiments tend to result in degradation of template and products. However, in the early Earth, there was likely a constant production of new strands that could counteract their degradation.

4.9 Replication of an RNA system †

If there was an RNA world or another pre-RNA informational polymer[77], the system we have developed should still work and be applicable with different informational

† This part of the project was done in collaboration with Bryce Clifton.

polymers. Since information transfer from one strand of a DNA duplex has been demonstrated, replication of both strands of an RNA duplex should be possible as well.

The same template, oligomers and protocol for replication used in Chapter 3 were used for this assay, with an adapted protocol from the DNA replication experiments. Briefly, the RNA was dried and resuspended in UAcW (1/2/4), followed by heat cycling and aqueous dilution to promote assembly and ligation respectively. The ligation products obtained from the solvent composed of urea and acetamide are comparable to the results obtained in glycholine (Figure 42), and the yield appears to be better in the new solvent. These results suggest that such system is applicable to DNA and RNA, and therefore could be extended to similar informational polymers that might have been present before RNA.

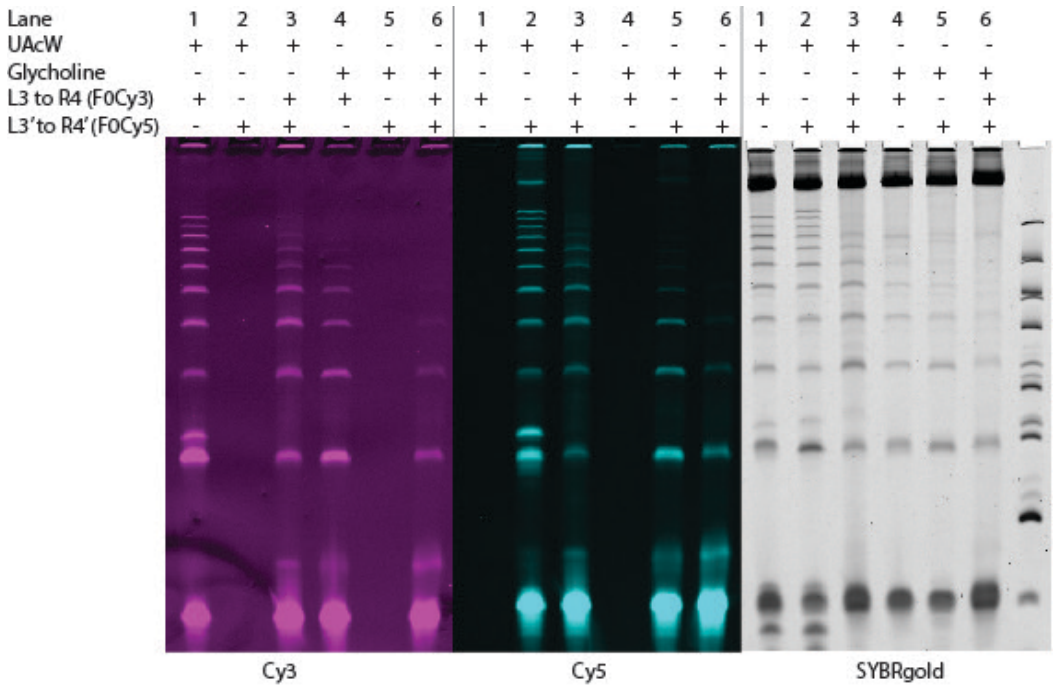


Figure 42. Replication of a 320 bp fragment from the 613 bp RNA template enabled by glycholine and UAcW (1/2/4).

Finally, to demonstrate that replication cycles are also possible with RNA in UAcW, a similar process to the one followed with DNA was executed. Initially, a test with and without addition of the ligase was done as a first approach to optimize the replication cycles, similar to the one done with the DNA system. As seen in Figure 43, there is no clear trend of increase in product yield for the full length (320 nt) product, however, if the 256 nt product is used as reference, there appears to be an increase when the ligase is added in each cycle with ATP and DTT. However, it is important to note that results for this system are less consistent as compared to the DNA system. This could be related to the purity of the oligomers used[11] and the higher tendency of RNA to hydrolyse. The T4 RNA ligase 2 used for RNA ligation is reported to be heat deactivated at 80 °C for 5 min as indicated by the manufacturer, which explains why adding it in each cycle would be necessary to get better yields.

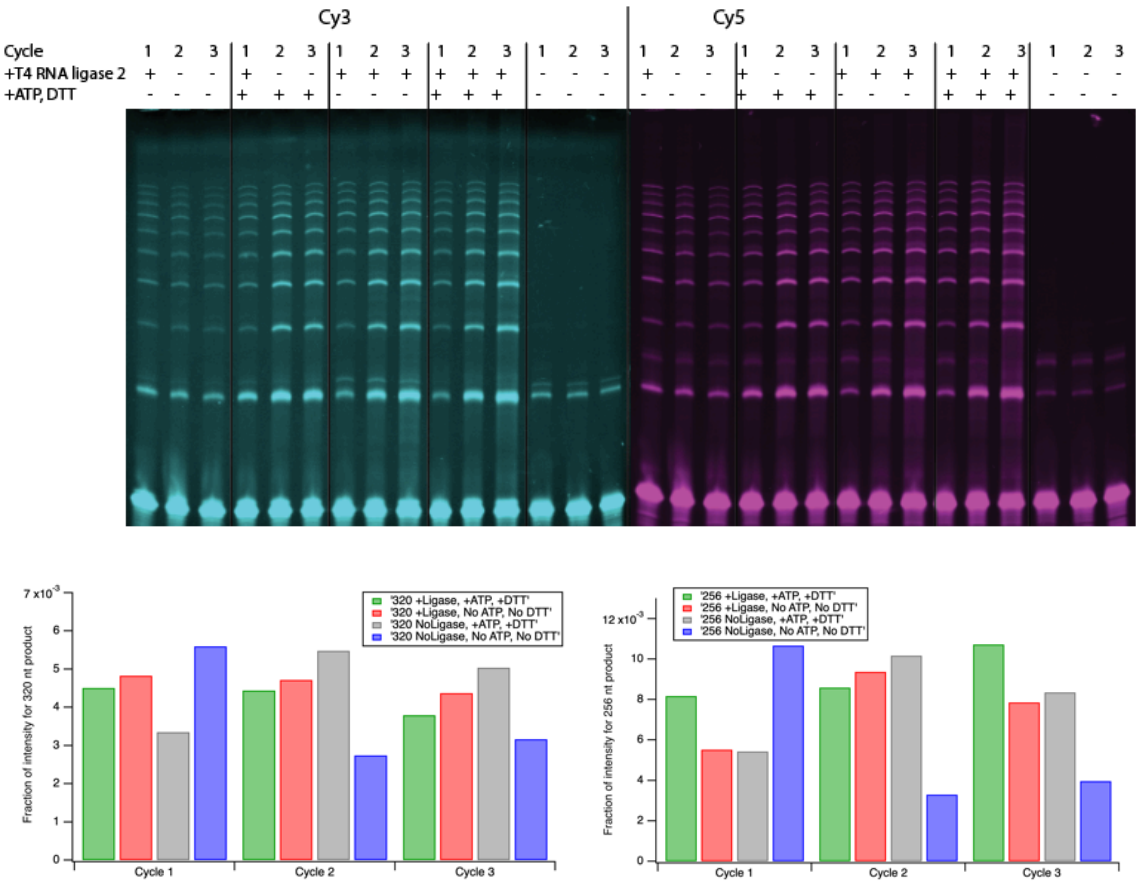


Figure 43. Denaturing gel and densitometry analysis from 3 replication cycles for the Cy3 labelled 320 nt and 256 nt RNA replication products using T4 RNA ligase 2, with and without addition of the ligase, ATP and DTT in each cycle during the rehydration step.

