COMPUTATIONAL MODELING OF RNA REPLICATION IN AN RNA WORLD

COMPUTATIONAL MODELING OF RNA REPLICATION IN AN RNA WORLD

By ANDREW S. TUPPER, B.S.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment of the Requirements for the Degree Doctor of Philosophy

McMaster University	Doctor of Philosophy (2020)
Hamilton, Ontario	Biochemistry and Astrobiology

TITLE: Computational Modeling of RNA replication in an RNA world

AUTHOR: Andrew S. Tupper, B.S. (Rensselaer Polytechnic Institute)

SUPERVISOR: Dr. Paul G. Higgs

NUMBER OF PAGES: xiii, 140

Lay Abstract

The biology of modern life is complex and diverse, to such an extent that our origin is a mystery. There are, however, underlying themes of life which provide clues to our origin. Based on modern life's reliance on RNA polymers to carry out vital cellular roles, the origin of life likely passed through an RNA world. A time in which life was simpler and dependent on RNA as a genetic material and a catalyst. Due to the great uncertainties surrounding Earth's history, and the current inability to recreate an RNA organism in a laboratory, we turn to computational modeling. While still in its infancy, computational modeling has the power to let us explore times, conditions, and chemistries which are currently unreachable. In this thesis, we utilize computational modeling to provide insight and find solutions to the problems which plague the replication of RNA in an RNA world.

Abstract

The biology of modern life predicts the existence of an ancient RNA world. A phase of evolution in which organisms utilized RNA as a genetic material and a catalyst. However, the existence of an RNA organism necessitates RNA's ability to self-replicate, which has yet to be proven. In this thesis, we utilize computational modeling to address some of the problems facing RNA replication. In chapter 2, we consider a polymerase ribozyme replicating by the Qβ bacteriophage mechanism. When bound to a surface, limited diffusion allows for survival so long as the termination error rate is below an error threshold. In Chapter 3, we consider the replication of short oligomers through an abiotic mechanism proposed in prebiotic experiments. When limited by substrate availability, competition results in the emergence of uniform RNA polymers from a messy prebiotic soup containing nucleotides of different chirality and sugars. In chapter 4, we consider the possibility of an RNA world lacking cytosine. Without cytosine, the ability of RNA to fold to complex secondary structures is limited. Furthermore, G-U wobble base pairing hinders the transfer of information during replication. Nevertheless, we conclude that an RNA world lacking cytosine may be possible, but more difficult for the initial emergence of life. In chapter 5, we analyze abiotic and viral mechanisms of RNA replication using known kinetic and thermodynamic data. While most mechanisms fail under non-enzymatic conditions, rolling-circle replication appears possible. In chapter 6, we extend our analysis of the rolling-circle mechanism to consider the fidelity of replication. Due to the thermodynamic penalty of incorporating an error, rolling-circle replication appears to undergo error correction. This results in highly accurate replication and circumvents Eigen's paradox. Rolling-circle replication therefore presents an appealing option for the emergence of RNA replication in an RNA world.

Acknowledgements

I would like to thank my family for their support and the sacrifices they made to provide me this opportunity. Without you, none of this would be possible.

I would also like to thank my supervisor Dr. Paul Higgs for his guidance and for his trust in me, which has given me the tools and freedom to formulate and test my own ideas. In addition, I would like to thank my supervisory committee members for their encouragement and enthusiasm throughout my thesis. As exhausting as they were, I will always miss the debates we had during my committee meetings.

Lastly, I would like to thank my amazing partner Caroline Cauret who has always been there when I needed her, even when I thought I was strong enough to go it alone. I know that I can be grumpy and unbearable when I face difficulty. The fact that you are always there for me is a testament to your good nature.

Table of Contents

Lists of Figures

List of all Abbreviations and Symbols

A	
AMP	Adenosine monophosphate (RNA form)
ATP	Adenosine triphosphate (RNA form)
C	
CoA	Coenzyme A
D	
DNA	Deoxy-ribonucleic acid
dsDNA	Double-stranded DNA polymer
dsRNA	Double-stranded RNA polymer
F	
$FADH_2$	Flavin adenine dinucleotide
G	
GMP	Guanosine monophosphate (RNA form)
L	
LHB	Late Heavy Bombardment
LUCA	Last universal common ancestor
M	
mRNA	Messenger-RNA

N	
NADH	Nicotinamide adenine dinucleotide
NADPH	Nicotinamide adenine dinucleotide phosphate
NTP	Any nucleotide in tri-phosphorylated state (RNA form)
R	
RNA	Ribonucleic acid
RNase P	Ribonuclease P enzyme
S	
SAM-e	S-Adenosyl methionine
ssDNA	Single-stranded DNA polymer
ssRNA	Single-stranded RNA polymer
T	
tRNA	Transfer-RNA

Preface

The presented ‘sandwich’ thesis was written to conform to the guidelines set by the McMaster University School of Graduate Studies. Each chapter is a published paper in a peer-reviewed scientific journal or a manuscript in preparation. Published papers are presented as they appeared in the journal and are prefaced by a statement which entails each authors contribution and the permission attained for their reproduction in this thesis. Manuscripts in preparation are formatted according to the intended journal of submission and are prefaced by a statement which entails each authors contribution.

Chapter 1: Introduction

Great uncertainty remains regarding the origin of life on Earth. The leading theory is founded on the idea of an "RNA world", a time in which primitive organisms used RNA as a genetic material and a catalyst (Gilbert 1986; Bernhardt 2012; Robertson and Joyce 2012; Higgs and Lehman 2015). The existence of an ancient RNA world is heavily supported by the biology of modern life. However, research into the emergence of life into an RNA world has faced great difficulty, both in the synthesis of the RNA building blocks and the replication of RNA polymers.

The introduction of this thesis is laid out as follows. In section 1.1, we consider the time interval for the origin of life, which attempts to constrain the 'when' part of the origin of life question. After which, we focus on the 'how' of the origin of life. In section 1.2, we consider a top-down approach, using what we know about modern life to predict our origins. From this, the idea of an ancient RNA world emerges. In section 1.3, we consider how prebiotic chemistry could lead to the emergence RNA nucleotides, RNA polymers, and RNA replication, the fundamental building blocks of the RNA world. In section 1.4, we review the use of computational modeling as it relates to RNA replication. Lastly, in Section 1.5, we conclude the introduction and discuss how chapters 2-6 of this thesis build upon current origins of life research.

1.1 Time Interval for the Origin of Life

To determine the 'when' of the origin of life question we will consider the astrophysical constraints and biological signatures which constrain the time interval for the origin of life. The astrophysical constraints are those which determine when the emergence of life on Earth is possible. For instance, to have life on Earth, there first needs to be an Earth. Additionally, the crust of the Earth needs to be sufficiently stable such that a habitable environment exists to host the emergence of life. The biological signatures are the historical evidences of past life. For instance, when evidence of life is found within a rock, there is reasonably confidence that the emergence of life occurred sometime prior to the formation of the rock. This includes the fossilized remains of organisms, but also the chemical traces that life leaves behind even after it is gone. By combining the astrophysical constraints and the biological signatures, a time interval for the origin of life is proposed. A summary of the astrophysical constraints and biological signatures is provided in Figure 1.1, which is taken from a review paper authored by myself and others at the Origins Institute at McMaster University (Pearce et al. 2018)

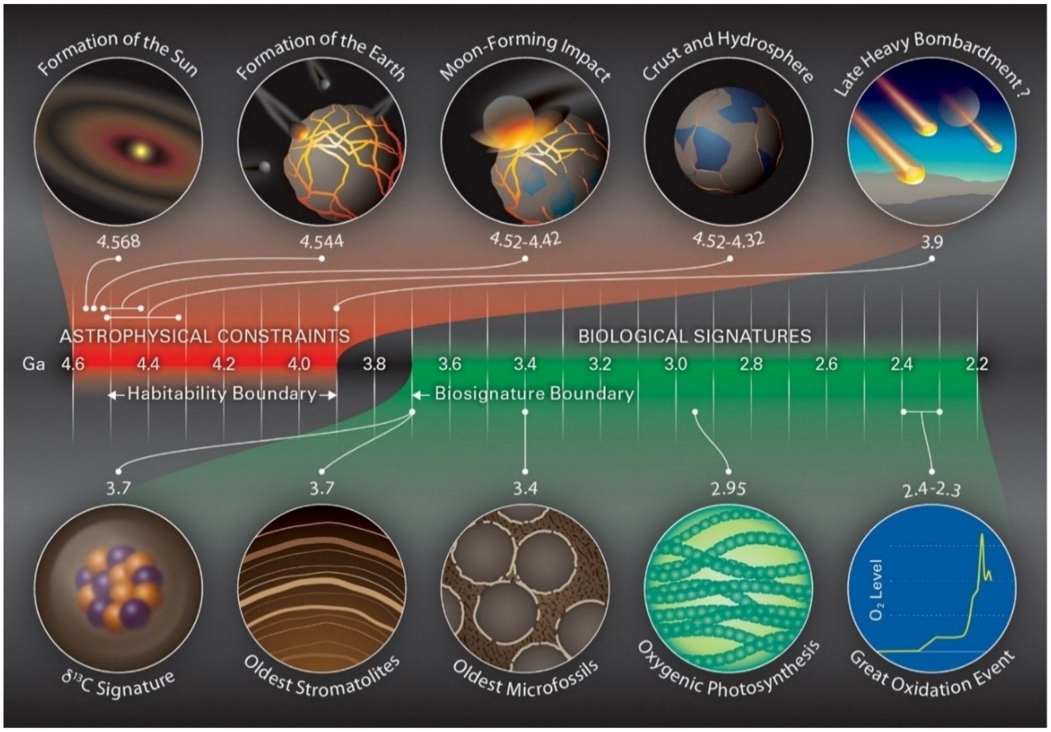


Figure 1.1.1: Time interval for the origin of life as predicted by the astrophysical constraints and biological signatures. Due to uncertainty regarding the existence and severity of the late heavy bombardment, the astrophysical constraints are uncertain, as given by the broad habitability boundary. Figure is taken from Pearce et al., (2018), in accordance with the terms defined by creative commons 4.0.

Based on the astrophysical constraints, a habitability boundary ranging from 4.52 Ga to 3.9 Ga is inferred. This is the predicted time interval in which the Earth became sufficiently habitable such that the emergence of life was possible. The earliest time of 4.52 Ga is based on the Earth becoming habitable just after the Moon forming impact (Jacobson et al. 2014; Barboni et al. 2017) and fast cooling of the crust (Lebrun et al. 2013; Monteux et al. 2016). Even if the Moon forming impact occurred later, and cooling took longer, zircon crystals suggest a stable hydrosphere and liquid surface water at 4.32 Ga (Cavosie et al. 2005; Harrison 2005). The presence of which is taken to be convincing evidence that the surface of the Earth was sufficiently stable such that the emergence of life could occur. The last astrophysical constraint is based on the cratering record of the Moon which predicts a "Late Heavy Bombardment" of meteorites, abbreviated LHB, at roughly 3.9 Ga (Zahnle et al. 2007). Depending on its severity, the LHB may have sterilized the surface of the Earth and delayed the emergence of life. However, it is not clear whether the LHB occurred (Zahnle et al. 2007; Spudis et al. 2011; Boehnke and Harrison 2016), and even if it did occur, life may have been able to survive the impacts since sterilization events would be localized (Abramov et al. 2013). Therefore, if the LHB did occur, and it was sterilizing, then life would have had to emerge after 3.9 Ga. Whereas without the LHB, life could have emerged as early as 4.52 Ga.

The biological signatures of past life can be used to form a biosignature boundary, marking the most recent time at which the emergence of life could have occurred on Earth. Starting with the least controversial evidence, microfossils dating to 3.43 Ga have been discovered in western Australia (Wacey et al. 2011). These fossils have been analyzed with electron microscopy, as well as mass spectrometry, and the remains of their cell lumina and cell walls have been confirmed. The 3.7 Ga date which marks the biosignature boundary is constrained by the discovery of layered rock structures resembling modern stromatolites (Nutman et al. 2016), as well as isotopic evidence of light carbon which is indicative of decomposed organic matter (Rosing 1999; Ohtomo et al. 2014). Potential evidence of life has also been found in zircon crystals dating to 4.1 Ga (Bell et al. 2015) and 4.25 Ga (Nemchin et al. 2008). These, however, are highly speculative and difficult to confirm given the rarity of zircon crystals and the harsh conditions under which they form. Therefore, the biosignature boundary is currently placed at 3.7 Ga, based on the joint evidence of stromatolites and carbon isotope dating.

When combined, the habitability boundary and the biosignature boundary constrain the emergence of life to a time interval ranging from 4.52 Ga to 3.7 Ga. This is a relatively broad time interval, but importantly, it shows just how ancient life is on Earth. More recently, a narrow time interval for the emergence of life at 4.36 ± 0.1 Ga has been proposed (Benner et al. 2020). This is based on the hypothesis that the Earth was hit by a second large impactor, named Moneta, after the Moon forming event. The Moneta impactor would have been roughly the size of the Moon, sterilized the Earth on impact, then left a shortly lived reducing atmosphere to kick-start the RNA world. The evidence in support for such an impactor, however, is limited. Furthermore, this narrow time interval is based on the hypothesis that a reducing atmosphere is required to produce the material necessary for the RNA world which is still under debate and will be discussed in section 1.3.1.

Due to the ancient origins of modern life, and the sparse rock record of the Hadean Earth, there is limited knowledge about the environmental conditions which led to the emergence of life. Furthermore, since the most ancient biological signatures are isotopic differences in rock and rock structures, there is limited information about the organisms which left them. Due to these limitations, origins of life research typically follow a bottom-up or top-down approach. The bottom-up approach starts with chemistry and tries to build up towards simple life. Whereas the top-down approach starts with the complexity of modern life and attempts to work backwards. Where these approaches meet is the RNA world.

1.2 Evidence for an Ancient RNA world

The time interval for the origin of life discussed previously was based on the geological record of the Earth and the greater solar system. We will now shift our focus to the biological record found within modern life. By rewinding the evolutionary clock, we can take a top-down approach to the origin of life and acquire information of our ancestral origins. We will start by considering the tree of life which shows how all life on Earth diverged from a common ancestor, LUCA, the Last Universal Common Ancestor, roughly 4 billion years ago. We will then consider the unifying traits of life, with a particular emphasis on the role of RNA in modern life. Since these traits are universal, they were likely present in LUCA and subsequently conserved throughout evolution. These unifying traits of modern life therefore provide a glimpse of what life was like ~4 billion years ago. Lastly, we will consider the important role in which RNA appears to have played in LUCA, and from this, the idea of an ancient RNA world.

1.2.1 The Tree of Life

With the advent of modern sequencing, the evolutionary relationships between organisms has been reimagined. Instead of the traditional five kingdom view (Monera, Protista, Fungi, Plantae, and Animalia), life is now separated across three domains: Archaea, Bacteria, and Eukaryota (Woese 1987; Woese et al. 1990). The rational for the latter being the result of phylogenetic reconstruction studies of the tree of life (Woese 1987; Hug et al. 2016). In the phylogenetic tree of life, the three domains are connected, with Eukaryotes being closer related to Archaea than bacteria. LUCA, the ancestor to the three domains is thought to be somewhere between the Bacteria and the Archaea-Eukaryote branches. As of now, there is still debate as to whether the three domains branch directly from LUCA, or whether Eukaryotes are a branch within Archaea (Williams et al. 2013; Raymann et al. 2015). Regardless of this distinction, the tree of life provides great utility to origins of life research. Since all modern life is related, we can look at the unifying traits of modern life and be reasonably confident that these were also present in LUCA.

This is important as it allows us to set an appropriate end goal for origins of life research.

1.2.2 Universal Traits of Life

Modern life as we know it depends on the central dogma of molecular biology (Figure 1.2.1). This simply states that the genetic information of life is stored in nucleic acids, and that information can be transferred to proteins, but the reverse information transfer never occurs (Crick 1970). In modern life, DNA stores the genetic information of the cell, and can be transcribed into RNA. The transcribed RNA can then be translated into proteins, which carry out the necessary enzymatic catalysis within the cell. As such, the information stored in DNA is converted to information stored in proteins, yet the information stored in proteins is never converted back into DNA. Since this is a universal trait to life, LUCA is expected to have also followed the central dogma of molecular biology. In the canonical interpretation of the central dogma, RNA plays an intermediary role which appears secondary to the roles of DNA and protein. Upon closer inspection we find that RNA is much more important than it first appears, being vital in the translation of proteins, the replication of DNA, and the metabolism of a cell.

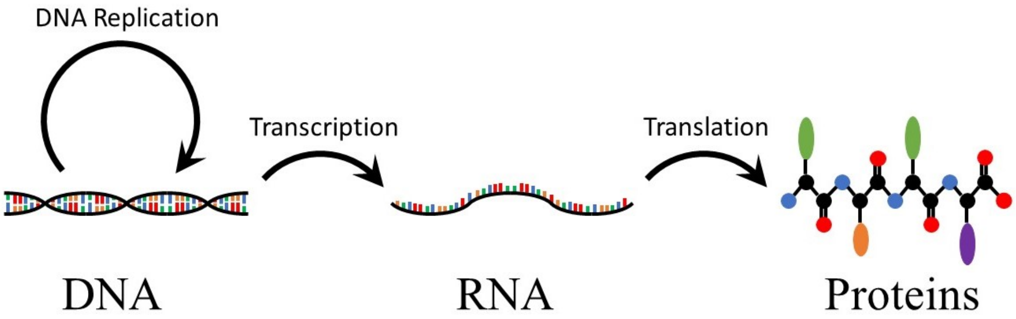


Figure 1.2.1: Central dogma of molecular biology. DNA can be replicated to form more DNA, with the help of DNA polymerase proteins. DNA can be transcribed into RNA by RNA polymerase proteins. And RNA can be translated into proteins by the ribosome. Information flows from nucleic acids to proteins, and never from proteins to nucleic acids.

To translate RNA into protein, modern life requires mRNA, tRNA, and ribosomes, all of which are composed primarily of RNA. The ribosome is particularly impressive due to its size and ability to catalyze protein synthesis, despite not being a protein enzyme. Instead, the ribosome is a universally conserved RNA enzyme, or ribozyme for short, utilizing RNA in its active site to catalyze protein synthesis (Ban et al. 2000; Cech 2000; Nissen et al. 2000; Moore and Steitz 2002). In addition to the ribosome, RNase P is a second universally conserved ribozyme. RNase P is also critical for translation as it catalyzes cleavage of the RNA backbone in the maturation of tRNA (Lai et al. 2010). Without RNA acting in this enzymatic capacity, the synthesis of proteins in modern life would not be possible.