Based on the results obtained, addition of ligase, ATP and DTT within each cycle appear to be the best working conditions for multiple cycles. Therefore, a duplicate of this experiment was performed, with three wet-dry cycles to verify its reproducibility and quantify the ligated products. Again, the trend is not as obvious as with the DNA system, but there is a slight increase of the ligated products after the third cycle.

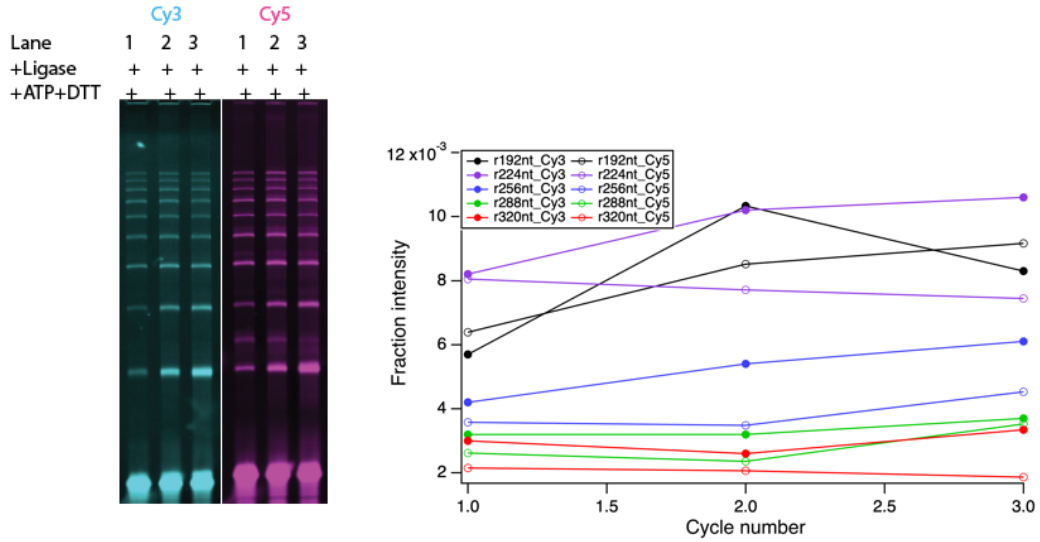


Figure 44. Replication cycles of a 320 nt fragment from a 613 bp RNA duplex.

To further explore RNA replication, a shorter double stranded template of 192 bp was transcribed and purified (the same sequences used in Chapter 3). This template was a short fragment from the 613 bp template used in previous experiments, so the same 32 nt oligomers could be used as substrates. The same protocol was used to run up to 5 replication cycles. As shown in Figure 45, there is an increase in the yield of ligated products with each cycle, and no degradation is observed.

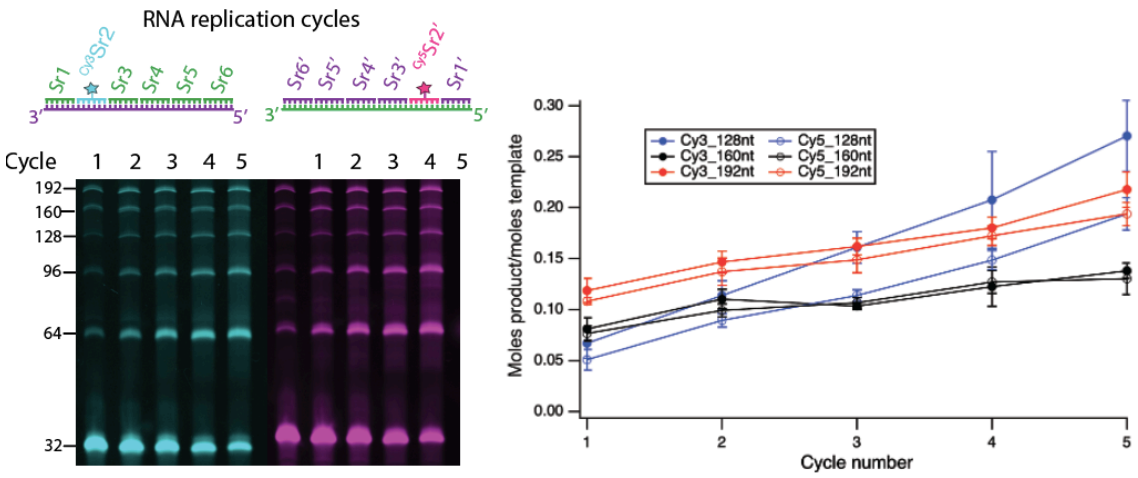


Figure 45. Five replication rounds of a 192 bp RNA duplex promoted through wet-dry cycling in UAcW.

4.10 Conclusions

A new solvent composed of prebiotic molecules was found to aid in the replication of DNA and RNA. This solvent enables solubilization of nucleic acids even under high salt concentrations containing divalent cations and lowers the melting temperature of the duplexes. This enables the use of mild conditions for strand denaturation and reannealing, simulating a more prebiotic environment. Furthermore, replication occurs under different solvent compositions and through different drying methods, demonstrating the robustness of this system.

Multiple replication cycles are possible through wet-dry cycling, increasing the product formation in each cycle, without the need to remove divalent cations such as Mg^{2+} for the dry states. Such a system could have been formed on the early Earth through recurring cycles of wet-dry and hot-cold conditions throughout the day or the season.

Components such as urea, acetamide, and salts could have been present 3.5 billion years ago and provide the adequate conditions for replication of informational polymers even in the presence of salts such as NaCl, $MgCl_2$ and $CaCl_2$.

4.11 Materials and methods

4.11.1 Preparation of UAcW

Urea and acetamide were purchased from Fisher Scientific and used without further purification. A mixture of urea, acetamide and 18.3 MΩ/cm water (Barnstead Nanopure™) in 1:2:4 molar ratio was prepared in a closed conical tube. The mix was then heated up to 80 °C and vortexed constantly until a homogeneous solution was formed.

4.11.2 DNA

The 641 DNA template was purchased as a gBlock from IDT. The sequence was taken from the plasmid pBluescript II SK(-), a, phagemid excised from lambda ZAPII (GenBank accession: X52330.1). This plasmid was chosen as a model of a mixed, arbitrary sequence (i.e., not designed to have a particular sequence or base pair composition). PCR of this sequence was performed using NEB Q5 Hot Start High Fidelity DNA polymerase, using the primers indicated in Table 10. Amplicons were then PAGE purified in native conditions (0.5x TBE, 7.5 % polyacrylamide).

The oligomers used for binding and ligation were ordered from IDT and purified through PAGE under denaturing conditions (0.5x TBE, 6 M urea, 15 % polyacrylamide).

4.11.2.1 Sequences

Table 9. 641 bp Template

641 bp DNA template	5'GATCGATCTCGCCCGCGAAATAATACGACTCACTATAGGGGAATAGTGTA TGCGGCGACCGAGTTGCTCTTGCCCGGCGTCGCAGGTACATCCAGCTGATG AGTCCCAAATAGGACGAAATGCCGAGCTCGGCGTCAATACGGGATAATACC GCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCG GGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAA CCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTT CTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAG GGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGA AGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTT AGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCAC CTGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGC GCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTT CTTCCCTTCCTTTCTCGCCACGTTCGCG-3'

Primers used for PCR of 641 bp DNA template

Table 10. Primers used for PCR of template

IVT_r613_S_FWD	GATCGATCTCGCCCGCGAAA
IVT_r613_S_REV	CGCGAACGTGGCGAGAAAGGA

10 sense and 10 anti sense oligomers were used to demonstrate replication

Table 11. Oligomers used for binding and ligation

S1	/5Phos/ AGAACTTTAAAAGTGCTCATCATTGGAAAACG
S2	/5Phos/ TTCTTCGGGGCGAAAACTCTCAAGGATCTTAC
S3	/5Phos/ CGCTGTTGAGATCCAGTTCGATGTAACCCACT
S4	/5Phos/ CGTGCACCCAACTGATCTTCAGCATCTTTTAC
^{Cy5}S5	/5Phos/ TTTCACCAGCGTTTCT/iCy5/GGGTGAGCAAAAACAG
S6	/5Phos/ GAAGGCAAAATGCCGCAAAAAAGGGAATAAGG
S7	/5Phos/ GCGACACGGAAATGTTGAATACTCATACTCTT
S8	/5Phos/ CCTTTTTCAATATTATTGAAGCATTTATCAGG
S9	/5Phos/ GTTATTGTCTCATGAGCGGATACATATTTGAA
S10	/5Phos/ TGTATTTAGAAAAATAAACAAATAGGGGTTCC
R1'	/5Phos/ CGTTTTCCAATGATGAGCACTTTTAAAGTTCT
R2'	/5Phos/ GTAAGATCCTTGAGAGTTTTCGCCCCGAAGAA
R3'	/5Phos/ AGTGGGTTACATCGAACTGGATCTCAACAGCG
R4'	/5Phos/ GTAAAAGATGCTGAAGATCAGTTGGGTGCACG
^{Cy3}R5'	/5Phos/ CTGTTTTTGCTCACC/iCy3/CAGAAACGCTGGTGAAA
R6'	/5Phos/ CCTTATTCCCTTTTTTGCGGCATTTTGCCTTC
R7'	/5Phos/ AAGAGTATGAGTATTCAACATTTCCGTGTCGC
R8'	/5Phos/ CCTGATAAATGCTTCAATAATATTGAAAAAGG
R9'	/5Phos/ TTCAAATATGTATCCGCTCATGAGACAATAAC
R10'	/5Phos/ GGAACCCCTATTTGTTTATTTTTCTAAATACA

4.11.3 Replication in UAcW with varying compositions

Samples were prepared by mixing 0.76 pmoles of the 641bp DNA template (300 ng) with each of the oligomers in a 4:1 molar ratio with respect to the template (5' phosphorylated to allow ligation with the T7 DNA Ligase). After drying under vacuum centrifugation for 2 hours, the samples were resuspended by adding 2 μL of UAcW and incubating at 35 °C for 1-2 hours. Samples were then heated up to 65 °C for 2 min (or 95 °C for aqueous buffer) and cooled down to 5 °C at a

constant rate over 20 min (-0.5 °C/8 s). Samples were diluted with water, 10x T4 DNA ligase buffer (to a concentration of 1x: 50 mM Tris-HCl , 10 mM $MgCl_2$, 1 mM ATP , 10 mM DTT, pH 7.5 @ 25°C) and 3000 U of T7 DNA ligase (New England Biolabs) to a final volume of 20 μL. After incubating for 1 hour at 25 °C, 6 μL of each reaction were mixed with 6 μL of 2x loading dye (concentration at 2x is: 95% formamide, 0.01% SDS, 0.5 mM EDTA pH 8.0). Samples were heated to 95 °C for 2 min and then placed on ice before loading into a 10% denaturing polyacrylamide gel (1x TBE, 8 M urea). The gel was pre-run for at least 30 min at 14 W, 45 V/cm before loading the samples. The gel was run at these same conditions for approximately 1 h and imaged on a Typhoon FLA 9500 laser scanner (GE Healthcare) at a resolution of 50 μm. Cy3 fluorescence was measured using the bandpass green (BPG1) emission filter, with an excitation wavelength of 532 nm; and the Cy5 fluorescence was measured using the long pass red (LPR) emission filter with an excitation wavelength of 635 nm. The gel was then stained with SYBR gold for 15 min and washed for another 15 min before scanning with the long pass blue filter (LPB) with an excitation wavelength of 473 nm.

4.11.4 Replication in UAcW with different drying conditions

Samples were prepared by mixing 0.76 pmoles of the 641bp DNA template with each of the oligomers in a 4:1 molar ratio with respect to the template (5' phosphorylated to allow ligation with the T7 DNA Ligase), 2 μL of 500 mM Tris-HCl, 100 mM $MgCl_2$, 2 μL of UAcW (1/2/4) and water to a final volume of 25 μL. Samples were then placed either in the 65 °C oven or in a chamber with controlled relative humidity for the indicated amount of time to remove water. Then they were rehydrated by addition of 2 μL of 10 mM ATP and 100 mM DTT, 3000 U of T7 DNA ligase (New

England Biolabs) and water to a final volume of 20 μL. Samples were left to react for 1 hour at 25 °C, and the same protocol as indicated above was followed to run the gel and analyze the samples.

4.11.5 Multiple replication rounds

Samples were prepared by mixing 1.14 pmoles of the 641bp DNA template (450 ng) with each of the oligomers in a 4:1 molar ratio with respect to the template (5' phosphorylated to allow ligation with the T7 DNA Ligase), 3 μL of 500 mM Tris-HCl and 100 mM $MgCl_2$, 3 μL of UAcW (1/2/4) and water to a final volume of 25 μL. For the first cycle, samples were dried for 30 min under vacuum, or 2.25 hours at 65 °C and then rehydrated by adding 3 μL of 10 mM ATP and 100 mM DTT, 3000 U of T7 DNA ligase (New England Biolabs) and water to a final volume of 30 μL. After leaving the samples to react for 1 hour, a 5 μL aliquot was taken, mixed with the loading buffer and stored in the freezer until all samples were ready to run in the gel. For subsequent cycles, a similar procedure was followed, adjusting the volume 25 μL before drying, then rehydrating to de adequate volume to keep the concentration of all components constant. For rehydration after the first cycle, only water was added (no ligase or buffer were added after cycle 1) with enough ATP and DTT to restore half the amount of what was present at the beginning of the reaction. Aliquots of 5 μL were taken between each cycle to further analyze in a polyacrylamide gel as indicated above.

4.11.6 Densitometry analysis

To estimate the total moles of product formed, the fraction of intensity from each band was divided by the fraction of the sequences containing the fluorophore label for each product size, and then multiplied by the total number of moles of each 32mer. For example, the fraction of total intensity would be divided by 1 for the 320 product and then multiplied by 4.49 pmoles which were initially added to the mixture, whereas for the 160 nt product the fraction of intensity would be divided by 0.83 (because 5/6 products formed would contain the fluorophore) and then divided by 4.49 pmoles. Then the resulting value could be divided by the total moles of template used to obtain the yield for each length.

CHAPTER 5. CONCLUSIONS

Within origins of life research, there are many areas of study that focus on different aspects, such as the origins of building blocks that could then have evolved into polymers, the formation of polypeptides[78,79] and nucleic acid oligomers[80], formation of protocells[81] and compartmentalization[82], nucleic acid replication[12], etc. Throughout this work, the importance of the environment to promote kinetic trapping of informational polymers and their replication has been highlighted, but there are many questions that remain unanswered.

The RNA world hypothesis states that RNA could have served two functions: information storage and catalysis. Even though there are naturally occurring ribozymes, how they were selected by nature to perform their current functions is still a mystery. Artificial ribozymes and DNAzymes have been selected in-vitro, under very stringent conditions to perform one particular function and through multiple rounds of selection. None of these conditions simulate a prebiotic environment that could provide insight on how functional sequences might have formed initially. Related to this problem is the fact that information tends to be lost as it gets transferred multiple times, as demonstrated by Spiegelman's experiment[83], in which replication of shorter sequences outcompeted replication of the long ones. Proposed ways to avoid this problem are related to environmental conditions, one of them uses a flow system with a thermal gradient that leads to accumulation of longer sequences that could later be replicated[22]; another proposed solution involves transient compartmentalization in microdroplets to diminish the replication of the short parasite sequences[84]. Our proposed environment doesn't

specifically target this problem, but we have demonstrated that with the aid of a low-water environment, templates of different lengths can be successfully replicated. In addition, structure of the templates, and not only the length, might also have played an important role for initial selection and replication of functional sequences.