The importance of RNA is also observed during the process of DNA replication. To initiate the replication of DNA, protein polymerases require a primer. However, life has yet to discover a way to synthesize a DNA primer. Instead, all DNA replication starts by first synthesizing an RNA primer (Frick and Richardson 2001; O'Donnell et al. 2013). This appears to be rather counterproductive. Firstly, the RNA primer becomes incorporated into an otherwise dsDNA helix and requires removal by additional protein enzymes. Secondly, the reliance on RNA primers prevents replication at the ends of linear DNA chromosomes since protein polymerases are unidirectional. As such, with every round of genome replication, linear chromosomes shrink in length (Watson 1972). To combat the ever-decreasing size of genomes, even more protein enzymes are required to extend the linear genomes with non-coding information (Greider and Blackburn 1985; Blackburn et al. 2006). These additional complexities could be avoided with the use of a DNA primer, yet life appears to be stuck with RNA primers. The utilization of RNA primers shows not only the ancestral importance of RNA in DNA replication, but also the limited capacity of evolution.

Delving deeper into the process of DNA replication, the *de novo* synthesis of DNA nucleotides first requires the synthesis of RNA nucleotides. This is perhaps not surprising given the similarity between the two molecules. In terms of chemical structure, there are only minor differences between an RNA nucleotide and a DNA nucleotide. This first of which is the replacement of the 2' hydroxyl group of ribose with a hydrogen atom, resulting in deoxyribose. In modern life, this requires a rather complex reduction reaction which is catalyzed by a ribonucleotide reductase protein (Poole et al. 2002; Lundin et al. 2015). For adenosine, guanosine, and cytidine, this is the only difference between an RNA nucleotide and a DNA nucleotide. To synthesize thymidine, a uridine nucleotide is first converted to deoxy-uridine, which is then methylated to form the thymidine nucleotide. Therefore, before modern life can even begin the process of DNA replication, it must first convert RNA nucleotides into DNA nucleotides, with the help of proteins.

The importance of RNA is also observed when considering the metabolism occurring within a cell. Nearly all chemical reactions within a cell are catalyzed by protein enzymes. However, within these protein enzymes there an abundance of RNA derived coenzymes, including NADH, NADPH, $FADH_2$, CoA, SAM-e, and ATP. These are molecules which are derived from RNA nucleotides and are required for the activity of the enzyme. The rational for the abundance of RNA derived coenzymes is that they are the ancient remains of ribozymes. Through evolution, these ribozymes were replaced by their modern protein enzyme counterparts through a stepwise takeover process which passed through an RNA-protein enzyme intermediate (White 1976; Benner et al. 1989; Jeffares et al. 1998; Raffaelli 2011). Since the active site is vital for enzymatic activity, it is most difficult to replace and thus remained, whereas the surrounding RNA was slowly replaced by protein. This evolution of ribozymes to protein enzymes is consistent with the modern ribosome and RNase P which have both incorporated proteins.

In addition to the role of RNA in active sites, RNA nucleotides are also abundant in cell signalling pathways, in particular, the cyclic forms of AMP (Nelson and Breaker 2017). It is important to note that most, but not all, of the previously mentioned coenzymes and signalling molecules do not depend on the chemistry of the 2’ hydroxyl group of the RNA nucleotide. In theory, these coenzymes and signalling molecules could be derived from DNA nucleotides. The fact that they are not may reveal something important about their evolutionary history.

1.2.3 The Ancient RNA world

Based on the universal traits of life, LUCA is expected to have been a relatively complex organism or “progenote” (Woese 1987; Doolittle 2000; Di Giulio 2011). Minimum genome studies suggest that LUCA likely contained at least 500-1500 genes, comparable in size to some modern organisms (Koonin 2003; Ouzounis et al. 2006). LUCA also followed the central dogma of biology, using DNA or perhaps RNA (Poole and Logan 2005) as its genetic material, proteins to catalyze its chemical reactions, and RNA as the intermediate for translation. Additionally, LUCA is thought to have used RNA extensively in the production of coenzymes and as signalling molecules for metabolism. Looking past LUCA, towards the origin of life, requires dissecting the complexity of the central dogma.

Since the central dogma of molecular biology relies on the cooperation of DNA, RNA, and proteins, it is too complex to emerge spontaneously. Instead, the central dogma is likely a product of evolution. If this is true, then one of the three key polymers likely emerged first, with the others arriving later. Based on the importance of RNA discussed previously, it is currently thought that RNA preceded the emergence of DNA and protein. Intuitively, this seems to be a rational idea. To synthesize proteins, RNA is required to act as mRNA, tRNA, ribosomes, and even RNase P. Similarly, RNA is required to replicate DNA, both in the synthesis of primers and in the *de novo* synthesis of DNA nucleotides. In which case, the possible orderings of emergence are RNA → proteins → DNA and RNA → DNA → proteins.

The order of emergence of DNA and proteins is still uncertain. If RNA is a poor catalyst, then proteins likely preceded DNA since proteins are required by modern life to synthesize DNA nucleotides and replicate DNA. However, it is still within the realm of possibility that DNA emerged prior to proteins. In which case, ribozymes, with the help of their coenzymes, catalyzed the early replication and synthesis of DNA nucleotides, which were then replaced by proteins. Regardless of the ordering, the emergence of RNA first appears to explain the critical role that RNA plays in metabolism and DNA replication. For instance, the reason many coenzymes and signaling molecules are derivatives of RNA, and not DNA, may simply be due to the absence of DNA at that time in evolution. Similarly, the *de novo* synthesis of DNA nucleotides was likely an addition to the already present synthesis of RNA nucleotides. And the use of RNA primers in DNA replication may be a remnant of a time in which RNA acted as the genetic polymer.

If RNA did precede the emergence of DNA and proteins, then life at one time was dependent on RNA to act as both a genetic polymer and a catalyst. Such a time is commonly referred to as the RNA world (Bernhardt 2012; Robertson and Joyce 2012; Higgs and Lehman 2015), which will be discussed more extensively in the next section. However, let us first consider the likelihood of RNA acting as a genetic polymer and a catalyst based on modern biology. While RNA is not used as a genetic material in any known life, it is used extensively in the viral world (Cameron et al. 2009; Hulo et al. 2011). RNA viruses utilize RNA as a genetic material and store the information required for protein synthesis. The viral world also contains viroids and satellite RNAs. These viral agents are infectious like viruses, except they are smaller, unprotected, and do not encode for proteins. Interestingly, viroids, satellite RNAs, and some viruses encode for self-cleaving ribozymes which they require for genome replication (Flores et al. 2011). The utilization of RNA genomes and ribozymes thus make the viral world attractive from an origins of life perspective (Diener 1989; Flores et al. 2014; Diener 2016; Berliner et al. 2018; Maurel et al. 2019). Based on the extensive use of RNA genomes in the viral world, and the importance of RNA to modern life, it seems reasonable to assume that life at one time could have used RNA as a genetic material.

As a catalyst, some RNA polymers are known to act as ribozymes to catalyze chemical reactions. In addition to the ribosome and RNase P mentioned previously, modern life is also known to utilize self-splicing ribozymes to parse mRNA (Cech 1987; Valadkhan 2007) and self-cleaving ribozymes are found extensively in genomes from all three domains of life (Perreault et al. 2011; Hammann et al. 2012). All of the naturally occurring ribozymes catalyze the cleavage or ligation of the RNA backbone, the sole exception being the ribosome (Doudna and Cech 2002; Talini et al. 2009). As such, the chemical repertoire of RNA catalysis appears limited. However, with the help of coenzymes and cofactors, RNA may have been able to perform the catalysis required for a primitive metabolism. The next section will further elaborate on the abundance of ribozymes discovered through *in vitro* evolution experiments.

1.3 From chemistry to an RNA world

In this section, we will discuss the chemistry which could have given rise to an RNA world. We will start with the origin of metabolism and the synthesis of RNA nucleotides both from a top-down and a bottom-up perspective. From there, we will consider how these RNA nucleotides can be ligated to form RNA polymers. And lastly, we will consider how these polymers could replicate non-enzymatically, or with the help of ribozymes. When combined with the previous section, the top-down and bottom-up approaches provide us a comprehensive view of the RNA world.

1.3.1 The Origin of Metabolism

To reach an RNA world, RNA nucleotides are essential, however, the mechanism which led to their synthesis remains unknown. There are currently two domains of thought on this topic resulting from top-down and bottom-up approaches. From the top-down perspective, modern metabolism is argued to be a blueprint for prebiotic RNA nucleotide synthesis (Ralser 2018). In this scenario, the origin of RNA nucleotides for the RNA world mimics that in modern life, except that it is initially occurring under non-enzymatic conditions. The bottom-up perspective argues that modern metabolism is too complex to emerge spontaneously and is instead a product of evolution (Orgel 2004; Benner et al. 2019). In which case, the original metabolism which gave rise to the RNA world is different than that used by modern biology. Here, we will attempt to summarize some of the main finding on each side.

In modern life, the proteins which catalyze metabolic reactions differ across the three domains of life, however, the general method of RNA nucleotide synthesis is universal, likely existing in LUCA (Weiss et al. 2016; Ralser 2018). This is intriguing from an origins of life perspective as it implies that modifying one's own metabolism through evolution is exceedingly difficult. When this reasoning is extended to times prior to LUCA, one expects that the ancestor to LUCA also contained the same metabolism. Rewinding the clock even further would imply that modern metabolism existing in some primitive form, under non-enzymatic conditions. Recent research suggests that such a primitive metabolism may indeed be possible non-enzymatically. In the presence of Fe(II), which is expected to be abundant prebiotically (Rouxel et al. 2005; Busigny et al. 2014), a non-enzymatic glycolysis and pentose phosphate system emerges (Keller et al. 2014; Keller et al. 2016). Similarly, portions of the Krebs cycle are catalyzed by iron and sulfur species (Keller et al. 2017), as well as UV light (Zhang and Martin 2006), and other reactive metal species (Muchowska et al. 2017). Even gluconeogenesis (Messner et al. 2017), and the synthesis of SAM-e (Laurino and Tawfik 2017) appear to be possible non-enzymatic. These non-enzymatic pathways are not highly refined as they are in modern life, but the interconversions between chemical species are there. In principle, this could allow for the non-enzymatic synthesis of RNA building blocks from simpler species. Once the RNA world is reached, ribozymes could refine the network and enhance the rate of RNA nucleotide synthesis. In this scenario, ribozymes are not inventing new chemistry. Instead, they are merely binding onto inorganic catalysts and enhancing the rate and or specificity of the chemical reactions (Ralser 2014).

In contrast to the top-down approach, the bottom-up approach attempts to synthesize RNA nucleotides from simpler chemical precursors, which are typically different than those used by modern life. The inspiration for such approaches dates back to Butlerow's synthesis of sugars from the formose reaction (Butlerow 1861), Stanley Miller's synthesis of amino acids from reducing gases (Miller 1953), and Oro's synthesis of purines from ammonium cyanide solutions (Oró 1960). For a historical review of prebiotic synthesis experiments, see Orgel, 2004. Here, we will focus on the recent advances in the synthesis of RNA nucleotides. In addition to synthesizing amino acids, the Miller-Urey reducing atmosphere has been shown to also result in the synthesis of nucleobases (Ferus et al. 2017). Borate minerals have also been shown to preferentially stabilize ribose sugars formed from the formose reaction (Ricardo et al. 2004; Grew et al. 2011). In combination, these reactions could have been producing the required nucleobases and sugars for an RNA world. Additionally, carbonaceous meteorites have been shown to contain both ribose sugars (Furukawa et al. 2019) and nucleobases (Callahan et al. 2011; Pearce and Pudritz 2015), leading to a second possible source of RNA nucleotide precursors. If these nucleobases and sugars can be combined, then RNA nucleotide synthesis may be possible. Under aqueous conditions, this addition is difficult (Fuller et al. 1972), however, wet-dry cycling make RNA nucleotide synthesis possible (Becker et al. 2019). Alternatively, the synthesis of RNA nucleotides can avoid free ribose and nucleobase entirely. Staring with cyanamide, glycolaldehyde, and other simple reactants, the synthesis of phosphorylated RNA nucleotides has been shown (Powner et al. 2009). The appeal of the bottom-up synthesis reactions is their simplicity in starting reagents, such that one could imagine synthesis occurring without the help of an experimenter of complex biology. And from this simplicity, life emerged, which through the complexity of evolution, led to the complexity of modern metabolism.

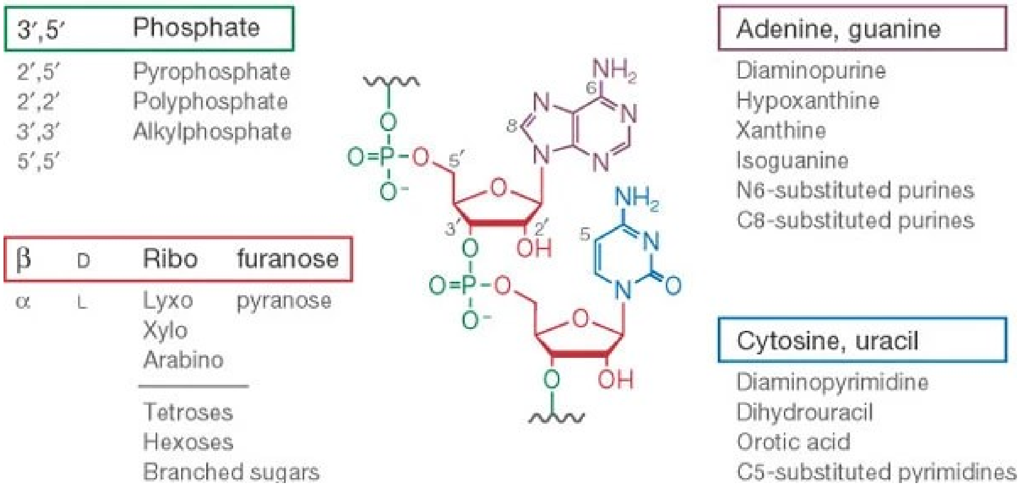


Figure 1.3.1: The prebiotic clutter of RNA nucleotide synthesis. Since the synthesis of RNA nucleotides results in many alternative sugars and nucleobases, the number of possible nucleotides is exceptionally large. Figure is taken without alteration from Joyce, (2002), with permission from Springer Nature.

Regardless of the mechanism of RNA nucleotide synthesis, there are some inherent problems which have yet to be overcome. The first is the messy nature of chemical reactions which results in a prebiotic clutter, see Figure 1.3.1 (Joyce 2002). Without protein enzymes (or ribozymes), chemical reactions produce an abundance of side products, whereas the desired product is typically rare. Forming RNA nucleotides is thus difficult given the many alternative combinations of sugars and nucleobases which are also formed in the same reaction mixture (Joyce 2002; Cleaves and Bada 2012; Krishnamurthy 2015). A further problem of RNA nucleotide synthesis is that ribose is a chiral molecule, existing in both right-handed and left-handed forms. While life only utilizes right-handed ribose, prebiotic chemistry is expected to produce both left and right-handed forms at equal abundance. Since chiral isomers have identical chemical properties, separating them is non-trivial. For an RNA world to exist, these problems of RNA nucleotide synthesis need to be solved.

1.3.2 The Formation of RNA Polymers

To reach an RNA world from chemistry, RNA nucleotides need to polymerize into RNA polymers. The types of RNA polymerization can be broken down into three classes (Higgs 2019). In the first class, RNA polymers are generated through the spontaneous ligation of nucleotides or oligomers. The polymers resulting from this type of synthesis have random sequences. In the second class, polymerization occurs while RNA nucleotides or oligomers are aligned on a template. In which case, the new polymer has a sequence which is complementary, or near complementary, to the template. In the third class, polymers are generated by the same mechanism as the second, except in this case, the reaction is catalyzed by a ribozyme. The second and third classes of RNA polymerization are important for RNA replication as they enable the transfer of information from a template to a new polymer. Here, we will focus on the first class of polymerization in which polymers form from spontaneous ligation of RNA nucleotides or oligomers.

Spontaneous polymerization of nucleotides has been considered under a variety of conditions, including aqueous solution, on (or within) clay minerals, within multilamellar lipids, and with wet-dry cycling. We will start our discussion with the expected equilibrium distribution of polymerization based on the thermodynamics of the process. If the process of polymerization is reversible, the resulting distribution of polymer lengths is expected to follow the geometric Flory-Shultz distribution (Higgs 2016; Spaeth and Hargrave 2020). This simply means that the ratio of concentrations between a polymer or length n, and a polymer of length $n + 1$, is a constant: $\frac{C_{n+1}}{C_n} = \frac{KC_{tot}}{1+KC_{tot}}$, where K is the equilibrium constant and

C_{tot} is the total concentration of nucleotides in solution. Since this ratio is less than unity, long polymers are expected to be exceedingly rare, especially when the equilibrium constant K is small. Intuitively, this makes sense. When bond formation is not favorable, $K < 1$, polymers are expected to be rare, as the theory predicts. To form long polymers in appreciable concentrations, bond formation needs to be favorable or the concentration of nucleotides needs to high.

Under aqueous conditions, the polymerization of RNA from nucleotide monophosphates, either 2'-phosphate, 3'-phosphate, or 5'-phosphate, is not detected in appreciable amounts (Giovanna et al. 2009). This is to be expected since phosphodiester bond formation is predicted to thermodynamically unfavorable (Dickson et al. 2000), which makes polymerization unfavorable, and nearly all nucleotides remain in the monomeric state. To achieve RNA polymerization in water, activated RNA nucleotides are required. Starting with nucleoside 2',3'-cyclic phosphate, the detecting of short oligomers is possible (Giovanna et al. 2009). In this case, the opening of the cyclic phosphate is thermodynamically favorable (Rudolph et al. 1971), which coupled to bond formation, allows polymerization to occur. Nucleoside 3',5'-cyclic phosphate also polymerize in aqueous solution, and to a greater extent, forming chains of at least 25 nucleotides (Giovanna et al. 2009; Costanzo et al. 2012; Costanzo et al. 2016). The longer lengths resulting from the increased energy liberated from opening the cycle (Rudolph et al. 1971).

As an alternative to aqueous conditions, some researchers have proposed polymerization with clay, lipid, or other minerals which contain localized environments that make bond formation favorable. In these localized environments, the nucleotides are free to polymerize. When the polymers are then released into the bulk solvent, they are out of thermodynamic equilibrium. Nevertheless, they persist in the aqueous phase since they are kinetically trapped. In short, the rate of polymer generation by the local environment is faster than the destruction in the aqueous environment, such that the normal aqueous equilibrium cannot be reached.