Another long-standing problem in replication is strand inhibition. This problem has prevented multiple replication rounds from occurring in multiple studies. The main approach that has been taken to solve this problem involves quick temperature changes[8,21] for PCR-like reactions and de-stabilization of the duplex by shortening strands or adding G:U wobble pairs[17]. We demonstrated that the strand inhibition problem can be overcome by the use of alternative solvents with no water or low water content, with mild heating and cooling conditions. This was demonstrated by the replication of DNA and RNA sequences in glycholine and UAcW, and multiple replication rounds were possible in the urea and acetamide based solvent.

An additional problem associated to replication within the RNA world hypothesis is the replicator-catalyst paradox[20]. Structured sequences tend to display more problems for information transfer; however, internal structure is needed for a sequence to have function and fulfil its dual role as proposed by the RNA world hypothesis[85]. In our system, replication of long and structured sequences was possible with glycholine. We demonstrated that internal structure promotes trapping of functional sequences and does not necessarily hinder replication of the strands. In addition, the transition between the duplex and the trapped sequences can be controlled by changing environmental conditions, such as cooling rate and water content of the solvent. Importantly, glycholine favours kinetic trapping of active DNA and RNA sequences in their active conformations despite

the presence of their complementary strands as demonstrated with the DNAzyme, the minimal hammerhead ribozyme and the hammerhead ribozyme embedded in a longer template strand. Nonetheless, we did observe a lower yield of information transfer with the most structured sequence, suggesting that there should be a degree of internal structure that can favour function, without hindering replication.

For the replication process, we used a ligase as a robust way of ligation of the assembled oligomers. Enzyme-free replication has been widely studied, but there is still not a method that could be universally used with as much efficiency as the ligases used in this project, representing another big challenge in origins of life research. The use of compounds that activate the phosphate group of a nucleotide, such as EDC, have been proposed as a way to do primer extension and ligate oligomers[21,86]; however, they result in multiple side products and are very sensitive to the structure and sequence of the nucleic acids being ligated[87]. Ribozymes have also been evolved to work as ligases and polymerases[8,15,19,88]; nonetheless, multiple turnover is hard to achieve, and the yields are low, in particular when long and structured sequences are being replicated. This demonstrates that even though significant advances have been made within prebiotic nucleic acid replication, there are still some challenges that need to be addressed.

We first hypothesized that viscosity was playing a central role in the kinetic trapping of the structured sequences and long templates. The results obtained for replication and trapping of functional sequences in chapters 2 and 3 fitted with this hypothesis. However, when the UAcW solvent was explored, we realized that there were other factors that might be involved in the process, since the viscosity UAcW is close to the viscosity of water. Some factors that might be affecting this process could be the

lowering in the melting temperature, and it has been reported that the rate of strand invasion is faster in solvents containing amides[75]. Therefore, further research needs to be done to explain why some solvents with low viscosity such as UAcW can enable trapping and replication of DNA and RNA.

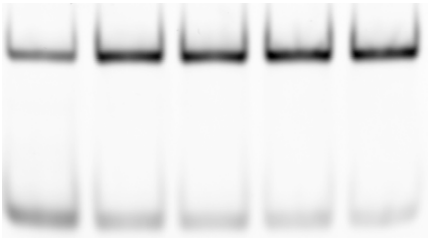
Finally, some of the principles that govern the trapping of active sequences presented in this dissertation could be further applied outside of origins of life research. Single-stranded DNA production during aptamer selection can be an expensive and laborious task. For example, sometimes a chemical modification is added to the primers so they run different on a gel or so they can bind to biotin, which can raise the cost of production and bias the selection. Digestion with lambda exonuclease is also associated with high costs, and incomplete digestion leads to accumulation of dsDNA[50,89]. Glycholine provides a new and easy way of trapping the active sequences without the need of removing the complementary sequences. Moreover, recently, reline, another viscous eutectic solvent was used for in vitro selection of aptamers that could be used for gluten detection in food[29], providing the first example of an aptamer selected specifically to work in a non-aqueous solvent.

APPENDIX A. ADDITIONAL DATA AND FIGURES.

Not all experimental data was added in the thesis to make it easier to read and follow. Within this section are some supplementary figures of experiments that were done, as well as the gels used for some of the plots displayed throughout this document.

Supplementary figures from Chapter 2.

Short I-R3					
0	0.5	1	2	4	time (h)
44.11	71.59	75.88	80.31	82.27	%Intermolecular duplex
55.89	28.41	24.12	19.69	17.73	%Intramolecular fold



Long-Cy3					L-Hairpin					
0	0.5	1	2	4	0	0.5	1	2	4	time (h)
36.84	38.23	38.39	36.76	38.40	25.84	27.58	26.54	25.40	26.47	Intermolecular duplex
63.16	61.77	61.61	63.24	61.60	74.16	72.42	73.46	74.60	73.53	Intramolecular fold

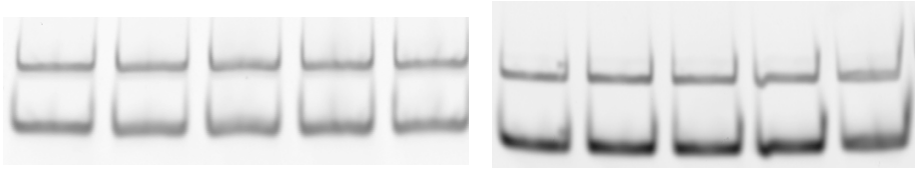


Figure 46. Sample gels from Figure 4 that show kinetic trapping of the single strands that form intramolecular structures. Samples were heated to 80 °C for 5 min and cooled down fast (-48 °C/min).

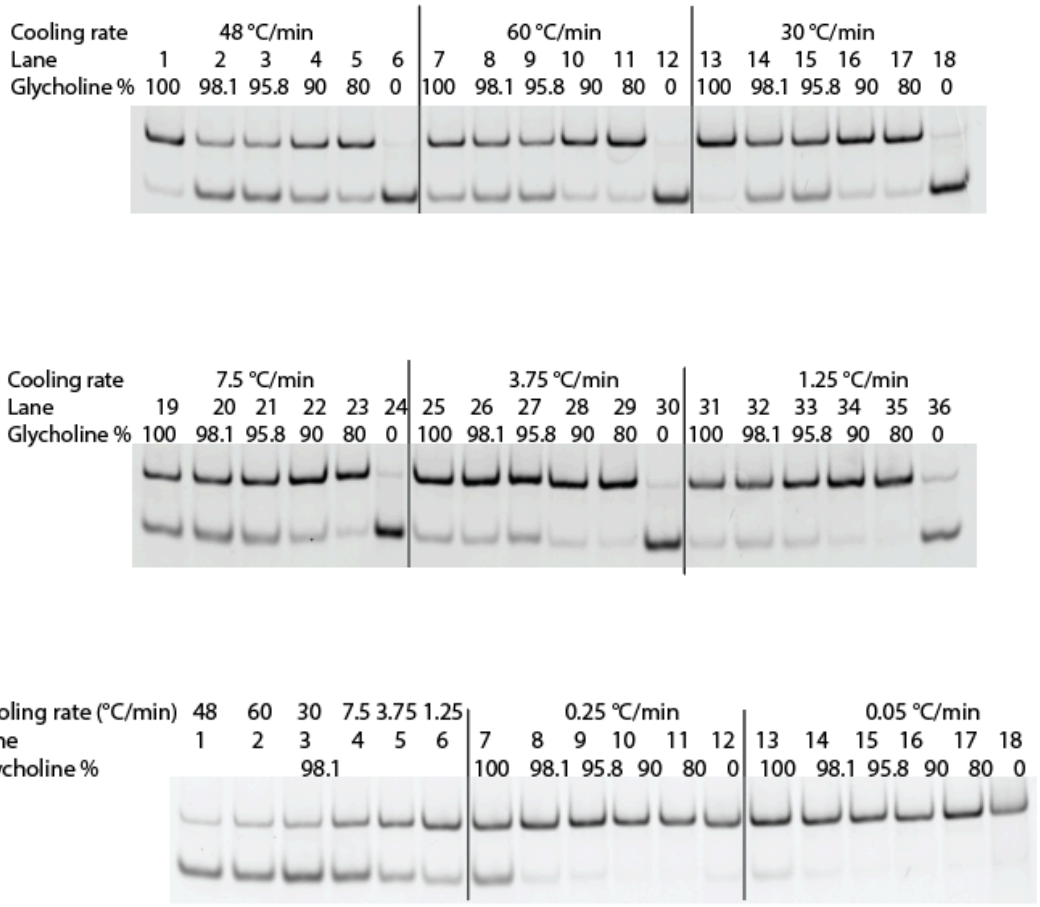


Figure 47. Raw gels shown for plot displayed Figure 7 for I-R3 DNAzyme. Kinetic trapping of the single strands followed by FAM with different cooling rates and water content.

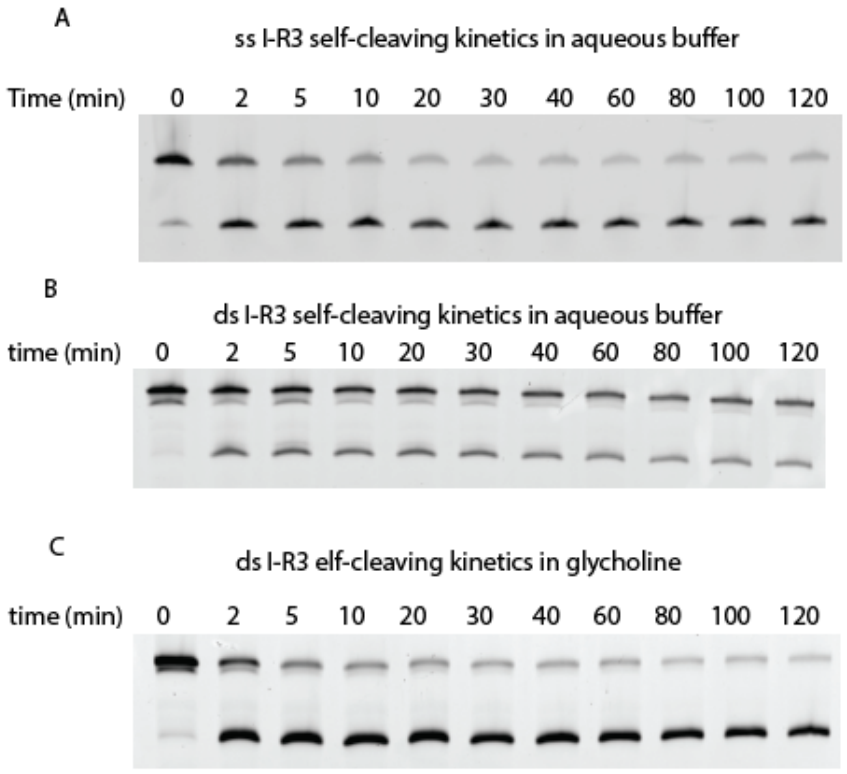


Figure 48. Full set of gels from Figure 8 for the long stem I-R3. Self-cleaving kinetics in aqueous buffer and glycholine after heating and cooling to room temperature at -0.5°C/min

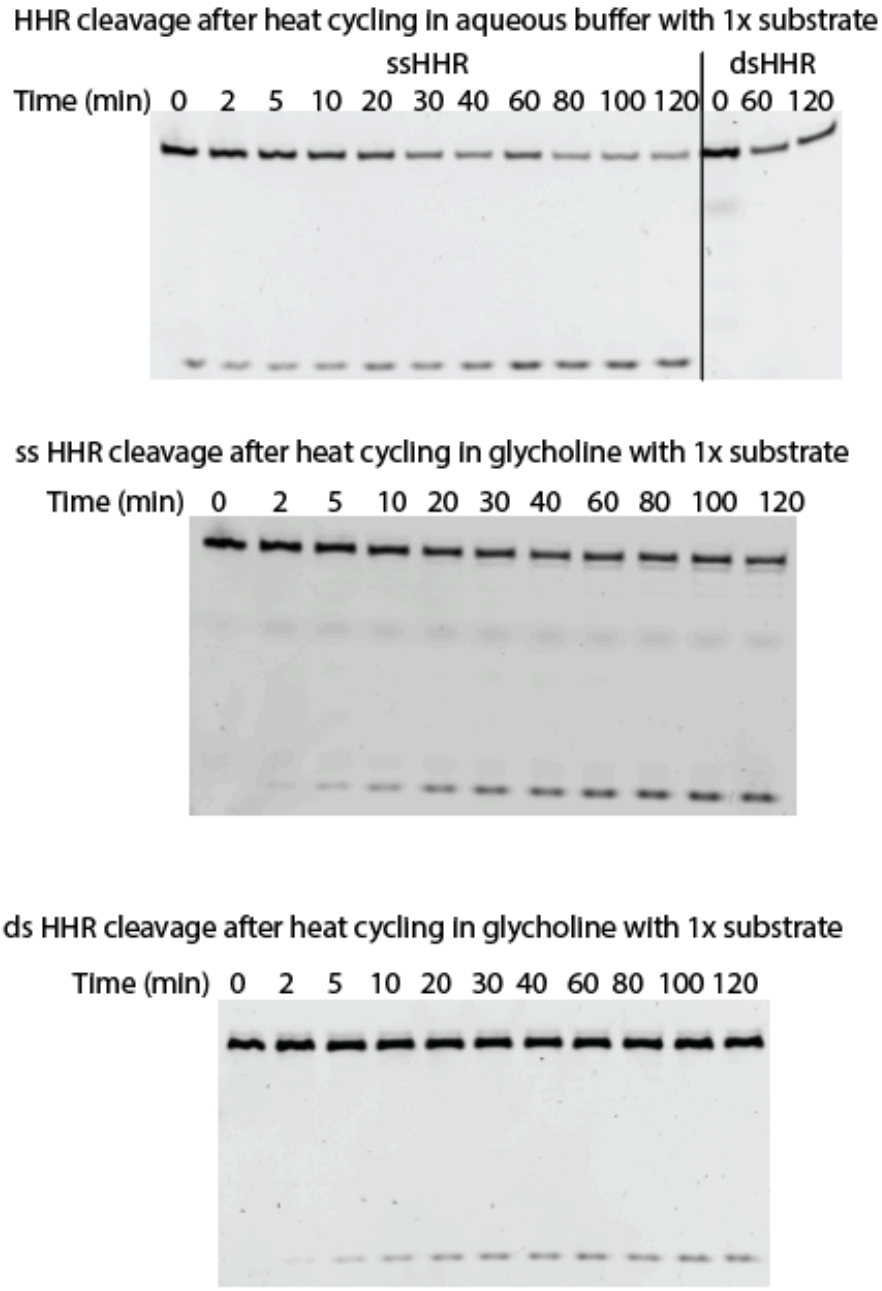


Figure 49. Gels following cleavage kinetics of the hammerhead ribozyme substrate after heat cycling in aqueous buffer and glycholine with 1:1 molar ratio of substrate to ribozyme. Densitometry analysis is plotted on Figure 14.