For clays, polymers of 30-50 nucleotides in length can be detected (Ferris 2002; Huang and Ferris 2006; Jheeta and Joshi 2014). This, however, still relies on activated nucleotides, and the specificity of the clay is worrying (Aldersley et al. 2017). In lipid environments, wet-dry cycling appears to generate polymers of 25-100 nucleotides from unactivated nucleoside 5'-monophosphates (Rajamani et al. 2008). In this case, the lipids act as an ordering agent in the dry phase to drive the polymerization of RNA. Under similar wet-dry conditions with salts, polymers up to 300 nucleotides are observed (Da Silva et al. 2015). In this case, the salt crystalizes upon drying allowing it to act as an ordering agent in the same way as the lipids.

It is important to state here that any of the previously discussed conditions could generate RNA polymers sufficiently long for an RNA world. The difference is the amount of time required. For example, polymerization under aqueous conditions resulted in the detection of 25-mers. Longer oligomers are expected to be present, but these were likely not detected because their concentrations were below the detection limit of the experimental setup. Therefore, if emerging life requires a specific 200-mer polymer, it may simply take longer in an aqueous environment, as opposed to a clay of lipid environment. Given that the time scale for the emergence of life is on the order of hundreds of millions of years, not the days to weeks typical of experimental research, this difference may not be crucial. In addition to time, it is also important to consider space. While wet-dry cycling environments outperform all others in terms of polymerization, these environments are also expected to be relatively rare compared to aqueous environments. What aqueous environments lack in efficiency, they may make up for in abundance.

1.3.3 The Emergence of RNA Replication

For the next step in the emergence of life, the RNA polymers created from spontaneous polymerization, or a subset of them, must be able to replicate. The replication of RNA polymers is key to the emergence of life as it allows for the transfer of information from a parent polymer to an offspring polymer. Which when combined with selection, allows for the evolution of RNA polymers and an increase in complexity. Experimentally, the replication of RNA polymers has proven difficult under both non-enzymatic and ribozyme catalyzed conditions and remains an active area of research.

Under non-enzymatic conditions, an RNA polymer (or for convenience a DNA polymer) is used as a template to synthesize a new polymer from monomers or oligomers. While template directed synthesis of the RNA polymers has been shown many times (Orgel 1995; Kozlov and Orgel 2000; Prywes et al. 2016; He et al. 2017), it has proven difficult to separate the resulting dsRNA duplex as required for successive round of replication. This is commonly referred to as the product inhibition problem of RNA replication, which is just one of eight unresolved problems facing non-enzymatic RNA replication (Szostak 2012). Under non-enzymatic conditions, multiple rounds of replication have only been achieved a few times and has been restricted to very short templates containing 4 bases (Zielinski and Orgel 1987), 6 bases (von Kiedrowski 1986; Achilles and von Kiedrowski 1993; Sievers and Von Kiedrowski 1994), and 24 bases (Edeleva et al. 2019). The longest of which requires the utilization of temperature cycling to drive strand separation.

Alternatively, it is possible that RNA replication is only achievable enzymatically through catalysis by ribozymes. In which case, most spontaneously generated RNA polymers would be “dead”, and replication would only start with the discovery of a specific ribozyme. Through *in vitro* evolution, ligases have been discovered which catalyze replication of themselves when given the two sequence specific fragments (Paul and Joyce 2002; Kim and Joyce 2004; Lincoln and Joyce 2009). A self-assembling ribozyme has also been engineered, which utilizes recombination reactions to make itself from a set of sequence specific oligomers (Draper et al. 2008; Hayden et al. 2008; Jayathilaka and Lehman 2018). Since the substrates in both cases are long sequence-specific oligomers, they are expected to be rarely formed from spontaneous polymerization. Unless these substrates can also replicate, the self-replication or self-assembly of these ribozymes under prebiotic conditions would likely not be possible. Furthermore, it is unclear how these ribozymes would allow for the maintenance of additional ribozymes as required for emerging life to increase in complexity and expand its RNA genome.

Great effort has also been put forth in the discovery of polymerase ribozymes which mimic the catalysis of modern protein polymerases. Current polymerase ribozymes trace their origins back to the *in vitro* evolution experiment by Bartel & Szostak, (1993) which found sequences with ligation activity from a starting pool of random sequences. Subsequent experiments improved on these sequences to generate the Class 1 ligase (Ekland et al. 1995; Ekland and Bartel 1995), which had a catalytic turnover rate of 100/min, comparable to modern protein polymerases. The Class 1 ligase, however, is limited to primer extension of sequences covalently attached to itself. Upon further modification, the round-18 polymerase was created which could bind onto generic RNA templates (Johnston et al. 2001). At its time, the round-18 polymerase was impressive, capable of extending a primer 14 times within 24 hours. However, this activity falls short of protein polymerases. Particularly problematic is its weak affinity to the template-primer duplex, and poor processivity, which makes synthesis slow (Lawrence and Bartel 2003). Nevertheless, the round-18 polymerase has served as a valuable starting point for the engineering of superior polymerase ribozymes (Lawrence and Bartel 2005; Zaher and Unrau 2007; Wochner et al. 2011; Attwater et al. 2013; Horning and Joyce 2016; Attwater et al. 2018). While none have reached the level of self-replication, as product inhibition remains a key issue (Cheng and Unrau 2010), the improvements made thus far provide hope that self-replicating polymerase ribozymes may be discovered in the near future.

1.4 Computational Modeling of RNA replication

Due to the experimental problems facing RNA replication, computational modeling may be of great utility as it can provide insight into what is to come and guide experimental research down a productive path. In RNA replication research, computational models tend to fall into one of two classes, which will be referred to as abstract models and realistic models. Abstract models take a high-level theoretical approach and consider just a few parameters which capture the dynamics of the simulation. These models provide a theoretical framework and can guide the long-term direction of experimental research. In contrast, realistic models attempt to build directly on experimental data to provide short-term direction to experimental research. While both classes of models have utility in RNA replication research, current literature is dominated by abstract models.

Many abstract computational models have considered self-replicating polymerase ribozymes, or more generally “replicases”, within a larger ribo-organism (reviewed by Szilágyi et al., 2017). The utility of these abstract models is that they can determine the fidelity requirements of replication, as well as the the genome organization of a ribo-organism, even though self-replicating ribozymes have yet to discovered. Once self-replication of RNA is achieved experimentally, these models may provide a guide for the development of a more complex ribo-organism. Abstract models have also been used to consider the evolution of complexity, for instance, how a second ribozyme could emerge alongside a polymerase ribozyme (Kim and Higgs 2016). Abstract models can also consider the advantages and disadvantages of replication environments, such as replicators bound to the surface of a rock, or replicators trapped inside a protocell (Takeuchi and Hogeweg 2009; Shah et al. 2019). In doing so, these abstract models attempt to determine the environment constraints for the emergence of life. For instance, if protocells are determined to be a requirement for replicator systems, then further research into the emergence of life should focus on environments which provide lipids in addition to nucleotides. In this thesis, chapter 2, and to a lesser extent, chapter 3, rely on abstract computational model.

In contrast to abstract computational modeling, realistic computational models attempt to build directly on the experimental research. In these models, chemical kinetics and thermodynamics often determine the rules of the simulations, whereas experimental data provides reference values for the parameters. The strength of realistic models is the interpretability of the model results, such that experimentalists can use the results to guide their short-term research. Besides those covered in chapters 4, 5, and 6 this thesis, the use of realistic computational models for RNA replication research is currently lacking in the literature. However, this trend appears to be changing as the amount of experimental data increases.

1.5 Aims of this Thesis

The aim of this thesis to use computational modeling to understand the emergence of RNA replication in an RNA world. For ease of discussion, the following chapters are listed in order of completion and not as a chronological story of the origin of life in an RNA world. One will notice that the early chapters utilize abstract computational models, whereas the later chapters utilize realistic computational models. This transition from abstract towards realistic modeling is one which should be encouraged in future RNA world research. With the increasing amount of experimental data, in conjunction with greater computational power, realistic models of RNA replication are expected to become a powerful tool in RNA world research.

In chapter 2, we consider the replication of hypothetical RNA polymerase ribozymes which are confined to a surface environment which limits diffusion. This chapter builds on previous work from our group (Shay et al. 2015; Kim and Higgs 2016) and incorporates the poor processivity observed in known polymerase ribozymes. We find that survival of polymerase ribozymes is still possible on a surface so long as the termination error rate, which is related to its processivity, is below a termination error threshold.

In chapter 3, we switch from the replication of polymerase ribozymes to the replication of short RNA oligomers. In this chapter, we show that template-directed replication can result in the emergence of uniform RNA polymers from a messy prebiotic soup. This emergence of uniform RNA is the result of the faster rate at which uniform polymers are predicted to replicate (Joyce et al. 1984; Bolli et al. 1997; Gavette et al. 2016; Kim et al. 2020).

In chapter 4, we consider the possibility of an RNA world lacking cytosine, as predicted by the absence of cytosine in carbonaceous meteorites (Pearce and Pudritz 2015) and the fast deamination of cytosine to uracil (Levy and Miller 1998). Using RNA folding software (Lorenz et al. 2011), we consider the likelihood of random polymers lacking cytosine being ribozymes. We also consider the replication of these polymers using a mutational model which incorporates the importance of G-U wobble pairing. While we cannot rule out the possibility of an RNA world lacking cytosine, it seems unlikely given the statistics of folded structures and the difficulty of maintaining ribozyme information.

In chapter 5, we turned our focus to the mechanism of non-enzymatic RNA replication. We start by considering the mechanisms of RNA replication used in the viral world and in prebiotic experiments. Due to product inhibition problem, long RNA polymers are only able to achieve sustained exponential growth via the rolling-circle replication found in modern viroids and viruses.

In chapter 6, we extended our analysis of non-enzymatic rolling-circle replication to consider the fidelity of RNA synthesis. Based on the kinetics of toehold-mediated displacement, rolling-circle synthesis is predicted to undergo a thermodynamically driven error-correction mechanism. The product sequences generated from this error-correction have extremely high fidelity and Eigen's paradox is avoided. When combined with the results from chapter 5, a case can be made that the emergence of life likely started with an RNA polymer capable of undergoing non-enzymatic rolling-circle replication.

Chapter 2: Error Thresholds for RNA replication

The contents of this chapter were published in *Journal of Theoretical Biology* in September 2017 (corresponding reference below). The manuscript, figures, and tables in this chapter are used with permission from the publisher Elsevier.

Paul Higgs and I contributed to the design of the model and the writing of the manuscript. I wrote the computer programs to run the simulations, accumulated the data, and made the corresponding figures. Paul Higgs wrote the paired-site approximation and corresponding program discussed in the appendix.

Tupper, A. S., & Higgs, P. G. (2017). Error thresholds for RNA replication in the presence of both point mutations and premature termination errors. *Journal of theoretical biology*, *428*, 34-42.

Journal of Theoretical Biology 428 (2017) 34–42

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/jtbi

Error thresholds for RNA replication in the presence of both point mutations and premature termination errors

Andrew S Tupper, Paul G Higgs*

Department of Biochemistry and Department of Physics and Astronomy, McMaster University, Hamilton, Ontario, Canada

ARTICLE INFO

Article history:
Received 15 February 2017
Revised 18 May 2017
Accepted 31 May 2017
Available online 9 June 2017

Keywords:
RNA world
Polymerase
Ribozyme
Error threshold
Termination error
Point mutation
Parasite

ABSTRACT

We consider a spatial model of replication in the RNA World in which polymerase ribozymes use neighbouring strands as templates. Point mutation errors create parasites that have the same replication rate as the polymerase. We have shown previously that spatial clustering allows survival of the polymerases as long as the error rate is below a critical error threshold. Here, we additionally consider errors where a polymerase prematurely terminates replication before reaching the end of the template, creating shorter parasites that are replicated faster than the functional polymerase. In well-known experiments where Qβ RNA is replicated by an RNA polymerase protein, the virus RNA is rapidly replaced by very short nonfunctional sequences. If the same thing were to occur when the polymerase is a ribozyme, this would mean that termination errors could potentially destroy the RNA World. In this paper, we show that this is not the case in the RNA replication model studied here. When there is continued generation of parasites of all lengths by termination errors, the system can survive up to a finite error threshold, due to the formation of travelling wave patterns; hence termination errors are important, but they do not lead to the inevitable destruction of the RNA World by short parasites. The simplest assumption is that parasite replication rate is inversely proportional to the strand length. In this worst-case scenario, the error threshold for termination errors is much lower than for point mutations. We also consider a more realistic model in which the time for replication of a strand is the sum of a time for binding of the polymerase, and a time for polymerization. When the binding step is considered, termination errors are less serious than in the worst case. In the limit where the binding time is dominant, replication rates are equal for all lengths, and the error threshold for termination is the same as for point mutations.

1. Introduction

The transition from chemistry to life on Earth may have occurred in an RNA World (Bartel and Unrau, 1999; Gilbert, 1986; Joyce, 2002; Higgs and Lehman 2015), with RNA taking the central role as both genetic storage and enzymatic function in the first organisms. Central to this idea is the existence of a polymerase ribozyme capable of synthesizing a complementary sequence from a template, thereby forming a self-replicating chemical system. Support for such a ribozyme has come from in-vitro evolution experiments which have made significant progress in recent years (Attwater et al., 2013; Johnston et al., 2001; Lawrence and Bartel, 2005; Wochner et al., 2011; Zaher and Unrau, 2007). In the most recent case, a polymerase ribozyme was created that is capable of synthesizing 206 nt extensions which are approximately the same length as itself (Attwater et al., 2013). This polymerase ribozyme however is not perfect and has a fidelity of 97.4% (accuracy of base additions) and processivity of 97.5% (probability of sequential nucleotide addition prior to dissociation with the template strand).

During polymerase-mediated replication, a point mutation error is the incorporation of an incorrect nucleotide into the growing product strand, whereas a termination error is the premature termination of replication before reaching the end of the template, which creates an incomplete sequence that is shorter than the template. Both kinds of errors create non-functional template strands that have the potential to overrun the replicating system if the error rates are too high. Following Eigen et al. (1988), we will use the term 'error threshold' to describe the maximum error rate for which the replicating system can survive without being overrun by mutations. The fidelity of replication in the point-mutation case is a well studied question, but the termination problem has received much less attention. Here we consider these two kinds of error in the same model. Spatial lattice models have been used extensively to study replicating RNA systems

* Corresponding author.
E-mail address: higgsp@mcmaster.ca (P.G. Higgs).

http://dx.doi.org/10.1016/j.jtbi.2017.05.037
0022-5193/© 2017 Elsevier Ltd. All rights reserved.

(Szabó et al., 2002; Könnyű et al., 2008, 2013; Ma and Hu, 2012; Ma et al., 2010a,b, Ma et al., 2007a,b; Takeuchi and Hogeweg, 2012; Walker et al., 2012; Wu and Higgs, 2012; Shay et al., 2015; Kim and Higgs, 2016; Colizzi and Hogeweg, 2016a,b). Here we consider a spatial model designed to more accurately represent the process of strand replication by allowing for termination errors as well as point mutations.

The standard error threshold theory (Eigen et al., 1988) deals with point mutations. This theory considers a well-adapted 'master sequence', *e.g.* a wild-type RNA virus, in competition with all the mutant sequences that surround it in sequence space. The mutant sequences have a lower replication rate than the master sequence, but they are generated by continual mutations from the master sequence. The concentration of master sequences in the mixture is found to go to zero at a critical value of the point mutation rate called the error threshold. The value of the error threshold depends on the ratio of replication rates of the master sequence and the mutant sequences, and this can be calculated fairly easily (Eigen et al., 1988).

We have previously studied point mutations for an RNA polymerase in the RNA World (Kim and Higgs, 2016; Shay et al., 2015). In this case, the polymerase is an RNA and the polymerase sequence is mutating, whereas in the standard theory, the polymerase is a protein that is not subject to mutation. The simplest assumption for the RNA World model is that sequences with point mutations are non-functional as catalysts but are equally good templates as the polymerase. This means that the mutant sequences are parasites of the polymerase. In the well-mixed version of this model, the polymerase is overrun by parasites for any non-zero error rate. Survival of the polymerase at finite error rate requires cooperating groups of polymerase sequences, either in a surface-based model with slow diffusion, or in a protocell model with group selection (see Higgs and Lehman, 2015, and references therein). We have shown by simulation (Kim and Higgs, 2016) that there is an error threshold in the two-dimensional surface-based problem, and that spatial clustering allows survival of the polymerase for error rates below this threshold. Calculation of the error threshold for the spatial lattice model is not easy because it depends on spatial correlations of the states of neighbouring sites, which cannot be determined exactly. In the appendix of this paper, we give a paired-site approximation to the lattice model that explains the error threshold behaviour of this model at least qualitatively.

The main aim of this paper is to compare termination errors with point mutations. The strands generated by premature termination are shorter than the template and therefore replicate faster. The simplest assumption is that the replication rate of a strand is inversely proportional to its length. Hence there will be selective pressure for parasites of shorter lengths. This selective pressure was shown in well-known experiments with Qβ RNA, which shrunk its genome by 83% after many rounds of selection, causing it to replicate 15 times faster (Mills et al., 1967). The Qβ example uses a protein polymerase that does not evolve in the experiment, whereas a polymerase ribozyme in the RNA World would have to compete with the short parasites generated by premature termination. The worry that motivates this paper is that parasites might evolve to shorter and shorter lengths until they inevitably destroy the polymerase. If this were true, this would essentially rule out the idea of an RNA World that depended on an RNA polymerase ribozyme. However, the central result that we show here is that, while short parasites are indeed lethal to the polymerase when alone, termination errors generate a mixture of parasites of different lengths, and this mixture is not always lethal. The polymerase survives in the presence of the mixed parasites up to a finite error threshold.

2. Methods

The model used is an extension of that in Kim and Higgs (2016). We use a square lattice in which each site can either be vacant or occupied by a single RNA strand. A strand is either a polymerase (P), the complement to a polymerase (C), or a parasite (X). Lengths of strands are measured in numbers of nucleotides. Polymerases and complements have a fixed length L_{pol}. Parasites may have any length up to and including L_{pol}. We assume that the surface environment limits diffusion, thus preventing strands from moving between lattice sites. The model allows only polymerase-catalysed replication and assumes that non-enzymatic template-directed replication is negligible. For a similar model that incorporates both kinds of replication, see Wu and Higgs (2012) and Shay et al. (2015). Polymerases replicate neighbouring template strands at rate $k(L)$, where L is the length of the template. The simplest case, which was also used by Kim and Higgs (2016), is to assume that $k(L)$ is inversely proportional to L. It is convenient to write

$$k(L) = k_{pol}\frac{L_{pol}}{L}, \tag{1}$$

where k_{pol} is the replication rate constant for a strand of length L_{pol}. We also introduce a cutoff length, L_{cut}, which is the minimum length of template that can be replicated. This is motivated by experiments with Qβ replicase, where the optimum RNA strands for replication have a length much shorter than the full virus RNA, but must be long enough to have a secondary structure that is recognized by the Qβ replicase protein (Biebricher and Luce 1992, 1993). For all the simulations in this paper, we have $L_{pol} = 100$ and $L_{cut} = 10$. For simplicity, we assume that one strand occupies one lattice site, irrespective of its length. However, if strands shorter than L_{cut} are generated by termination errors, these strands are too short to be replicated again, and they are also assumed to be too short to take up a lattice site. These very short strands are simply ignored because they play no role in the model. The form for $k(L)$ in Eq. (1) is the worst-case scenario for survival of the polymerase, because it gives the short parasites the maximum advantage. Later in the paper we will consider more realistic forms for $k(L)$ in which the replication rate is less strongly length dependent.

The model also incorporates loss of strands. Strand sites are turned into vacancies at a rate that is assumed constant for all lengths, and is set to 1. This provides a scale for comparison of the other rates in the model. The loss rate represents either the escape of a strand from the surface or the breakdown of a strand back to individual monomers.

The simulations proceed in time steps of length δt. In one time step, we visit every strand in a random order and give it a chance to be a template. For each strand, we select one of the eight neighbouring sites at random from the Moore neighbourhood. If this site is occupied by a polymerase, the template is replicated with a probability $k(L)\delta t$. Only strands next to polymerases can be replicated. The new strand is the complement of the template: a P creates a C, and a C creates a P. Note that C strands are non-functional, but they are necessary for replication of the polymerases. If a point mutation occurs, a parasite of length L_{pol} is generated, instead of a P or C. If a termination error occurs, a parasite of a length less than L_{pol} is generated (further details below). If the template is a parasite of length L, accurate replication creates another parasite of the same length. Since all parasites of a given length are equivalent in this model we do not keep track of their plus and minus forms. Point mutations are not relevant to parasites for the same reason. Premature termination of replication of a parasite can generate shorter parasites.

When a new strand is created, we select a second random neighbour site of the template strand different from the site occupied by the polymerase. If the second neighbour site is a vacancy, the new strand is placed on this site. If the second site is already occupied by a strand, the new strand is eliminated and no change occurs. After giving each strand a chance to replicate, we again go through each strand in a random order and give it a chance to be lost/broken down. This occurs with a probability δt, because the loss rate is defined as 1. This completes one time step δt.

We now discuss errors in more detail. For point mutations, we assume an error probability m_{point} per nucleotide. The probability of at least one point mutation occurring during replication of a sequence of length L is

$$M(L) = 1 - (1 - m_{\text{point}})^{L}. \tag{2}$$

P and C sequences have length L_{pol}. Hence, when a P or C is replicated, there is a probability $M(L_{pol})$ that the new strand is a parasite of length L_{pol}.

For termination errors, we assume that there is a probability m_{term} of premature termination at each nucleotide. Let $p(l)$ be the probability of generation of a new strand of length l from a template of length L.

$$\begin{aligned} p(l) &= (1 - m_{term})^{l} m_{term} && (\text{for } 0 \le l \le \text{L} - 1) \\ p(L) &= (1 - m_{term})^{L} && (\textit{accurate replication}) \end{aligned} \tag{3}$$

This is normalized so that $\sum_{l=0}^{L} p(l) = 1$. There is a subtlety here regarding the replication rate. The rate of production of the new strand should depend on its own length, l, not on the length of the template, L. Hence, when a template is about to be replicated, we first determine the length of the product l as a random value from the distribution $p(l)$. If $l < L$, the probability of the replication occurring is $k(l)\delta t$, rather than $k(L)\delta t$. In simulations that include both kinds of error, we first check for premature termination, then if the strand is accurately replicated according to Eq. (3), we check for point mutations according to Eq. (2).

All results reported here are from simulations using a square lattice of size 1024 × 1024 with periodic boundaries to limit edge effects. The time step is $\delta t = 0.001$ in all cases, except for runs with very high polymerization rates ($k_{pol} \ge 1000$), in which case it was necessary to decrease the time step to $\delta t = 0.0001$.

3. Polymerase survival with point mutation errors

We will first consider the simplest case with only point mutations and no termination errors. Due to the clustering of polymerases that arises in this spatial model, it is less likely for a parasite to be adjacent to a polymerase ribozyme than for a polymerase to be adjacent to another polymerase or complement. This means that parasites have a disadvantage, and they die out if there is no continued mutation ($m_{point} = 0$). For moderate m_{point}, there is coexistence of the parasites with the polymerase and complement. An example of this situation is shown in Fig. 1(a). For the corresponding animation see Video S1.

Fig. 2 shows the time averaged strand concentrations as a function of m_{point} when all other variables are fixed. The error threshold is close to $m_{point} = 2.5 \times 10^{-3}$ for these parameters, which means the mutation probability per strand is $M = 0.22$ for a strand of length $L_{pol} = 100$. If the error threshold occurs at M of order 1, then the maximum per-base error rate is m_{point} of order $1/L_{pol}$. This is a similar conclusion to the usual error threshold in the master-sequence landscape. Eigen et al. (1988) showed that the minimum fidelity in that case is $Q = 1/\sigma_0$, where σ_0 is the relative replication rate (superiority parameter) of the master sequence to the mutants. In our case, however, the polymerase survives not because of a replication rate advantage, but due to an advantage arising from spatial clustering.

The example in Fig. 2 is calculated for $k_{pol} = 25$. Fig. 3 shows the way the error threshold in m_{point} depends on k_{pol} (with L_{pol} fixed at 100). Firstly, we note that, even in absence of errors, there is a minimum value of the replication rate necessary for the survival of the polymerase. This is close to 8.6. Below this, the rate of loss/breakdown of strands is faster than the multiplication rate. The error threshold in Fig. 3 is therefore zero below $k_{pol} = 8.6$. Above this, the error threshold increases as k_{pol} increases, because speed partially compensates for accuracy. If the replication rate is larger, then the number of accurate copies produced from one strand in its lifetime is larger, which helps the survival of the polymerase, even if mutant copies are also produced. However, Fig. 3 shows that the error threshold begins to decrease again very slowly at very high replication rates above $k_{pol} = 50$. This is because the spatial structure of the model breaks into very small clusters of polymerase and complement that are effectively "walled-in" by parasites. The parasites stop the polymerase clusters from growing, even though the parasites cannot spread. The concentration of the polymerases actually decreases with k_{pol} over this range, and the error threshold also decreases slightly in consequence.

We do not have an exact calculation of strand concentrations or the error thresholds for this model. An approximate solution can be found using a pair approximation that considers correlations in the states of pairs of neighbouring sites but ignores correlations beyond two sites. This calculation is shown in the Appendix, and the strand concentrations as a function of m_{point} are shown in Fig. A1. The result is qualitatively similar to Fig. 2, although the value of the error threshold found from the approximation is substantially higher than that found from simulation of the lattice model.

4. The existence of lethal parasites

As pointed out in the previous section, parasites of the same length as the polymerase have a disadvantage due to spatial clustering of the polymerases. Therefore, these parasites die out unless they are continually created by mutation. However, shorter parasites have an advantage due to faster replication which can overcome the disadvantage due to clustering. We already gave examples (Kim and Higgs, 2016) where short independent parasites coexist with the polymerase, and where very short parasites destroy the polymerase entirely. Here, we investigate how short the parasites need to be in order to be lethal.

A simulation of just P and C sequences was allowed to reach a steady state without mutations or parasites of any kind. A small number of parasites of a fixed length L was then added in order to see if these parasites could invade the replicating system. The mutation rates were kept at zero, so the only parasites present are copies of the initial few that were added.

Three outcomes are possible, depending on L. For $L > L_{max}$, the parasites die out. For $L_{min} \le L \le L_{max}$, the parasites coexist with the polymerase and complement. For $L < L_{min}$, the parasites destroy the polymerases and everything dies out. Fig. 4 shows the time averaged concentrations of the strands as a function of L. For these parameters, where $k_{pol} = 25$ and $L_{pol} = 100$, we estimate $L_{min} = 14$ and $L_{max} = 76$. This problem can also be studied approximately using the pair approximation shown in the Appendix. The results shown in Fig. A2 are qualitatively similar to those in Fig. 4.

In contrast to the case of point mutations (Fig. 1a), the spatial dynamics in the case of coexisting parasites and polymerases without mutations gives rise to travelling waves (Fig 1b and 1c). Polymerases and complements form the leading edge of the waves while parasites survive on the trailing edge. Travelling waves in similar models to this have been observed previously (Takeuchi and Hogeweg, 2012; Colizzi and Hogeweg, 2016a,b). For the longest parasites in the range $L_{min} \le L \le L_{max}$, the traveling waves are small and constantly colliding (Fig. 1b) whereas shorter parasites result in larger traveling waves (Fig 1c). The large scale travelling waves for the shorter parasites are only possible if the lattice size is large enough, *i.e.* short parasites are more lethal if the lattice size

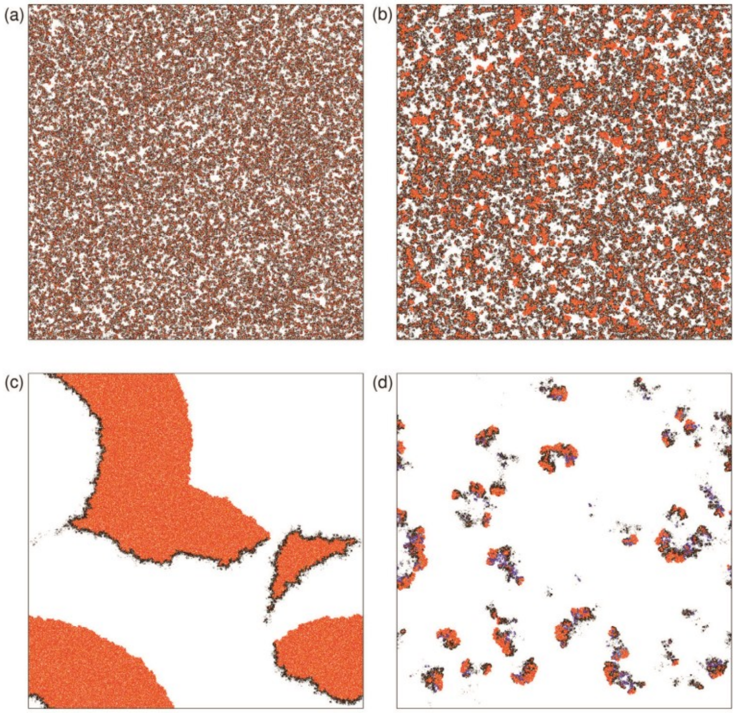


Fig. 1. Snapshots from simulations, where polymerases are shown in red, complements in orange, and parasites in black. In all cases $k_{pol} = 25$ and $\delta t = 0.001$. (a) Small clusters are seen in the case of point mutation rate $m_{point} = 1.6 \times 10^{-3}$ and no termination errors. (b) Small chaotic waves emerge when parasites of fixed length 45 coexist with polymerases and complements and no replication errors are allowed. (c) Large travelling waves are seen when parasites are shortened to length 15 under the same constraint of no replication errors. (d) When termination errors occur with $m_{term} = 1.0 \times 10^{-5}$, small waves emerge which constantly collide, split, and die. Parasites are coloured according to length, from black (short) to long (light blue). "(For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)".

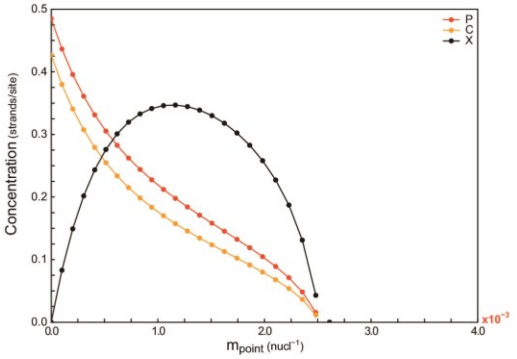


Fig. 2. Concentrations of polymerase, complements, and parasites are shown as a function of m_{point}, with $k_{pol} = 25$ and $\delta t = 0.001$. The point mutation error threshold occurs around 2.5×10^{-3} where the average polymerase population goes to zero. All simulations were run until $t = 1000$ and repeated 100 times for each value of m_{point} shown. Each data point represents time and simulation averaged strand concentration if at least 5% of the trials had a surviving polymerase population, and 0 otherwise.

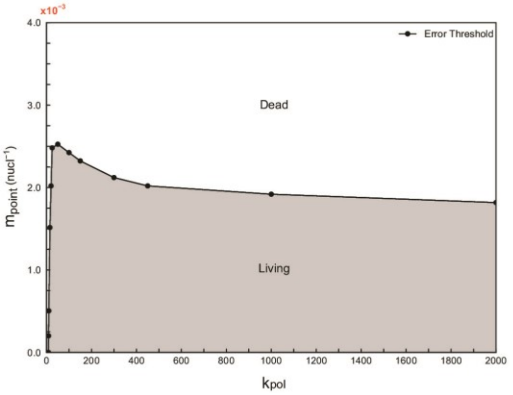


Fig. 3. The point mutation error threshold is shown as a function of k_{pol}. For each value of k_{pol}, simulations were run until $t = 1000$ and repeated 100 times for increasing values of m_{point}. Each data point represents the largest value of m_{point} for which at least 5% of the trials had a surviving polymerase population.

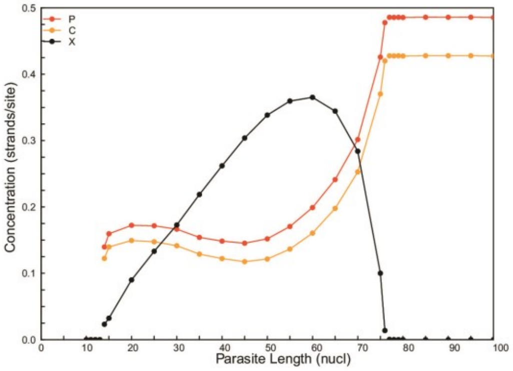


Fig. 4. Equilibrium concentrations of polymerase, complements, and parasites are shown as a function of parasite length. Coexistence of polymerases and parasites is possible when parasites have lengths in the range 14–76 nucleotides. All simulations were initialized with a small population of parasites with fixed length added to a polymerase population, and no point mutations or termination errors were allowed. Each simulation was run until $t = 1000$ and repeated 100 times for each length of parasite shown. Each data point represents time and simulation averaged strand concentration if at least 5% of the trials had a surviving polymerase population, and 0 otherwise. Common parameters used are $k_{pol} = 25$ and $\delta t = 0.001$.

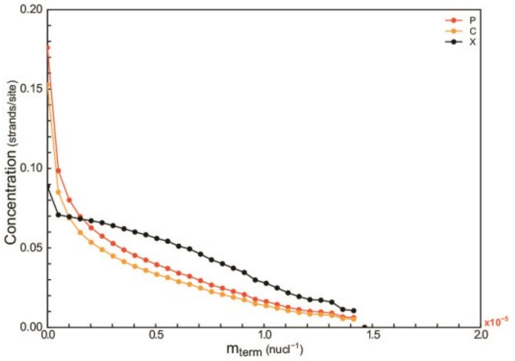


Fig. 5. Concentrations of polymerase, complements, and parasites are shown as a function of m_{term}, with $k_{pol} = 25$ and $\delta t = 0.001$. The termination error threshold occurs around 1.5×10^{-5}, a factor of 100 times smaller than the point mutation error threshold. All simulations were run until $t = 1000$ and repeated 100 times for each value of m_{term} shown. Each data point represents time and simulation averaged strand concentration if at least 5% of the trials had a surviving polymerase population, and 0 otherwise.

is smaller. Animations of these cases are also available. Video S2 shows the case where $L < L_{min}$, where introduction of a few short parasites is lethal. Videos S3 and S4 show travelling waves.

The existence of short, lethal parasites is unsettling, as termination errors will inevitably generate parasites with $L < L_{min}$. One solution to this problem would be to evolve a polymerase that does not bind to templates that are too short. As we discussed in the introduction, this does seem to be the case for RNA replication by the Qβ replicase protein (Biebricher and Luce 1992, 1993). This feature is included in our model via the cutoff length L_{cut}, which is the minimum length template strand that can bind to the polymerase and be replicated. It is clear that if we set L_{cut} larger than L_{min}, the parasites that are short enough to be lethal cannot be replicated, which eliminates the problem of lethal parasites immediately. However, the main aim of this paper is to look at the case of termination errors, which continually generate parasites of a mixture of lengths. We will now show that the short parasites are *not* lethal when present in this mixture, and that the polymerase can coexist with the mixture of parasites, even if $L_{min} > L_{cut}$.

5. Polymerase survival with termination errors

In this section, we consider simulations with termination errors but no point mutations. Parasites are generated of all lengths shorter than the template, with a probability distribution $p(l)$ as described in the Methods section. When parasites of all lengths are present, a new equilibrium emerges in which competition between parasites of varying lengths prevents the lethality of short strands. This stable state (Fig 1d) is visually distinct from the previous cases (Figs 1a–c), in that small traveling waves arise that are separated by large amounts of empty space. The wave structures are rather irregular and can split to generate new waves. Waves also collide sometimes, which tends to lead to death of the colliding waves because they are surrounded by parasites on all sides. For animations of polymerases surviving with termination errors see Videos S5-S7.

For the parameters in this example, $L_{min} = 14$, and $L_{cut} = 10$. This means that potentially lethal parasites in the range $L = 10–13$ can be replicated. Nevertheless, the system survives. Fig. 5 shows the concentration of polymerase and complement as well as the total concentration of parasite strands of all lengths as a function of m_{term}. These simulations were done in the following way. For each value of m_{term}, simulations were run until $t = 1000$ and repeated 100 times. Each data point represents the time and simulation averaged parasite concentration of all simulations which had a surviving polymerase population. Common parameters used are $k_{pol} = 25$ and $\delta t = 0.001$.