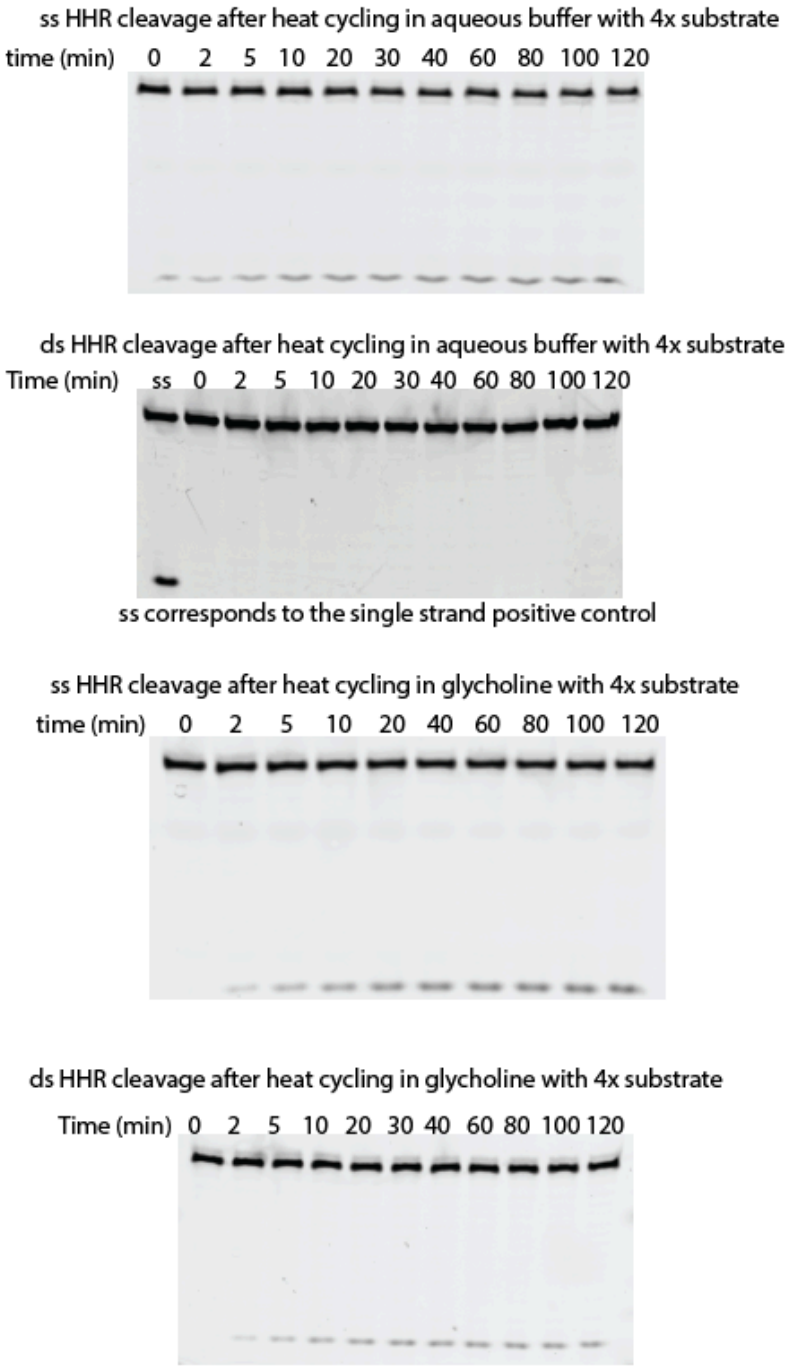


Figure 50. Gels following cleavage kinetics of the hammerhead ribozyme substrate after heat cycling in aqueous buffer and glycholine with 4:1 molar ratio of substrate to ribozyme. Densitometry analysis is plotted on Figure 14

REFERENCES

1 Nelson, D. E. & Cox, M. M. *Lehninger Principles of Biochemistry*. 3rd edn, (Worth Publishers, 2000).

2 Doudna, J. A. & Cech, T. R. The chemical repertoire of natural ribozymes. *Nature* **418**, 222-228 (2002).

3 Hammann, C., Luptak A Fau - Perreault, J., Perreault J Fau - de la Peña, M. & de la Peña, M. The ubiquitous hammerhead ribozyme.

4 Fedor, M. J. Structure and function of the hairpin ribozyme.

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

6 Gilbert, W. Origin of Life - the Rna World. *Nature* **319**, 618-618 (1986).

7 Been, M. D. & Cech, T. R. RNA as an RNA polymerase: net elongation of an RNA primer catalyzed by the Tetrahymena ribozyme.

8 Horning, D. P. & Joyce, G. F. Amplification of RNA by an RNA polymerase ribozyme. *Proceedings of the National Academy of Sciences* **113**, 9786-9791, doi:10.1073/pnas.1610103113 (2016).

9 Deck, C., Jauker, M. & Richert, C. Efficient enzyme-free copying of all four nucleobases templated by immobilized RNA. *Nature Chemistry* **3**, 603-608, doi:10.1038/nchem.1086 (2011).

10 Prywes, N., Blain, J. C., Del Frate, F. & Szostak, J. W. Nonenzymatic copying of RNA templates containing all four letters is catalyzed by activated oligonucleotides. *eLife* **5**, e17756, doi:10.7554/eLife.17756 (2016).

11 He, C., Gallego, I., Laughlin, B., Grover, M. A. & Hud, N. V. A viscous solvent enables information transfer from gene-length nucleic acids in a model prebiotic replication cycle. *Nature Chemistry* **9**, 318-324, doi:10.1038/nchem.2628 (2017).

12 Szostak, J. W. The eightfold path to non-enzymatic RNA replication. *Journal of Systems Chemistry* **3**, 2, doi:10.1186/1759-2208-3-2 (2012).

13 Dose, C., Ficht, S. & Seitz, O. Reducing product inhibition in DNA-template-controlled ligation reactions. *Angewandte Chemie-International Edition* **45**, 5369-5373, doi:10.1002/anie.200600464 (2006).

14 Grossmann, T. N., Strohbach, A. & Seitz, O. Achieving Turnover in DNA-Templated Reactions. *Chembiochem* **9**, 2185-2192, doi:10.1002/cbic.200800290 (2008).

15 Attwater, J., Wochner, A. & Holliger, P. In-ice evolution of RNA polymerase ribozyme activity. *Nature Chemistry* **5**, 1011-1018, doi:10.1038/nchem.1781 (2013).

16 Joyce, G. F. & Orgel, L. E. Non-enzymatic template-directed synthesis on RNA random copolymers: Poly(C,A) templates. *Journal of Molecular Biology* **202**, 677-681, doi:http://dx.doi.org/10.1016/0022-2836(88)90297-5 (1988).

17 Zhou, L., O'Flaherty, D. K. & Szostak, J. W. Assembly of a Ribozyme Ligase from Short Oligomers by Nonenzymatic Ligation. *Journal of the American Chemical Society* **142**, 15961-15965, doi:10.1021/jacs.0c06722 (2020).

18 Wochner, A., Attwater, J., Coulson, A. & Holliger, P. Ribozyme-catalyzed transcription of an active ribozyme. *Science* **332**, 209-212, doi:10.1126/science.1200752 (2011).

19 Attwater, J., Raguram, A., Morgunov, A. S., Gianni, E. & Holliger, P. Ribozyme-catalysed RNA synthesis using triplet building blocks. *eLife* **7**, e35255, doi:10.7554/eLife.35255 (2018).

20 Ivica, N. A. *et al.* The Paradox of Dual Roles in the RNA World: Resolving the Conflict Between Stable Folding and Templating Ability. *Journal of Molecular Evolution* **77**, 55-63, doi:10.1007/s00239-013-9584-x (2013).

21 Edeleva, E. *et al.* Continuous nonenzymatic cross-replication of DNA strands with in situ activated DNA oligonucleotides. *Chemical Science* **10**, 5807-5814, doi:10.1039/C9SC00770A (2019).