It can be seen in Fig. 5 that the system survives with nonzero m_{term} up to a critical error threshold, in a similar way as for point mutations. The reason for this seems to be the fragmentary structure of the travelling waves. Since medium-length parasites coexist with polymerases in a traveling wave, the emergence of a short parasite is no longer lethal as it now has to compete for vacant sites with all non-lethal parasites present on the wave edge. We also observed occasions where enough short parasites accumulated to encapsulate and destroy a wave. However, since waves are widely separated from each other, the death of an individual wave does not result in the death of the entire polymerase population. Wave death is offset by the formation of new waves, which occurs when a wave splits due to the emergence of an internal parasite, or the escape of a polymerase from the trailing edge of the wave. This is an example of multi-level evolution acting at the higher level of the wave as well as the lower level of the single molecule (see also Takeuchi and Hogeweg, 2012; Colizzi and Hogeweg, 2016a,b).

The key point up to now is that the system with continued creation of parasites of all lengths by termination errors is stable up to a finite error threshold, even though the system with only very short parasites would be unstable, even with zero mutation rate. We can therefore be satisfied that the RNA World is not inevitably destroyed by the existence of termination errors. Nevertheless, comparison of Figs. 2 and 5 shows an important point. The error threshold for termination errors is two orders of magnitude smaller than for point mutations for the parameters we investigated ($m_{term} = 1.5 \times 10^{-5}$ in comparison to $m_{point} = 2.5 \times 10^{-3}$). This is not surprising, since termination errors result in parasites with a larger replication rate than those resulting from point errors. It does raise a substantial worry as to whether such low values of m_{term} could be achieved by the earliest ribozymes.

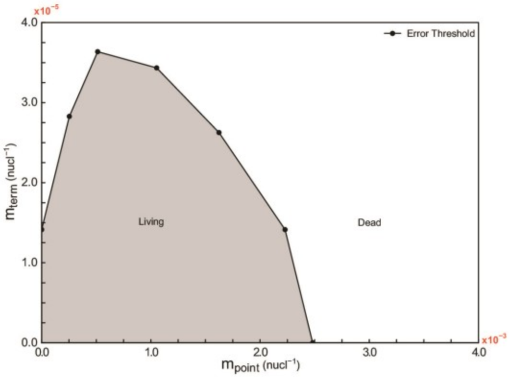


Fig. 6. The termination error threshold is shown as a function of point mutation rate. The shaded gray area corresponds to the combinations of point mutation and termination error rates for which the polymerase survives. For each value of m_{point}, simulations were run until $t = 1000$ and repeated 100 times for increasing values of m_{term}. Each data point represents the largest value of m_{term} for which at least 5% of the trials had a surviving polymerase population. Common parameters used are $k_{pol} = 25$ and $\delta t = 0.001$.

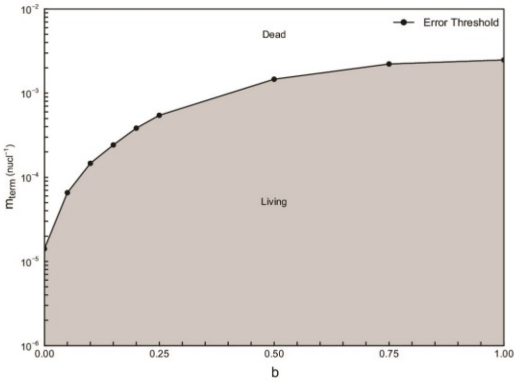


Fig. 7. The termination error threshold is shown as a function of the binding variable b. For each value of b, simulations were run until $t = 1000$ and repeated 100 times for increasing values of m_{term}. Each data point represents the largest value of m_{term} for which at least 5% of the trials had a surviving polymerase population. Common parameters used are $k_{pol} = 25$ and $\delta t = 0.001$.

Another difference between Figs. 2 and 5 is that, for the termination errors, the concentration of parasites approaches a nonzero constant as m_{term} tends to zero, whereas for point mutations, the concentration of parasites tends to zero as m_{point} tends to zero. This is because medium length parasites coexist with the polymerase when there is no mutation, whereas parasites of length L_{pol} do not coexist with the polymerase in absence of continued mutation. All simulations reported here were initialized with 15% polymerases, 15% complementary sequences, and 15% parasites of random lengths. The point at $m_{term} = 0$ in Fig. 5 appears at a nonzero parasite concentration because parasites were initially present and medium length parasites coexist with the polymerase indefinitely, even without mutation (as we showed in Fig. 4). Clearly, if we began with no parasites, the parasite concentration would remain zero when $m_{term} = 0$. In contrast, in the point mutation case in Fig. 2, it doesn't matter whether a few parasites are present initially or not, because parasites of length L_{pol} would disappear anyway when $m_{point} = 0$.

6. Polymerase survival with point and termination errors

Now that we have shown polymerase ribozymes can survive in the presence of each type of error separately, we will consider the case in which both errors are possible. While it would seem that combining point mutations and termination errors would inevitability result in a new error threshold that is lower than either individually, this is not the case. The presence of a certain amount of point mutations in fact increases the termination error threshold by a factor of more than 2 in the best cases relative to the case with only termination errors (Fig. 6). Increasing the point error rate results in a new source of long parasites that limit the replicative advantage of short parasites that are competing for vacant sites on the trailing edge of a wave. This competition hinders short parasites, thereby allowing the polymerase population to survive for higher termination error rates.

7. Distinguishing binding and nucleotide addition steps

So far, we supposed that the replication rate was inversely proportional to the strand length. This is the worst-case scenario, because it gives maximum advantage to short parasites. Here we consider a slightly more realistic model of replication that distinguishes an initial step of binding of the polymerase to the template and a step of nucleotide addition. Let T_B be the mean time for binding. We assume that this time is the same for templates of all lengths of at least L_{cut}, and that templates shorter than L_{cut} cannot bind at all. Let T_A be the mean time for adding one nucleotide. The mean time for copying a polymerase sequence of length L_{pol} is $T_{pol} = T_B + T_A L_{pol}$. The fraction of time spent in the binding step is $b = T_B/T_{pol}$, while the fraction spent in the nucleotide addition steps is $(1 - b) = T_A L_{pol}/T_{pol}$. We will keep the model simple by treating replication as a single effective step with a rate that is the inverse of the mean time. Hence for the polymerase, the rate is $k_{pol} = 1/T_{pol}$, and for a strand of length L, the rate is

$$k(L) = \frac{1}{T_B + T_A L} = k_{pol} \frac{L_{pol}}{bL_{pol} + (1 - b)L}. \tag{4}$$

If we set $b = 0$, there is no time spent on binding, and we are in the worst case (same as Eq. (1)). If we set $b = 1$, the time for nucleotide addition is negligible compared to the binding time. In this case, all strands replicate at rate k_{pol} irrespective of their length. For an intermediate value of b, $k(L)$ increases for shorter parasites, but less drastically than in the worst case.

Fig. 7 shows results of simulations in which the replication rate follows Eq. (4) with different values of b. Termination errors occur at rate m_{term} and no point mutations are included. The termination error threshold is shown as a function of the binding proportion b. When $b = 0$, the error threshold is equivalent to our previous case of termination errors (Fig. 5). As the binding proportion increases, the termination error threshold increases by two orders of magnitude, *i.e.* when $b > 0$, the system is much more tolerant of termination errors than in the worst possible case, and short parasites are much less dangerous. When $b = 1$, all parasites have the same replication rate, and their length is not important. Hence termination errors are equivalent to point mutation errors, and the error threshold for m_{term} when $b = 1$ in Fig. 7 is the same as the error threshold for m_{point} in Fig. 2. These results highlight the importance of the binding step, as even a 10% binding proportion is enough to increase the termination error threshold by a factor of 10.

8. Discussion

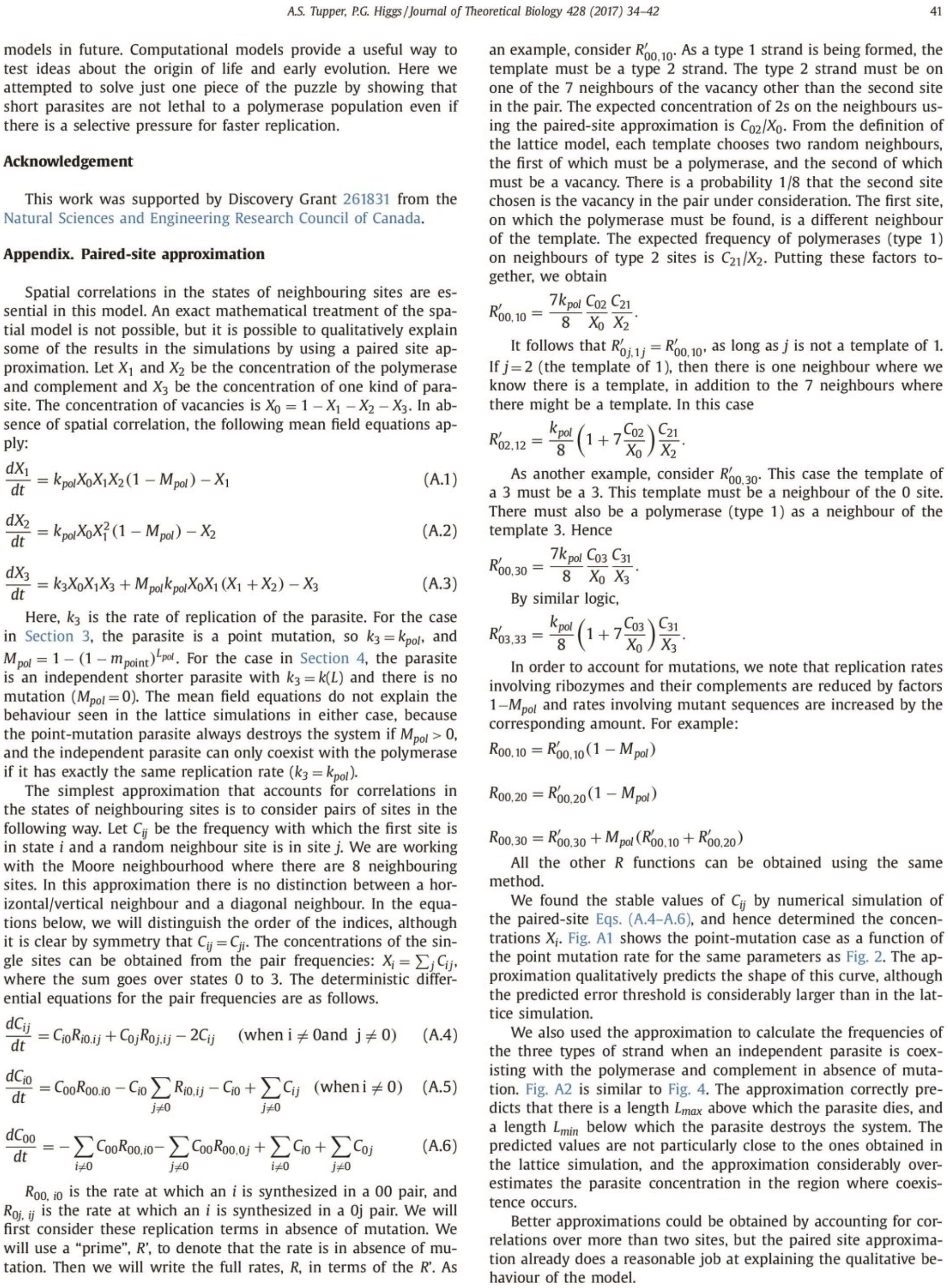
Parasites are likely to be important in the RNA world as it is easy for a sequence to be a template and difficult for it to be a polymerase. We have presumed that mutations that prevent the function of the polymerase will produce sequences that are still viable templates. Therefore, no special adaptation of the template sequence is required for it to act as a parasite. Nevertheless, selection on parasites will tend to increase their replication rate, and decreasing the template length is an easy way to do this without requiring any special adaptation. Hence, termination errors will provide a constant source of shorter and shorter parasites. Using our computational model we have shown that survival of a polymerase ribozyme is possible in a surface model when both point mutations and termination errors are considered. Hence, the tendency for selection of shorter parasites does not inevitably kill the polymerases.

In both the cluster patterns that arise in the point mutation case and the travelling wave patterns that arise in the termination error case, the polymerases are more likely to be next to other polymerases and complements and less likely to be next to parasites than they would be in the well-mixed case. This gives an advantage to the polymerases that allows them to survive up to a finite error threshold, whereas the polymerases are destroyed by parasites in the well mixed case for any non-zero error rate. There are several other models that show that spatial pattern formation can allow the survival of polymerases in the presence of parasites (Takeuchi and Hogeweg, 2012; Colizzi and Hogeweg, 2016a,b). These models also show travelling wave patterns similar to ours. These models allow the rate parameters of the parasites to evolve without explicitly considering the length of the template. Including the length as a parameter in our case makes the comparison possible between the termination errors and point mutations. We have also investigated the factors that affect the error threshold, which was not done previously.

We showed in Section 3 that the error threshold in the per-sequence error rate M that is achievable by spatial clustering can be relatively large, but this still implies a per base error rate that must be very small and varies inversely with the length of the template. The value of the error threshold depends on many details, including whether the functional sequence is maintained by replication rate advantage or by clustering (as in this paper), the presence of neutral networks in the fitness landscape (Reidys et al., 2001; Wilke 2001; Takeuchi et al., 2005; Szilagyi et al., 2014) and the possibility of recombination (Santos et al., 2004). While these factors make quantitative differences, they do not really change the nature of the problem: replication must be accurate in order to sustain an RNA World. Szilagyi et al. (2014) calculate the phenotypic error threshold, and conclude that known ribozymes of lengths up to about 200 could be successfully replicated by currently known polymerases with per-base error rates of a few percent (Johnston et al., 2001; Wochner et al., 2011). However, they assume that there is a very large replication rate advantage to the sequences with the correct secondary structure, and that fitness only depends on the secondary structure. Furthermore, they assume that the polymerase is fixed, and is not itself subject to replication, whereas in the RNA World, the polymerase has to replicate other copies of itself, so a mechanism such as spatial clustering or compartmentalization is required to prevent the invasion of parasites. Hence, the conclusions of Szilagyi et al. (2014) seem somewhat optimistic to us. Nevertheless, we do not wish to argue against the RNA World hypothesis. In our view, there is a lot of evidence that supports the existence of an RNA World in the early stages of life on Earth (Higgs and Lehman, 2015), and theoretical treatments demonstrate that small per-base error rates are required for this to work. Therefore it becomes an experimental problem to demonstrate that sufficiently accurate polymerase ribozymes are possible. Work on error rates in non-enzymatic replication is also relevant here (Rajamani et al., 2010), which demonstrates that stalling of replication after an error slows down the replication of sequences with errors, giving an advantage to correctly copied sequences and an increase in the error threshold. This same effect could also occur in ribozyme catalysed replication.

We would like to summarize several important points that emerge from our studies in this paper. Short parasites are clearly lethal in this model when they are introduced into a connected system of polymerases and complements. Nevertheless, when parasites of all lengths are present, as is the case for termination errors, the lethality of short parasites is prevented because the system is broken up into separate fragments, and the lethal parasites cannot travel through the whole system. Our model allows us to compare the error thresholds from point mutations and termination errors. In general termination errors are more dangerous than point mutations because they create faster replicating parasites. Hence The error threshold in m_{term} was found to be two orders of magnitude less than that for m_{point} in the worst case, where replication rate varies inversely with template length. Competition between parasites of different lengths was further extended to the case in which we considered both errors simultaneously and showed that the addition of a point mutation rate in fact increased the termination error threshold above that which occurs with termination errors alone. This is again because the presence of non-lethal, long parasites created by point mutations changes the spatial pattern in which the shorter parasites evolve. An interesting factor that we did not yet consider is that point mutations are likely to impose a stalling effect on a polymerase similar to the stalling effect seen in non-enzymatic replication (Rajamani et al., 2010), and this might lead to an increase in the termination likelihood. Lastly, we showed that when a binding step was incorporated in the replication process, termination errors become much less dangerous than in the worst case because the relative advantage of the short parasites is reduced. We found that the termination error threshold approaches the point mutation threshold if the binding step is long compared to the polymerization step.

This paper therefore supports the idea that a polymerase ribozyme replicating on a surface may have supported the early stages of life. However, one limitation that should be borne in mind, is that in the present paper, there is no diffusion of strands across the surface. We already showed in the case with point mutations only (Kim and Higgs, 2016) that the incorporation of strand diffusion benefits parasites and reduces the error threshold. An alternative mechanism that prevents parasites from destroying polymerase systems is to place the replicators in protocells, in which case group selection at the cell level can overcome individual selection at the strand level (Takeuchi and Hogeweg, 2009; Higgs and Lehman (2015) and references therein). In an interesting recent experimental realization of the protocell case (Matsumura et al., 2016), it was also found that parasites of high replication rate do not inevitably destroy the system. Thus, both spatial surface models and protocell models agree that polymerase systems can survive, although it is still not clear which of these two factors was more important in actual evolutionary history, and whether there were surface-based replicating systems prior to the origin of cells. Another important question for early life is how a system with a single kind of functional ribozyme could evolve additional functions. Kim and Higgs (2016) showed that an unlinked strand functioning as a nucleotide synthetase can coexist and cooperate with the polymerase in some cases. We suggested that building up a metabolism controlled by many ribozymes might be easier in a proto-cell model than a surface based model, although this has not yet been fully tested. It will be also interesting to consider the relative sizes of error thresholds in proto-cell and surface based



models in future. Computational models provide a useful way to test ideas about the origin of life and early evolution. Here we attempted to solve just one piece of the puzzle by showing that short parasites are not lethal to a polymerase population even if there is a selective pressure for faster replication.

Acknowledgement

This work was supported by Discovery Grant 261831 from the Natural Sciences and Engineering Research Council of Canada.