22 Kreysing, M., Keil, L., Lanzmich, S. & Braun, D. Heat flux across an open pore enables the continuous replication and selection of oligonucleotides towards increasing length. *Nat. Chem.* **7**, 203-208, doi:10.1038/nchem.2155 http://www.nature.com/nchem/journal/v7/n3/abs/nchem.2155.html#supplementary-information (2015).

23 Gállego, I., Grover, M. A. & Hud, N. V. Folding and imaging of DNA nanostructures in anhydrous and hydrated deep-eutectic solvents. *Angew. Chem. Int. Ed.* **54**, 6765-6769, doi:10.1002/anie.201412354 (2015).

24 Mamajanov, I., Engelhart, A. E., Bean, H. D. & Hud, N. V. DNA and RNA in anhydrous media: duplex, triplex, and G-quadruplex secondary structures in a deep eutectic solvent. *Angew. Chem. Int. Ed.* **49**, 6310-6314, doi:10.1002/anie.201001561 (2010).

25 Lannan, F. M., Mamajanov, I. & Hud, N. V. Human telomere sequence DNA in water-free and high-viscosity solvents: G-quadruplex folding governed by Kramers rate theory. *Journal of the American Chemical Society* **134**, 15324-15330 (2012).

26 Tateishi-Karimata, H. & Sugimoto, N. Structure, stability and behaviour of nucleic acids in ionic liquids. *Nucleic Acids Res*, doi:10.1093/nar/gku499 (2014).

27 Nakano, S.-I. & Sugimoto, N. The structural stability and catalytic activity of DNA and RNA oligonucleotides in the presence of organic solvents. *Biophysical reviews* **8**, 11-23, doi:10.1007/s12551-015-0188-0 (2016).

28 Smith, E. L., Abbott, A. P. & Ryder, K. S. Deep Eutectic Solvents (DESs) and Their Applications. *Chemical Reviews* **114**, 11060-11082, doi:10.1021/cr300162p (2014).

29 Svigelj, R. *et al.* Selection of Anti-gluten DNA Aptamers in a Deep Eutectic Solvent. *Angewandte Chemie International Edition* **57**, 12850-12854, doi:10.1002/anie.201804860 (2018).

30 Harroun, S. G. *et al.* Programmable DNA switches and their applications. *Nanoscale* **10**, 4607-4641, doi:10.1039/C7NR07348H (2018).

31 Bath, J. & Turberfield, A. J. DNA nanomachines. *Nature Nanotechnology* **2**, 275-284, doi:10.1038/nnano.2007.104 (2007).

32 Alberts, B., Johnson, A., Lewis, J. & al., e. *Molecular Biology of the Cell*. 4th edition edn, (Molecular Biology of the Cell, 2002).

33 Liu, D. S. & Balasubramanian, S. A proton-fuelled DNA nanomachine. *Angewandte Chemie-International Edition in English* **42**, 5734-5736 (2003).

34 Wang, C., Huang, Z., Lin, Y., Ren, J. & Qu, X. Vol. 22 2792–2798 (Advanced Materials, 2010).

35 Viasnoff, V., Meller, A. & Isambert, H. DNA nanomechanical switches under folding kinetics control. *Nano Lett.* **6**, 101-104, doi:10.1021/nl052161c (2006).

36 Yurke, B., Turberfield, A. J., Mills, A. P., Simmel, F. C. & Neumann, J. L. A DNA-fuelled molecular machine made of DNA. *Nature* **406**, 605-608 (2000).

37 Turberfield, A. J. *et al.* DNA fuel for free-running nanomachines. doi:10.1103/PhysRevLett.90.118102. (2003).

38 Gu, H., Furukawa, K., Weinberg, Z., Berenson, D. F. & Breaker, R. R. Small, highly active DNAs that hydrolyze DNA. *Journal of the American Chemical Society* **135**, 9121-9129, doi:10.1021/ja403585e (2013).

39 Fedor, M. J. & Williamson, J. R. The catalytic diversity of RNAs. *Nature Reviews Molecular Cell Biology* **6**, 399-412, doi:10.1038/nrm1647 (2005).

40 Tang, J. & Breaker, R. R. Structural diversity of self-cleaving ribozymes. *Proceedings Of The National Academy Of Sciences Of The United States Of America* **97**, 5784-5789 (2000).

41 Canny, M. D. *et al.* Vol. 126 10848-10849 (Journal of the American Chemical Society, 2004).

42 Scott, W. G., Finch, J. T. & Klug, A. The crystal structure of an all-RNA hammerhead ribozyme: a proposed mechanism for RNA catalytic cleavage. *Cell* **81**, 991-1002 (1995).

43 Achenbach, J. C., Chiuman W Fau - Cruz, R. P. G., Cruz Rp Fau - Li, Y. & Li, Y. DNAzymes: from creation in vitro to application in vivo.

44 Hollenstein, M. 20777-20804. (Molecules, 2015).

45 Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research* **31**, 3406-3415, doi:10.1093/nar/gkg595 (2003).

46 Ma, L., Liu, B., Huang, P.-J. J., Zhang, X. & Liu, J. DNA Adsorption by ZnO Nanoparticles near Its Solubility Limit: Implications for DNA Fluorescence Quenching and DNAzyme Activity Assays. *Langmuir* **32**, 5672-5680, doi:10.1021/acs.langmuir.6b00906 (2016).

47 Lilley, D. M. J. Folding and catalysis by the hairpin ribozyme. *FEBS Letters* **452**, 26-30, doi:https://doi.org/10.1016/S0014-5793(99)00544-X (1999).

48 Martick, M. & Scott, W. G. Tertiary Contacts Distant from the Active Site Prime a Ribozyme for Catalysis. *Cell* **126**, 309-320, doi:10.1016/j.cell.2006.06.036 (2006).

49 Müller, U. F. Re-creating an RNA world. *Cellular and Molecular Life Sciences CMLS* **63**, 1278-1293, doi:10.1007/s00018-006-6047-1 (2006).

50 Marimuthu, C., Tang, T. H., Tominaga, J., Tan, S. C. & Gopinath, S. C. Vol. 137 1307-1315 (Analyst., 2012).

51 Ferris, J. P., Hill, A. R., Liu, R. H. & Orgel, L. E. Synthesis of long prebiotic oligomers on mineral surfaces. *Nature* **381**, 59-61 (1996).

52 He, C., Lozoya-Colinas, A., Gallego, I., Grover, M. A. & Hud, N. V. Solvent viscosity facilitates replication and ribozyme catalysis from an RNA duplex in a model prebiotic process. *Nucleic Acids Res* **47**, 6569-6577, doi:10.1093/nar/gkz496 (2019).

53 sheely, M. L. Glycerol Viscosity Tables. *Industrial & Engineering Chemistry* **24**, 1060-1064, doi:10.1021/ie50273a022 (1932).

54 Ferreira, A. G. M. *et al.* The viscosity of glycerol. *The Journal of Chemical Thermodynamics* **113**, 162-182, doi:https://doi.org/10.1016/j.jct.2017.05.042 (2017).

55 Yadav, A. & Pandey, S. Densities and Viscosities of (Choline Chloride + Urea) Deep Eutectic Solvent and Its Aqueous Mixtures in the Temperature Range 293.15 K to 363.15 K. *Journal of Chemical & Engineering Data* **59**, 2221-2229, doi:10.1021/je5001796 (2014).

56 Agieienko, V. & Buchner, R. Densities, Viscosities, and Electrical Conductivities of Pure Anhydrous Reline and Its Mixtures with Water in the Temperature Range (293.15 to 338.15) K. *Journal of Chemical & Engineering Data* **64**, 4763-4774, doi:10.1021/acs.jced.9b00145 (2019).

57 Priyakumar, U. D., Hyeon, C., Thirumalai, D. & Mackerell, A. D., Jr. Urea destabilizes RNA by forming stacking interactions and multiple hydrogen bonds with nucleic acid bases. *Journal of the American Chemical Society* **131**, 17759-17761, doi:10.1021/ja905795v (2009).

58 Lim, W. K., Rösgen, J. & Englander, S. W. Urea, but not guanidinium, destabilizes proteins by forming hydrogen bonds to the peptide group. *Proceedings of the National Academy of Sciences* **106**, 2595, doi:10.1073/pnas.0812588106 (2009).

59 Kaiser, R. I., Maity, S. & Jones, B. M. Synthesis of Prebiotic Glycerol in Interstellar Ices. *Angewandte Chemie International Edition* **54**, 195-200, doi:10.1002/anie.201408729 (2015).

60 Forsythe, J. G. *et al.* Ester-Mediated Amide Bond Formation Driven by Wet-Dry Cycles: A Possible Path to Polypeptides on the Prebiotic Earth. *Angewandte Chemie-International Edition* **54**, 9871-9875, doi:10.1002/anie.201503792 (2015).

61 Becker, S. *et al.* Wet-dry cycles enable the parallel origin of canonical and non-canonical nucleosides by continuous synthesis. *Nature Communications* **9**, 163, doi:10.1038/s41467-017-02639-1 (2018).

62 Miller, S. L., Urey, H. C. & Oró, J. Origin of organic compounds on the primitive earth and in meteorites. *Journal of Molecular Evolution* **9**, 59-72, doi:10.1007/bf01796123 (1976).

63 He, C., Lozoya, A., Gallego, I., Grover, M. A. & Hud, N. Solvent viscosity facilitates replication and ribozyme catalysis from an RNA duplex in a model prebiotic process. *Nucleic Acids Research* **47**, 6569–6577 (2019).

64 Abbott, A. P., Barron, J. C., Ryder, K. S. & Wilson, D. Eutectic-based ionic liquids with metal-containing anions and cations. *Chemistry-a European Journal* **13**, 6495-6501, doi:10.1002/chem.200601738 (2007).

65 Feng, M. *et al.* One-step preparation of an antibacterial chitin/Zn composite from shrimp shells using urea-Zn(OAc)2·2H2O aqueous solution. *Green Chemistry* **20**, 2212-2217, doi:10.1039/C8GC00767E (2018).

66 Abbott, A., Capper, G. & Gray, S. Vol. 7 803-806 (Chem Phys Chem, 2006).

67 Bakirtzi, C., Triantafyllidou, K. & Makris, D. Novel lactic acid-based natural deep eutectic solvents: Efficiency in the ultrasound-assisted extraction of antioxidant polyphenols from common native Greek medicinal plants. *Journal of Applied Research on Medicinal and Aromatic Plants* **3**, 120-127, doi:10.1016/j.jarmap.2016.03.003 (2016).

68 Smith, E. L., Abbott, A. P. & Ryder, K. S. Deep Eutectic Solvents (DESs) and Their Applications. *Chem. Rev.*, doi:10.1021/cr300162p (2014).

69 Narayanan, N. S. V., Ashok Raj, B. V. & Sampath, S. Physicochemical, spectroscopic and electrochemical characterization of magnesium ion-conducting, room temperature, ternary molten electrolytes. *Journal of Power Sources* **195**, 4356-4364, doi:https://doi.org/10.1016/j.jpowsour.2010.01.070 (2010).

70 Butzow, J. J. & Eichhorn, G. L. Interactions of metal ions with polynucleotides and related compounds. IV. Degradation of polyribonucleotides by zinc and other divalent metal ions. *Biopolymers* **3**, 95-107, doi:10.1002/bip.360030110 (1965).

71 Ondier, G. O., Siebenmorgen, T. J. & Mauromoustakos, A. Low-temperature, low-relative humidity drying of rough rice. *Journal of Food Engineering* **100**, 545-550, doi:https://doi.org/10.1016/j.jfoodeng.2010.05.004 (2010).

72 Estrada, J. A. & Litchfield, J. B. HIGH HUMIDITY DRYING OF CORN: EFFECT ON DRYING RATE AND PRODUCT QUALITY. *Drying Technology* **11**, 65-84, doi:10.1080/07373939308916803 (1993).

73 Slonitsky, S. V., Maevsky, A. A., Mantulenki, V. B. & Frisman, E. V. Vol. 14 754-758 (Molekuliarnaia biologiia, 1979).

74 Blake, R. D. & Delcourt, S. G. Thermodynamic Effects of Formamide on DNA Stability. *Nucleic Acids Research* **24**, 2095-2103, doi:10.1093/nar/24.11.2095 (1996).

75 Dave, N. & Liu, J. Fast Molecular Beacon Hybridization in Organic Solvents with Improved Target Specificity. *The Journal of Physical Chemistry B* **114**, 15694-15699, doi:10.1021/jp106754k (2010).

76 Jones , C., Nomosatryo S. , Crowe S.A. , Bjerrum , C. J. & D.E. Canfield , D. E. Vol. 43 135-138 (Geology, 2015).

77 Hud, N. V., Cafferty, B. J., Krishnamurthy, R. & Williams, L. D. The origin of RNA and 'My Grandfather's Axe'. *Chem. Biol.* **20**, 466-474 (2013).

78 Campbell , T. D. *et al.* Vol. 10 4508 (2019).

79 Forsythe, J. G. *et al.* Collision cross section calibrants for negative ion mode traveling wave ion mobility-mass spectrometry. *Analyst* **14**, 6853-6861, doi:10.1039/c5an00946d (2015).

80 Ferris, J. P. & Ertem, G. Oligomerization of Ribonucleotides on Montmorillonite - Reaction of the 5'-Phosphorimidazolide of Adenosine. *Science* **257**, 1387-1389 (1992).

81 Mansy, S. S. *et al.* Template-directed synthesis of a genetic polymer in a model protocell. *Nature* **454**, 122-125, doi:10.1038/nature07018 (2008).

82 Frankel, E. A., Bevilacqua, P. C. & Keating, C. D. Polyamine/Nucleotide Coacervates Provide Strong Compartmentalization of Mg2+, Nucleotides, and RNA. *Langmuir* **32**, 2041-2049, doi:10.1021/acs.langmuir.5b04462 (2016).

83 Mills, D. R., Peterson, R. L. & Spiegelman, S. An extracellular Darwinian experiment with a self-duplicating nucleic acid molecule. *Proceedings of the National Academy of Sciences of the United States of America* **58**, 217-224, doi:10.1073/pnas.58.1.217 (1967).

84 Matsumura, S. *et al.* Transient compartmentalization of RNA replicators prevents extinction due to parasites. *Science* **354**, 1293, doi:10.1126/science.aag1582 (2016).

85 Gesteland, R. & Atkins, J. F. *The RNA World. The Nature of Modern RNA Suggests a Prebiotic RNA World*. (Cold Spring Harbor Laboratory Press, 1993).

86 Jauker, M., Griesser, H. & Richert, C. Copying of RNA Sequences without Pre-Activation. *Angewandte Chemie-International Edition* **54**, 14559-14563, doi:10.1002/anie.201506592 (2015).

87 Obianyor, C., Newnam, G., Clifton, B. E., Grover, M. A. & Hud, N. V. Towards Efficient Nonenzymatic DNA Ligation: Comparing Key Parameters for Maximizing Ligation Rates and Yields with Carbodiimide Activation. *ChemBioChem* **n/a**, doi:10.1002/cbic.202000335 (2020).

88 Flynn-Charlebois, A. *et al.* Deoxyribozymes with 2'-5' RNA ligase activity. *Journal of the American Chemical Society* **125**, 2444-2454 (2003).

89 Wang, T., Chen, C., Larcher, L. M., Barrero, R. A. & Veedu, R. N. Three decades of nucleic acid aptamer technologies: Lessons learned, progress and opportunities on aptamer development. *Biotechnology Advances* **37**, 28-50, doi:https://doi.org/10.1016/j.biotechadv.2018.11.001 (2019).