Appendix. Paired-site approximation

Spatial correlations in the states of neighbouring sites are essential in this model. An exact mathematical treatment of the spatial model is not possible, but it is possible to qualitatively explain some of the results in the simulations by using a paired site approximation. Let X_1 and X_2 be the concentration of the polymerase and complement and X_3 be the concentration of one kind of parasite. The concentration of vacancies is $X_0 = 1 - X_1 - X_2 - X_3$. In absence of spatial correlation, the following mean field equations apply:

$$\frac{dX_1}{dt} = k_{pol}X_0X_1X_2(1 - M_{pol}) - X_1 \tag{A.1}$$

$$\frac{dX_2}{dt} = k_{pol}X_0X_1^2(1 - M_{pol}) - X_2 \tag{A.2}$$

$$\frac{dX_3}{dt} = k_3X_0X_1X_3 + M_{pol}k_{pol}X_0X_1(X_1 + X_2) - X_3 \tag{A.3}$$

Here, k_3 is the rate of replication of the parasite. For the case in Section 3, the parasite is a point mutation, so $k_3 = k_{pol}$, and $M_{pol} = 1 - (1 - m_{\text{point}})^{L_{pol}}$. For the case in Section 4, the parasite is an independent shorter parasite with $k_3 = k(L)$ and there is no mutation ($M_{pol} = 0$). The mean field equations do not explain the behaviour seen in the lattice simulations in either case, because the point-mutation parasite always destroys the system if $M_{pol} > 0$, and the independent parasite can only coexist with the polymerase if it has exactly the same replication rate ($k_3 = k_{pol}$).

The simplest approximation that accounts for correlations in the states of neighbouring sites is to consider pairs of sites in the following way. Let C_{ij} be the frequency with which the first site is in state i and a random neighbour site is in site j. We are working with the Moore neighbourhood where there are 8 neighbouring sites. In this approximation there is no distinction between a horizontal/vertical neighbour and a diagonal neighbour. In the equations below, we will distinguish the order of the indices, although it is clear by symmetry that $C_{ij} = C_{ji}$. The concentrations of the single sites can be obtained from the pair frequencies: $X_i = \sum_j C_{ij}$, where the sum goes over states 0 to 3. The deterministic differential equations for the pair frequencies are as follows.

$$\frac{dC_{ij}}{dt} = C_{i0}R_{i0,ij} + C_{0j}R_{0j,ij} - 2C_{ij} \quad (\text{when i} \neq 0\text{and j} \neq 0) \tag{A.4}$$

$$\frac{dC_{i0}}{dt} = C_{00}R_{00,i0} - C_{i0}\sum_{j\neq0} R_{i0,ij} - C_{i0} + \sum_{j\neq0} C_{ij} \quad (\text{when i} \neq 0) \tag{A.5}$$

$$\frac{dC_{00}}{dt} = -\sum_{i\neq0} C_{00}R_{00,i0} - \sum_{j\neq0} C_{00}R_{00,0j} + \sum_{i\neq0} C_{i0} + \sum_{j\neq0} C_{0j} \tag{A.6}$$

$R_{00,\,i0}$ is the rate at which an i is synthesized in a 00 pair, and $R_{0j,\,ij}$ is the rate at which an i is synthesized in a 0j pair. We will first consider these replication terms in absence of mutation. We will use a "prime", R', to denote that the rate is in absence of mutation. Then we will write the full rates, R, in terms of the R'. As an example, consider $R'_{00,10}$. As a type 1 strand is being formed, the template must be a type 2 strand. The type 2 strand must be on one of the 7 neighbours of the vacancy other than the second site in the pair. The expected concentration of 2s on the neighbours using the paired-site approximation is C_{02}/X_0. From the definition of the lattice model, each template chooses two random neighbours, the first of which must be a polymerase, and the second of which must be a vacancy. There is a probability 1/8 that the second site chosen is the vacancy in the pair under consideration. The first site, on which the polymerase must be found, is a different neighbour of the template. The expected frequency of polymerases (type 1) on neighbours of type 2 sites is C_{21}/X_2. Putting these factors together, we obtain

$$R'_{00,10} = \frac{7k_{pol}}{8}\frac{C_{02}}{X_0}\frac{C_{21}}{X_2}.$$

It follows that $R'_{0j,1j} = R'_{00,10}$, as long as j is not a template of 1. If $j = 2$ (the template of 1), then there is one neighbour where we know there is a template, in addition to the 7 neighbours where there might be a template. In this case

$$R'_{02,12} = \frac{k_{pol}}{8}\left(1 + 7\frac{C_{02}}{X_0}\right)\frac{C_{21}}{X_2}.$$

As another example, consider $R'_{00,30}$. This case the template of a 3 must be a 3. This template must be a neighbour of the 0 site. There must also be a polymerase (type 1) as a neighbour of the template 3. Hence

$$R'_{00,30} = \frac{7k_{pol}}{8}\frac{C_{03}}{X_0}\frac{C_{31}}{X_3}.$$

By similar logic,

$$R'_{03,33} = \frac{k_{pol}}{8}\left(1 + 7\frac{C_{03}}{X_0}\right)\frac{C_{31}}{X_3}.$$

In order to account for mutations, we note that replication rates involving ribozymes and their complements are reduced by factors $1 - M_{pol}$ and rates involving mutant sequences are increased by the corresponding amount. For example:

$$R_{00,10} = R'_{00,10}(1 - M_{pol})$$

$$R_{00,20} = R'_{00,20}(1 - M_{pol})$$

$$R_{00,30} = R'_{00,30} + M_{pol}(R'_{00,10} + R'_{00,20})$$

All the other R functions can be obtained using the same method.

We found the stable values of C_{ij} by numerical simulation of the paired-site Eqs. (A.4–A.6), and hence determined the concentrations X_i. Fig. A1 shows the point-mutation case as a function of the point mutation rate for the same parameters as Fig. 2. The approximation qualitatively predicts the shape of this curve, although the predicted error threshold is considerably larger than in the lattice simulation.

We also used the approximation to calculate the frequencies of the three types of strand when an independent parasite is coexisting with the polymerase and complement in absence of mutation. Fig. A2 is similar to Fig. 4. The approximation correctly predicts that there is a length L_{max} above which the parasite dies, and a length L_{min} below which the parasite destroys the system. The predicted values are not particularly close to the ones obtained in the lattice simulation, and the approximation considerably overestimates the parasite concentration in the region where coexistence occurs.

Better approximations could be obtained by accounting for correlations over more than two sites, but the paired site approximation already does a reasonable job at explaining the qualitative behaviour of the model.

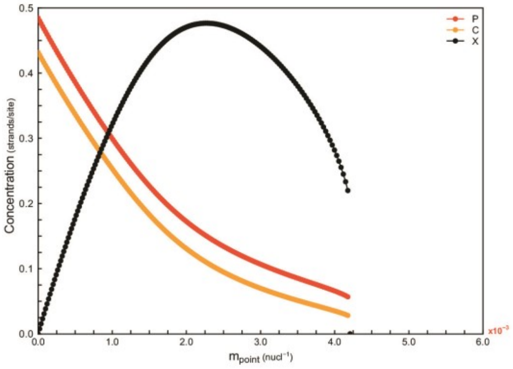


Fig A1. Error threshold as a function of mutation according to the paired-site approximation. Polymerase P$=X_1$, Complement C$=X_2$, Parasite $X=X_3$.

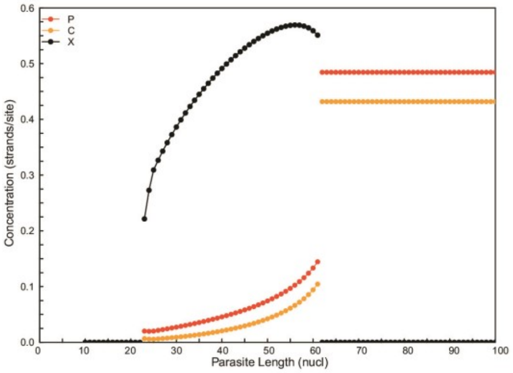


Fig A2. Range of L for coexistence of an independent parasite with a polymerase. Polymerase P$=X_1$, Complement C$=X_2$, Parasite $X=X_3$.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jtbi.2017.05.037.

References

Attwater, J., Wochner, A., Holliger, P., 2013. In-ice evolution of RNA polymerase ribozyme activity. Nat. Chem. 5 (12), 1011–1018.

Bartel, D.P., Unrau, P.J., 1999. Constructing an RNA world. Trends Biochem. Sci. 24 (12), M9–M13.

Biebricher, C.K., Luce, R., 1992. *In vitro* recombination and terminal elongation of RNA by Qβ replicase. EMBO J. 11, 5129–5135.

Biebricher, C.K., Luce, R., 1993. Sequence analysis of RNA species synthesized by Qβ replicase without template. Biochemistry 32, 4848–4854.

Colizzi, E.S., Hogeweg, P., 2016a. Parasites sustain and enhance RNA-like replicators through spatial self-organisation. PLoS Comput. Biol. 12 (4), e1004902.

Colizzi, E.S., Hogeweg, P., 2016b. High cost enhances cooperation through the interplay between evolution and self-organization. BMC Evol. Biol. 16, 31.

Eigen, M., McCaskill, J., Schuster, P., 1988. Molecular quasi-species. J. Phys. Chem. 92 (24), 6881–6891.

Gilbert, W., 1986. Origin of life: the RNA world. Nature 319 (6055).

Higgs, P.G., Lehman, N., 2015. The RNA World: molecular co-operation at the origin of life. Nat. Rev. Genet. 16, 7–17.

Johnston, W.K., Unrau, P.J., Lawrence, M.S., Glasner, M.E., Bartel, D.P., 2001. RNA-catalyzed RNA polymerization: accurate and general RNA-templated primer extension. Science 292 (5520), 1319–1325.

Joyce, G.F., 2002. The antiquity of RNA-based evolution. Nature 418 (6894), 214–221.

Kim, Y.E., Higgs, P.G., 2016. Co-operation between polymerases and nucleotide synthetases in the RNA world. PLoS Comput. Biol. 12 (11), e1005161.

Könnyű, B., Czárán, T., Szathmáry, E., 2008. Prebiotic replicase evolution in a surface-bound metabolic system: parasites as a source of adaptive evolution. BMC Evol. Biol. 8, 267.

Könnyű, B., Czárán, T., 2013. Spatial aspects of prebiotic replicator coexistence and community stability in a surface-bound RNA world model. BMC Evol. Biol. 13, 204.

Lawrence, M.S., Bartel, D.P., 2005. New ligase-derived RNA polymerase ribozymes. RNA 11 (8), 1173–1180.

Ma, W., Hu, J., 2012. Computer simulation on the cooperation of functional molecules during the early stages of evolution. PLoS One 7 (4), e35454.

Ma, W., Yu, C., Zhang, W., Hu, J., 2010. A simple template-dependent ligase ribozyme as the RNA replicase emerging first in the RNA world. Astrobiology 10 (4), 437–447.

Ma, W., Yu, C., Zhang, W., Zhou, P., Hu, J., 2010. The emergence of ribozymes synthesizing membrane components in RNA-based protocells. Biosystems 99 (3), 201–209.

Ma, W., Yu, C., Zhang, W., 2007. Monte Carlo simulation of early molecular evolution in the RNA world. Biosystems 90 (1), 28–39.

Ma, W., Yu, C., Zhang, W., Hu, J., 2007. Nucleotide synthetase ribozymes may have emerged first in the RNA world. RNA. 13 (11), 2012–2019.

Matsumura, S., Kun, A., Ryckelynck, M., Coldren, F., Szilagyi, A., Jossinet, F., Rick, C., Nghe, P., Szathmary, E., Griffiths, A.D., 2016. Transient compartmentalization of RNA replicators prevents extinction due to parasites. Science 354, 1293–1296.

Mills, D.R., Peterson, R.L., Spiegelman, S., 1967. An extracellular Darwinian experiment with a self-duplicating nucleic acid molecule. Proc. Nat. Acad. Sci. 58 (1), 217–224.

Reidys, C., Forst, C.V., Schuster, P., 2001. Replication and mutation on neutral networks. Bull. Math. Biol. 63, 57–94.

Rajamani, S., Ichida, J.K., Antal, T., Treco, D.A., Leu, K, Nowak, M.A., Szostak, J.W., Chen, I.A., 2010. Effect of stalling after mismatches on the error catastrophe in nonenzymatic nucleic acid replication. J. Am. Chem. Soc. 132, 5880–5885.

Santos, M., Zintzaras, E., Szathmáry, E., 2004. Recombination in primeval genomes: a step forward but still a long leap from maintaining a sizeable genome.. J. Mol. Evol. 59, 507–519.

Shay, J.A., Huynh, C., Higgs, P.G., 2015. The origin and spread of a cooperative replicase in a prebiotic chemical system. J. Theor. Biol. 364, 249–259.

Szabó, P., Scheuring, I., Czárán, T., Szathmáry, E., 2002. In silico simulations reveal that replicators with limited dispersal evolve towards higher efficiency and fidelity. Nature 420 (6913), 340–343.

Szilagyi, A., Kun, A., Szathmary, E., 2014. Local neutral networks help maintain inaccurately replicating ribozymes. PLoS One 9 (10), e109987.

Takeuchi, N., Poorthuis, P.H., Hogeweg, P., 2005. Phenotypic error threshold; additivity and epistasis in RNA evolution. BMC Evol. Biol. 5, 9.

Takeuchi, N., Hogeweg, P., 2009. Multilevel selection in models of prebiotic evolution II: A direct comparison of compartmentalization and spatial self-organization. PLoS Comp. Biol. 5 (10), e1000542.

Takeuchi, N., Hogeweg, P., 2012. Evolutionary dynamics of RNA-like replicator systems: a bioinformatic approach to the origin of life. Phys. Life Rev. 9, 219–263.

Walker, S.I., Grover, M.A., Hud, N.V., 2012. Universal sequence replication, reversible polymerization and early functional biopolymers: a model for the initiation of prebiotic sequence evolution. PLoS One 7 (4), e34166.

Wilke, C.O., 2001. Selection for fitness *versus* selection for robustness in RNA secondary structure folding. Evolution 55, 2412–2420.

Wochner, A., Attwater, J., Coulson, A., Holliger, P., 2011. Ribozyme-catalyzed transcription of an active ribozyme. Science 332 (6026), 209–212.

Wu, M., Higgs, P.G., 2012. The origin of life is a spatially localized stochastic transition. Biol. Direct 7, 42.

Zaher, H.S., Unrau, P.J., 2007. Selection of an improved RNA polymerase ribozyme with superior extension and fidelity. RNA 13 (7), 1017–1026.

Chapter 3: Templating and the Emergence of RNA

The contents of this chapter were published in a special issue of *Life* in October 2017 titled *The RNA World and the Origin of Life* (corresponding reference below). The manuscript, figures, and tables in this chapter are used with permission from MDPI, in accordance with the MDPI Open Access Policy.

Paul Higgs and I conceived of and designed the experiments. Kevin Shi and I wrote the computer programs to perform simulations and analyzed the data. Paul Higgs wrote the manuscript.

Tupper, A. S., Shi, K., & Higgs, P. G. (2017). The role of templating in the emergence of RNA from the prebiotic chemical mixture. *Life*, *7*(4), 41.

life

MDPI

Article

The Role of Templating in the Emergence of RNA from the Prebiotic Chemical Mixture

Andrew S. Tupper [1], Kevin Shi [2] and Paul G. Higgs [2,*]

[1] Origins Institute and Department of Biochemistry and Biomedical Science, McMaster University, Hamilton, ON L8S 4L8, Canada; tuppea2@mcmaster.ca

[2] Origins Institute and Department of Physics and Astronomy, McMaster University, Hamilton, ON L8S 4K1, Canada; shik6@mcmaster.ca

* Correspondence: higgsp@mcmaster.ca; Tel.: +1-905-525-9140

Received: 1 August 2017; Accepted: 26 October 2017; Published: 31 October 2017

Abstract: Biological RNA is a uniform polymer in three senses: it uses nucleotides of a single chirality; it uses only ribose sugars and four nucleobases rather than a mixture of other sugars and bases; and it uses only 3′-5′ bonds rather than a mixture of different bond types. We suppose that prebiotic chemistry would generate a diverse mixture of potential monomers, and that random polymerization would generate non-uniform strands of mixed chirality, monomer composition, and bond type. We ask what factors lead to the emergence of RNA from this mixture. We show that template-directed replication can lead to the emergence of all the uniform properties of RNA by the same mechanism. We study a computational model in which nucleotides react via polymerization, hydrolysis, and template-directed ligation. Uniform strands act as templates for ligation of shorter oligomers of the same type, whereas mixed strands do not act as templates. The three uniform properties emerge naturally when the ligation rate is high. If there is an exact symmetry, as with the chase of chirality, the uniform property arises via a symmetry-breaking phase transition. If there is no exact symmetry, as with monomer selection and backbone regioselectivity, the uniform property emerges gradually as the rate of template-directed ligation is increased.

Keywords: RNA world; non-enzymatic template-directed replication; chirality; regioselectivity; prebiotic chemistry; symmetry breaking

1. Introduction

An important feature of life is that it makes frequent use of a well-defined set of molecules, but it does not use many other kinds of molecules that have rather similar chemical properties. This has been called the 'lego principle' [1], and it has been proposed that this feature could be used as a signature of life on other planets. This principle applies to amino acids, where the set of possible amino acids is much larger than the 20 used in biological proteins [2], and to nucleotides, where there is a huge diversity of sugars [3] and nucleobases [4,5] that could potentially form polymers similar to RNA and DNA. Non-living chemistry is governed by thermodynamics and reaction kinetics. Similar molecules will have similar free energies of formation and will undergo similar reactions; hence, they will be produced in similar quantities. Living biochemistry is autocatalytic. A subset of molecules is able to catalyze formation of more of the same set of molecules that it requires for growth and reproduction. Because there is continued flow of energy and matter into a living system, the relative frequencies of different molecules can be maintained far from what would be found in thermodynamic equilibrium.

One example of the way autocatalysis maintains biased sets of molecules is the homochirality observed in nucleic acids and proteins [6]. Theoretical models assume that the formation of molecules of a given enantiomer (*D* or *L*) is catalyzed by molecules of the same handedness. In the simplest model of Frank [7], monomers are autocatalytic; however, more complex models have been studied in which the catalysts are dimers [8] or polymers [9,10]. All these models share the feature that it is the monomer synthesis reaction that is catalyzed (i.e., the synthesis of ribose or ribonucleotides, if we apply this to RNA). Here we consider the alternative that the asymmetric autocatalysis comes from template-directed synthesis of complementary oligomers, rather than from the catalysis of nucleotide synthesis. We assume that oligomers of uniform chirality (either *D* or *L*) are efficient templates for the ligation of shorter oligomers of the same chirality, whereas oligomers of mixed chirality are unable to act as templates. We show that if the template-directed reaction rate is fast, a symmetry-breaking phase transition occurs in which one or other enantiomer dominates the system.

Biological nucleic acids show two other kinds of uniform properties in addition to chirality. Firstly, biology uses a uniform set of monomers, rather than a mixture of many other similar molecules with different sugars, different bases, or both. Only four nucleotides are used in genetic information storage and transcription in DNA and RNA (although many modified nucleotides are used in specific positions in structural RNAs, such as tRNAs). We refer to the question of why these particular nucleotides are used as the monomer selection problem. Secondly, biology uses regular $3'$-$5'$ bonds between ribose sugars rather than a mixture of $3'$-$5'$ and $2'$-$5'$ bonds. We refer to this as the backbone regioselectivity problem. The central point of this paper is that the monomer selection and backbone regioselectivity problems are similar problems to the chirality problem, and we may use a similar theory to explain all three. Our theory depends on two propositions: (1) that uniform oligomers of one kind are templates that catalyze synthesis of further oligomers of the same kind (i.e., the same chirality, the same monomers, or the same bond type); and (2) that uniform oligomers are good templates, but mixed oligomers (i.e., mixed chirality, mixed monomers, or mixed bond types) are not. These two propositions are supported by experiment in several ways, as we will now discuss.

In the case of chirality, Bolli et al., [11] studied template-directed ligation of tetramers of pyranosyl-RNA, and showed that the ligation of the homochiral tetramers is faster by at least two orders of magnitude than the ligation of tetramers in which one of the nucleotides has the opposite chirality. In this example, both propositions are clearly satisfied. With standard RNA, Joyce et al. [12] studied the polymerization of G monomers using a poly(C) template of the *D* enantiomer, and showed that template-directed synthesis of G oligomers is efficient when the G monomers are also of the *D* enantiomer, and not when they are of the *L* enantiomer. When a racemic mixture of D/L monomers was used, the *L* monomers inhibited the growth of the G oligomers to some extent, but the template-directed reaction was still an improvement over the case with no template at all. The chiral inhibition effect was presented as a problem for theories of the origin of life in [12], however we see this effect as part of the solution, in the sense that some kind of chiral inhibition is necessary to drive the symmetry breaking between *D* and *L*. Without this, we would expect the two systems to mix and coexist.

The two propositions also apply with regard to the monomer selection problem. Oligomer duplexes of RNA and DNA have been studied in many combinations [13]. It is found that duplexes of pure RNA or pure DNA have higher melting temperatures than hybrid duplexes (in which one strand is pure RNA and one is pure DNA) and mixed duplexes (in which strands are mixtures of ribonucleotides and deoxyribonucleotides). RNA and DNA are rather similar in structure, and hybridization between the two is clearly possible. Hybridization between the two allows for transcription in today's organisms, and is also essential for the transfer of information from RNA to DNA that is proposed to occur at the end of the RNA World. Scenarios have also been proposed in which DNA arose concurrently with RNA, rather than as a late successor [13]. However, the essential point here is that, even with two very similar kinds of monomers, each has a preference for its own kind, and this will tend to drive separation between the two, as we will see in the theoretical models below.

Several other alternative nucleic acid-like polymers have been studied that differ more substantially in structure from RNA. Melting temperatures of duplexes of these polymers differ significantly, and may be either higher or lower than for RNA duplexes [14]. This suggests that RNA may be optimized for the conditions in which replication was originally occurring [15,16]. Some alternative nucleic acids can form hybrid duplexes with RNA, and some cannot. For example, Schöning et al. [17] studied four pairs of complementary oligomers, each made with TNA (α-threofuranosyl nucleic acid), RNA, and DNA backbones, and measured the melting temperatures of the 24 possible hybrid duplexes. They found a general tendency to favour uniform duplexes, although the results are complex, and we will return to them in the discussion section.

Monomer selection also involves a choice between nucleotides that differ in bases but have the same backbone. It is clear that strands made from the standard ACGU ribonucleotide alphabet are effective as templates for strands of the same alphabet. Non-enzymatic, template-directed synthesis of RNA and DNA has been studied [18], particularly with regard to the fidelity of sequence replication. A good monomer alphabet for genetics should allow little mismatch pairing. The possibility of GU pairing in the ACGU alphabet of RNA, and the absence of this possibility in the ACGT alphabet of DNA, is a potential reason for the transfer of information from RNA to DNA during evolutionary history [18]. Furthermore, a number of other base pairs have been studied that can be used as an extension of the coding repertoire of RNA and DNA [19,20]. These bases are 'orthogonal' to the standard bases, i.e., there should be little or no pairing between them and the standard bases. This means that sequences made of either the standard alphabet or the alternative alphabet would be templates for complementary sequences in the same alphabet, which is our proposition 1. However alternative pairs that are compatible with the RNA or DNA helix may not satisfy proposition 2, because the mixed sequences may still be good templates in this case. Studies of template-directed synthesis [21] have shown that the rate of nucleotide addition after a mismatch is much lower than after a match. This illustrates a close parallel between the monomer selection problem and the chirality problem, where the addition of a monomer of the wrong chirality can slow down the growth of an oligomer [12]. Although these effects cause slower average growth of oligomers, they actually increase the degree of uniformity of the longer oligomers. Mismatch inhibition increases the overall fidelity of RNA replication [21], and we expect that chiral inhibition would increase the enantiomeric excess of longer oligomers for the same reason.

Early studies on oligomerization of activated guanosine nucleotides using poly(C) templates [22,23] found that G-G bond formation is regioselective with a majority of 3′-5′ bonds, i.e., the template favours formation of the same bond type in the complementary strand (proposition 1 above). It was additionally shown [24] that the A structure of the nucleic acid helix is important in bringing the nucleotides together in the correct conformation for 3′-5′ bond formation. Furthermore, there have been several experimental studies that look at the effects of mixing 2′-5′ and 3′-5′ bonds. When regular 3′-5′ strands are used as templates, the template-directed synthesis of the complementary strand is regioselective for 3′-5′ bonds [25]. The duplex melting temperature is found to decrease systematically by around 15 °C as 2′-5′ bonds are added to the regular 3′-5′ structure [26,27]. This also suggests that strands of uniform bond type should be better templates than mixed strands (proposition 2). If 2′-5′ bonds and 3′-5′ bonds are similar with respect to rates of formation and hydrolysis, then the possibility arises that that strands of uniform bond type can be selected by symmetry breaking, as we discuss in this paper. However, it is already known that 2′-5′ bonds may be much less stable to hydrolysis than 3′-5′ bonds [28], hence, there is no true symmetry in this problem, and it is possible for one bond type to be more frequent than the other without there being a symmetry-breaking transition. A recent experimental study has shown that iterative degradation and repair of bonds gradually converts 2′-5′ bonds to 3′-5′ bonds [29] because bond formation in the context of the existing helix is regioselective. This effect is included in the model of backbone regioselectivity given here.

The aim of this paper is to present a simple computational model that is able to treat the problems of chirality, monomer selection, and backbone regioselectivity in the same way. We propose that all three types of order emerge when the rate of non-enzymatic template-directed synthesis is high, due to the fact that uniform strands are templates for their own kind, and that uniform strands are better templates than mixed strands. This mechanism is likely to occur at the level of oligomers, and does not require the synthesis of strands that are long enough to function as specific ribozymes.

2. Materials and Methods

2.1. Reaction Schemes

The reaction schemes used in this paper are summarized in Figure 1. For the chirality problem, we begin with equal concentrations of monomers of the two enantiomers *D* and *L*. These monomers can react to form oligomers of all possible sequences. Uncatalyzed polymerization is controlled by a reaction rate constant k_{pol}. Let *i* and *j* represent any two oligomers or single monomers, and let *ij* represent the longer oligomer formed by joining these two (see example in Figure 1i). Let C_i, C_j, and C_{ij} be the concentrations of these three oligomers. The rate of formation of *ij* by uncatalyzed polymerization is $k_{pol}C_iC_j$. The reverse hydrolysis reaction occurs at rate $k_{hyd}C_{ij}$. Hydrolysis of *ij* can also occur at any other point in the sequence. For simplicity, we suppose that the same rate constant k_{pol} applies for joining any two oligomers and that the same rate constant k_{hyd} applies for hydrolysis at any point in any sequence.

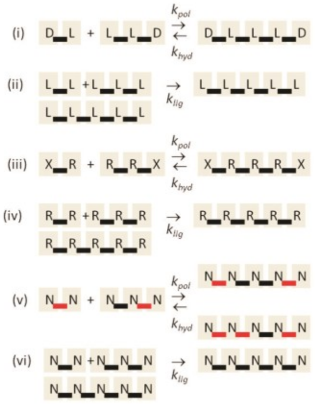


Figure 1. Reaction schemes for synthesis of RNA oligomers, involving uncatalyzed polymerization, k_{pol}; hydrolysis, k_{hyd}; and template-directed ligation, k_{lig}. (**i**,**ii**) the chirality problem; (**iii**,**iv**) the monomer selection problem; (**v**,**vi**) the backbone regioselectivity problem.

In addition to polymerization and hydrolysis, we suppose that uniform strands can act as templates for template-directed ligation of two shorter oligomers of the same kind. Thus, in Figure 1ii, an *LLLLL* pentamer can be a template for the ligation of *LL* and *LLL* to make another *LLLLL*. The same also applies for uniform *D* sequences. In order for the reaction to be template-directed, we suppose that both short oligomers *i* and *j* must be fully uniform (only *D* or only *L* this case) and that the template must contain a uniform sequence at least as long as the concatenated *ij* sequence. Thus, the sequence *DLLLLLLDDD* is also a template for joining *LL* and *LLL*. For simplicity, we suppose that all strands that possess a uniform sequence of the minimum required length are equally good templates. Template-directed ligation is controlled by rate constant k_{lig}. The rate of ligation of *i* and *j* by templating is $k_{lig}C_iC_jC_{ij}^{temp}$, where C_{ij}^{temp} is the total concentration of all strands that are templates for formation of *ij*, using the rules above. The templating reaction is treated as a single-step reaction. We do not keep track of duplex states and we assume that *ij* separates immediately from the template once formed. It would be possible to add separate steps of duplex formation and separation to the model, but initially we are aiming for the simplest possible model.

A final feature of the chirality model is that it must be possible to convert monomers from one enantiomer to the other in some way. The model is initialized with equal concentrations of *L* and *D*. If no interconversion is possible, then no chiral symmetry breaking can occur. Direct racemization of sugars and nucleotides is extremely slow, but breakdown of a monomer to achiral precursors can also occur, as well as synthesis of monomers of both enantiomers from the precursors. We do not include precursors in this model. We assume that interconversion of L to D and D to L can occur by an effective single-step reaction at rates $k_{int}C_L$ and $k_{int}C_D$, respectively. Only single monomers can be interconverted; monomers contained in oligomers cannot be changed until they are hydrolyzed back to single monomers. Including interconversion in this simple way means that the total concentration of nucleotides (including those in oligomers) is fixed, but the concentrations of D and L can each increase or decrease.

For the monomer selection problem, we consider two monomers R and X, where R represents a ribonucleotide, and X represents an alternative nucleotide. The simplest case for the monomer selection problem is identical to the chirality problem, except that the labels D and L are replaced by R and X (see Figure 1iii,iv). We are assuming that formation of mixed RX sequences is possible by uncatalyzed polymerization, but that only uniform R or X sequences can be joined by templating. In the chirality case, there is complete symmetry initially between D and L; hence, the reaction rate constants should be equal for oligomers of D and L. This is not necessarily true for R and X because they are chemically different. Therefore, in the results section, we will consider cases in which one or other type of monomer is a better template than the other. For simplicity, we allow the interconversion between R and X to occur in a single effective step, although breakdown and re-synthesis of could be included in a more complex model. The interconversion rate constants in the two directions need not be equal in this case.

The backbone regioselectivity problem requires a few small modifications to the model. There is only one kind of monomer, which we denote N for nucleotide, and two kinds of bonds. We have coloured the bonds red and black in Figure 1v, representing 2′-5′ and 3′-5′ bonds, respectively. Joining two oligomers can occur via either type of bond formation, leading to two longer oligomers with different bond sequences. In the simplest case, we suppose the rate constants k_{pol} and k_{hyd} are the same for the two bond types, but this can be relaxed. The templating reaction occurs only if the two short oligomers are of uniform bond type, and only formation of the bond of the same type is catalyzed by the template (Figure 1vi). To be a template, the templating sequence must contain a uniform sequence of bonds at least as long as the sequence of bonds that is formed by the ligation reaction. Thus, in Figure 1vi, the pentamer formed has four black bonds, so the template must have at least four black bonds. There is no equivalent of the interconversion reaction in the regioselectivity problem because there is only one type of monomer. When a dimer with one kind of bond is hydrolyzed, the two monomers can be rejoined with either bond type, so the number of bonds of each type is not fixed.

2.2. Computational Methods

In this paper, we study the three theoretical models described above using two kinds of computer simulations: reaction kinetics and Monte Carlo. In the reaction kinetics method, we keep track of the concentration of each monomer and oligomer species as a function of time, and solve the deterministic reaction kinetics equations by iterating forwards in small time steps of length δt. The change in concentration of oligomer ij due to formation from i and j and hydrolysis back to i and j is:

$$\delta C_{ij} = \left(k_{pol}C_iC_j + k_{lig}C_iC_jC_{ij}^{temp} - k_{hyd}C_{ij}\right)\delta t \quad (1)$$

There is an equal and opposite change in concentration of i and j from these reactions:

$$\delta C_i = \delta C_j = -\delta C_{ij} \quad (2)$$

We consider each possible pair i and j and sum up the total concentration changes of all molecules from all possible reactions. The concentration of the template in Equation (1) is different for each i and j. The template concentration does not change in the reaction for which it is a template, but templates are made by equivalent reactions involving other oligomers. Additionally, there is a change of the two monomer concentrations due to the interchange reaction described above:

$$\delta C_L = -\delta C_D = k_{\text{int}}(C_D - C_L)\delta t \tag{3}$$

A slight modification is required for Equation (1) for the case of backbone regioselectivity. Whereas there is only one way to link two oligomers for the chirality and monomer selection problems, there are two ways to do this for the backbone regioselectivity problem. We therefore change the first term of Equation (1) so that each of the two reactions occurs at rate $\frac{1}{2}k_{pol}C_iC_j$. The total rate of linking i and j is still $k_{pol}C_iC_j$, which means that the distribution of lengths of oligomers is still the same as the other two models when there is no ligation term (See Appendix A). On the other hand, only one of the two kinds of bonds is catalyzed by the template-directed reaction. Therefore, the rate of this term remains a $k_{lig}C_iC_jC_{ij}^{temp}$, as with the other two models.

In the reaction kinetics method, we deal with all possible reactions deterministically; therefore, it is necessary to specify a maximum possible length, l_{max}, of oligomers that can be formed, in order to keep the number of possible sequences finite. If i and j have a total length greater than l_{max}, this polymeration reaction is not permitted to occur. For the results presented here, we set l_{max} = 6. Thus the reaction system consists of 2^6 hexamers, plus all the oligomers shorter than 6. We specified the total monomer concentration as C_{tot} = 10 monomers per unit volume (arbitrary units). The rate constants for polymerization and hydrolysis were fixed at k_{pol} = 1 per unit time per unit concentration and k_{hyd} = 1 per unit time. The behaviour of the model was studied for different values of $k_{lig.}$

In the Monte Carlo method, we begin with a finite total number of nucleotides, N_{tot}, and follow individual joining and hydrolysis reactions, keeping track of the sequences of all the oligomers that are formed. In this case, the computation is finite because of the finite number of molecules; therefore, it is not necessary to specify a maximum strand length. The oligomers are assumed to be reacting in a well-mixed solution of volume V. Reactions occur randomly with probabilities such that the expected rates are the same as the reaction kinetics method. We set N_{tot} = 100,000 and V = 10,000; hence C_{tot} = 10, as in the reaction kinetics method.

The Monte Carlo program also proceeds in small time steps of length δt. In each time step, each current strand is given a possibility of hydrolysis. For a strand of length n nucleotides, there are $n - 1$ bonds. Hydrolysis occurs with probability

$$p_{hyd} = (n-1)k_{hyd}\delta t \tag{4}$$

and one of the $n - 1$ bonds is chosen at random.

Let $N_{strands}$ be the number of strands in the simulation (counting single monomers as a strand). $N_{strands}$ varies during the simulation and must be calculated at each time step, whereas the total number of nucleotides, N_{tot}, is fixed. In each time step, each current strand is given a possibility to participate as the 'left' oligomer of a polymerization reaction (i). One of the $N_{strands} - 1$ other strands is chosen to be the 'right' oligomer (j). The probability of the polymerization reaction occurring is

$$p_{pol} = \frac{(N_{strands}-1)}{V}k_{pol}\delta t \tag{5}$$

If the reaction occurs, the new strand ij is formed by linking i and j.

The i + j reaction can also occur via a template-directed reaction. There are $N_{strands} - 2$ other strands that could be the template. If all strands were templates, the probability of this reaction occurring would be:

$$p_{lig} = \frac{(N_{strands}-1)(N_{strands}-2)}{V^2}k_{lig}\delta t \tag{6}$$

For each strand i, a second strand j is chosen at random, and a third strand k is chosen to be a potential template. If k is a template for the i + j reaction (as defined by the rules in Section 2.1), then the new strand ij is formed by the ligation reaction with probability p_{lig}. If k is not a template, this reaction does not occur. At the beginning of the simulation, when all nucleotides are single monomers, $N_{strands}/V$ can be larger than 1. Also, we need to consider cases where k_{lig} is much larger than 1 in order for the symmetry breaking to occur. Therefore, we begin with a very small δt (5×10^{-5}) to ensure that p_{lig} is a probability that is less than 1. As the simulation proceeds, the number of strands decreases, and it is possible to make δt larger, to increase the efficiency of the program. We adjust δt so that p_{lig} remains close to 5%, or so that p_{pol} is close to 5% in cases where $p_{pol} > p_{lig}$.

The interchange reaction is straightforward in the Monte Carlo simulation: each single monomer has a probability $k_{int}\delta t$ of switching from D to L or vice versa.

3. Results

3.1. Chirality

The parameters in the results shown here are $C_{tot} = 10$, $k_{pol} = 1$, $k_{hyd} = 1$, $k_{int} = 1$, and $l_{max} = 6$, as described above. Figure 2 shows the faction of nucleotides ϕ_D, ϕ_L, and ϕ_M at equilibrium. The fraction of nucleotides contained in uniform D oligomers is

$$\phi_D = \frac{1}{C_{tot}} \sum_i n_i C_i \tag{7}$$

where n_i is the number of nucleotides in sequence i, C_i is the concentration of sequence i, and the sum is over all sequences that are uniform-D oligomers (dimers and longer). Similarly, we can define the fraction of nucleotides in uniform-L oligomers, ϕ_L, and in mixed oligomers, ϕ_M. In the latter case, the sum is over all sequences that contain at least one D and one L.

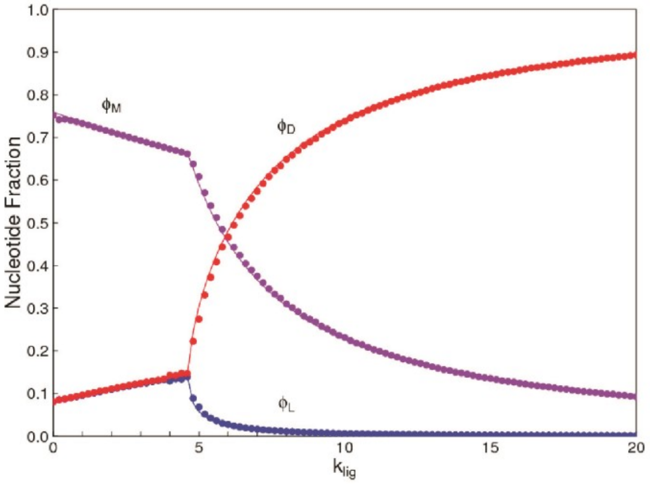


Figure 2. Fractions of nucleotides ϕ_D, ϕ_L, and ϕ_M, in uniform D and L strands and in mixed strands. Smooth lines are calculated from reaction kinetics. Symbols are from Monte Carlo simulations. Below the phase transition, D and L are equal. Above the phase transition, there is a chiral bias in favour of D. An equivalent solution with the bias in the other direction is also possible.

When $k_{lig} = 0$, a majority of nucleotides end up in mixed oligomers. The fraction of nucleotides in uniform D and L oligomers are both low, and equal to one another (Figure 2). There is a symmetry-breaking phase transition close to $k_{lig} = 4.6$ in this example. Above this point, the fraction of monomers in one kind of uniform oligomers is much higher than in the other kind (in this case $\phi_D > \phi_L$), and the fraction in mixed oligomers also decreases rapidly.

Symmetry breaking occurs because the symmetric mixture is unstable when k_{lig} is large. The simulations demonstrate this by beginning with a small bias towards D monomers—$C_D = 0.505C_{tot}$ and $C_L = 0.495C_{tot}$. When k_{lig} is low, the bias disappears, and the equilibrium solution is perfectly symmetric. When k_{lig} is high, the bias increases, and the equilibrium has a large excess of D over L. There is an equivalent equal-and-opposite solution with an excess of L over D, which is reached if we began with a slight bias towards L. In any real case, there is a finite volume with a finite number of molecules, and small fluctuations are bound to create a slight asymmetry one way or the other. This asymmetry is magnified to a large excess of either L or D because the reaction system is inherently unstable. The prediction is that the system will become biased towards L or D with equal probability if it begins with equal concentrations of the two. It is not necessary to begin with any source of initial asymmetry. On the other hand, if the reaction system begins with a slight bias one way for other reasons (as is apparently observed for organic molecules in meteorites [30]), the autocatalytic reaction amplifies the asymmetry in the direction that is previously specified (theoretical examples of this are given in [10]).

As a further check that our simulations have reached an equilibrium, we also started simulations where $C_D = C_{tot}$ and $C_L = 0$. These simulations were run for a long time with no interchange reaction permitted, so that a distribution of uniform D strands was created. The interchange reaction was then turned on, and the simulation was continued to equilibrium. For low k_{lig}, the same symmetric solution was reached as before, and for high k_{lig}, the same chirally biased solution was reached as before. This shows that the biased solution can be reached either by amplifying a very small initial bias, or by allowing a system with a 100% chiral bias to relax to the equilibrium state.

We also used the Monte Carlo simulation to study the same case. The points in Figure 2 are from Monte Carlo, and the lines are from the reaction kinetics. These two methods give identical results, which is an important check on validity of both methods. In the Monte Carlo case, we calculate time averages of quantities once equilibrium has been reached. We can begin with exactly equal numbers of D and L monomers ($N_{tot}/2$ = 50,000 of each type). It is not necessary to begin with a small bias in one direction, because the interchange reactions allow fluctuations of these numbers to arise. If k_{lig} is above the phase transition, the fluctuations are amplified and the system shifts to a state with either high D or high L with equal probability. For the purposes of comparison with the reaction kinetics method in Figure 2, we plotted the higher concentration as D and the lower one as L in each case, but the labeling of D and L is arbitrary because there is a true symmetry in this problem and the symmetry breaking can go either way.

The reactions for hydrolysis, polymerization, and template-directed ligation are unimolecular, bimolecular, and trimolecular, respectively. Thus, the relative rates of these reactions depend on the concentration. The importance of templating is increased when the concentration is high. Figure 3 shows the nuclotide fractions as a function of total monomer concentration, C_{tot}, with $k_{lig} = 10$, $k_{pol} = 1$, $k_{hyd} = 1$, and $k_{int} = 1$. At low concentration, D and L are equal, and chiral symmetry breaking occurs as C_{tot} is increased. If this is repeated for different k_{lig} rates, the concentration at which the symmetry breaking occurs is roughly inversely proportional to k_{lig}.

Figures 2 and 3 were done with maximum strand length $l_{max} = 6$, in order to allow comparison of results between the two simulation methods. Using the Monte Carlo program, we repeated the simulation with $l_{max} = 10$, and with no maximum length restriction ($l_{max} = \infty$). The results are similar to the case with $l_{max} = 6$, although the phase transition occurs at a higher value of k_{lig}. This is shown in Figure 4 in terms of enantiomeric excess, *ee*. The *ee* is defined as the difference in concentrations of the two enantiomers relative to the sum: $ee = (D - L)/(D + L)$. Figure 4 shows the *ee* for nucleotides incorporated into oligomers (length 2 or above). There are two equal and opposite solutions when k_{lig} is high. In our model, there is direct interchange between single monomers of D and L. Therefore, the concentrations of the two single monomers remain equal. For this reason, the *ee* of single monomers is always zero, even above the phase transition, whereas the *ee* of the nucleotides in oligomers becomes non-zero above the phase transition. The main point of Figure 4 is that the bifurcation point shifts to slightly higher k_{lig} as the l_{max} is shifted from 6 to 10, to ∞, but there is no qualitative change. Imposing the maximum length restriction simplifies the simulations slightly, but states of biased chirality arise whether or not there is a maximum length. The equilibrium distribution of lengths of oligomers for these cases is discussed in the Appendix of this paper.

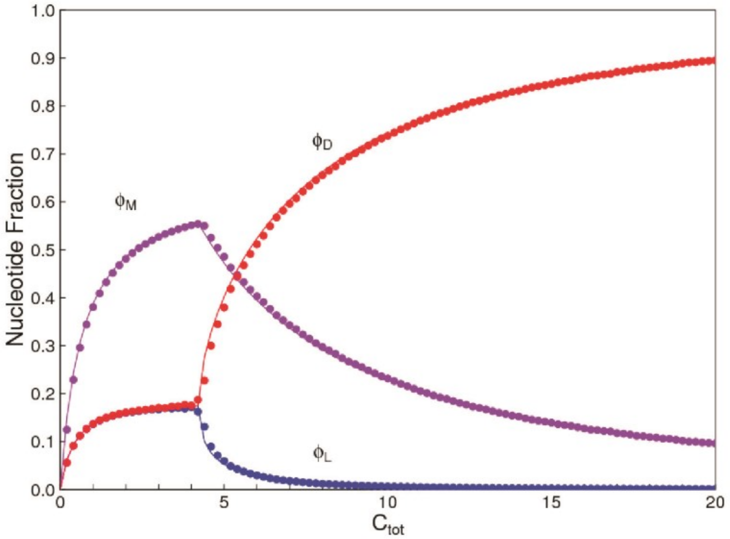


Figure 3. Fraction of nucleotides ϕ_D, ϕ_L, and ϕ_M, in uniform D and L strands and in mixed strands as a function of concentration. Smooth lines are calculated from reaction kinetics. Symbols are from Monte Carlo simulations.

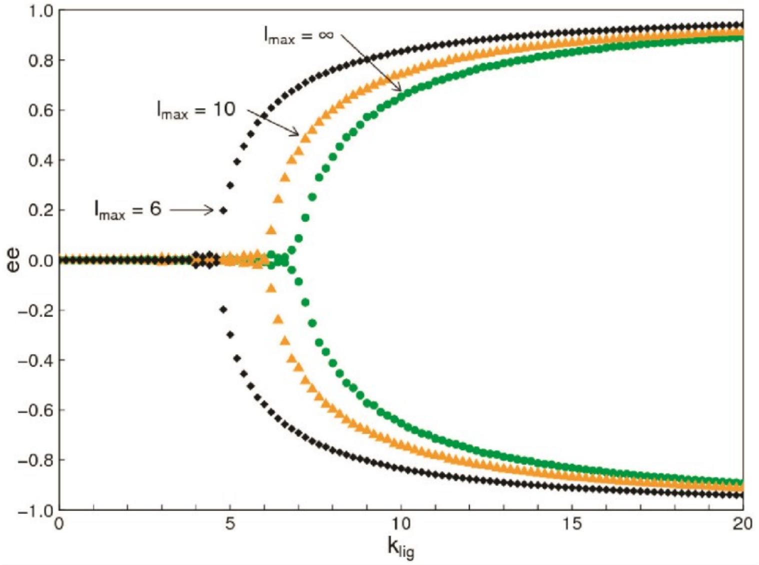


Figure 4. Enantiomeric excess (*ee*) of nucleotides contained in oligomers. The cases of $l_{max} = 6$, 10, and ∞ are very similar, but the chiral symmetry breaking occurs at slightly higher k_{lig} when l_{max} is larger. The *ee* is zero below the phase transition, and can either be positive or negative once the symmetry is broken.

For the results shown above, we considered D and L nucleotides without specifying the base sequence. In order to deal more explicitly with complementary sequence pairing, we also considered a case in which there are eight kinds of nucleotides: A, U, G, and C bases, each of D and L form. Polymerization and hydrolysis reactions occurred at equal rates independently of the base and the chirality. For template-directed reactions to occur, both the oligomers i and j had to be of uniform chirality, and the same chirality as each other, but they could have any base sequence. A third sequence could be a template if it contained a uniform sequence of the same chirality as i and j, with a base sequence that was complementary to the ij sequence that is formed by ligation. Only AU and GC pairs were permitted when checking for complementarity. The results of this case, shown in Figure 5, are very similar to Figure 2, except that the phase transition occurs at k_{lig} close to 40 rather than 4.6 in Figure 2. The main effect of introducing the detailed requirement of complementary pairing is to reduce the concentration of strands that can be catalysts for any two oligomers being ligated. As a result, higher values of k_{lig} are required in order for the symmetry-breaking effects of the template-directed reaction to become apparent. However, the essential symmetry-breaking effect is the same as in the simpler model.

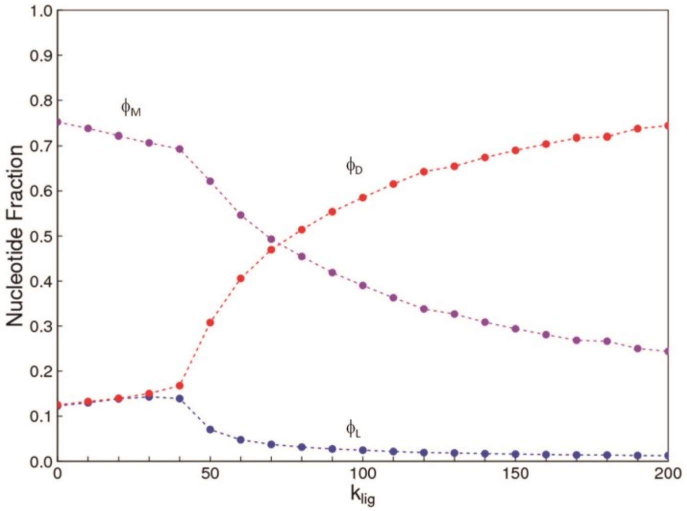


Figure 5. Fractions of nucleotides ϕ_D, ϕ_L, and ϕ_M, in uniform D and L strands and in mixed strands for the case where A, C, G, and U bases of both enantiomers are explicitly included and complementary pairing is enforced between templates and oligomers being ligated.

3.2. Monomer Selection

In this problem, R represents a ribonucleotide and X represents an alternative nucleotide. If the rate constants for the two nucleotides are the same, the problem is identical to the chirality problem. The symmetry between R and X is broken as k_{lig} is increased. In Figure 6, we also considered the case where the ligation rates for the two types of polymer are slightly different: $k_{ligX} = 0.9k_{ligR}$. This means that R has an advantage over X because the ligation rate of uniform X sequences is less than that for uniform R sequences. It can be seen that the phase transition is 'rounded out' when there is no exact symmetry between R and X. When k_{lig} is far above or far below the phase transition, the concentrations of R and X are close to that for the symmetric case.

Figure 6 also shows a second solution at high k_{ligR} where the symmetry is broken in the opposite direction, i.e., X dominates even though it has a lower ligation rate than R. This solution only exists above a minimum ligation rate (close to 6 in the figure), whereas the solution where R dominates exists for all values of k_{ligR}. The two solutions are obtained by beginning with monomers of all R or all X and allowing a steady state to be reached before the interchange reaction is turned on. The interchange reaction is then turned on and the simulation is then continued to equilibrium. For ligation rates where there is only one solution, the same solution is reached starting from both extremes. When there are two solutions, these are reached from the two different starting points. It should be remembered that, even in the simplest case in Figure 2, there are two solutions when k_{lig} is above the transition point. As these are equal and opposite, only one set of curves appears in Figure 2, but the two opposite solutions are visible in the *ee* graph (Figure 4). In Figure 6, the two solutions are not equal and opposite; therefore, two sets of curves are visible.

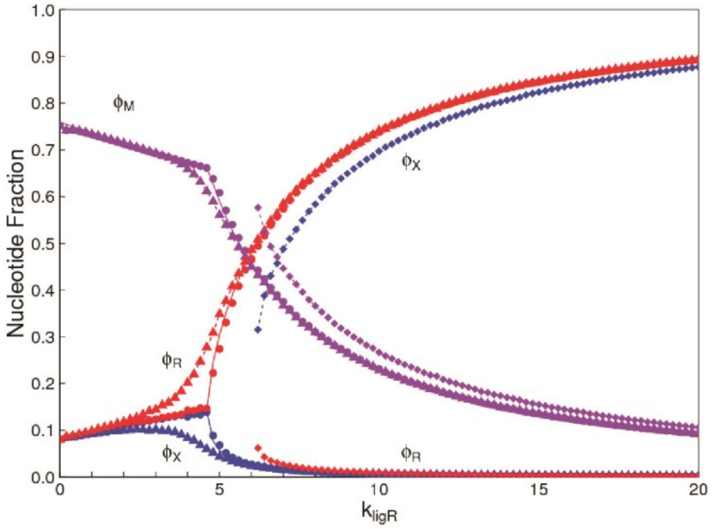


Figure 6. Comparison of the monomer selection problem with the chirality problem. Red points: uniform *R* strands; Blue points: uniform *X* strands; Violet points: mixed strands. Circles show the solution in the case where there is perfect symmetry between *R* and *X* (This is identical to Figure 2). Triangles and squares show the two possible solutions in the case where $k_{ligX} = 0.9k_{ligR}$.

A second way in which *R* and *X* can differ is by their frequency in the monomer mixture prior to polymerization. If the rate constants for interchange from *X* to *R* and from *R* to *X* are k_{intXR} and k_{intRX}, respectively, then the frequencies of monomers at equilibrium under the interchange reaction, in absence of any polymerization will be unequal:

$$\frac{C_X}{C_R} = \frac{k_{intRX}}{k_{intXR}} \tag{8}$$

Figure 7 shows the case where there is a two-fold advantage to *X* in terms of equilibrium monomer frequencies ($C_R = \frac{1}{2}C_X = \frac{1}{3}C_{tot}$), but a two-fold advantage to *R* in terms of ligation rates $k_{ligX} = 0.5k_{ligR}$. In this case, the solution where *X* dominates exists over the whole range of ligation rates, and a second solution where *R* dominates exists only for high ligation rates. Thus, the frequency of the monomers in the reaction mixture is a major factor that determines which type of monomer dominates when templating occurs.

The number of monomer types in the mixture need not be limited to only two. In Figure 8, we consider the case where there are two alternatives, X_1 and X_2, in addition to *R*. We assume that all the interchange rates are equal, so all three types have equal average concentration in the monomer mixture ($C_R = C_{X_1} = C_{X_2} = \frac{1}{3}C_{tot}$), and we give *R* a small advantage in the ligation rate ($k_{ligX_1} = k_{ligX_2} = 0.9k_{ligR}$). In this case, it is the solution where *R* dominates that is possible over the full range of ligation rates, and the two solutions where the other monomer types dominate exist only at high ligation rates.

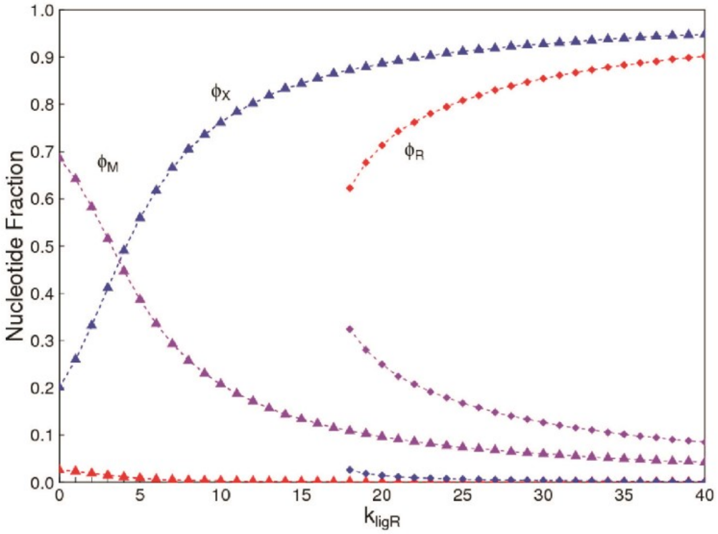


Figure 7. Two possible solutions in the monomer selection problem where $C_X = 2C_R$ and $k_{ligX} = 0.5k_{ligR}$. Red points: uniform R strands; Blue points: uniform X strands; Violet points: mixed strands. Triangles: solution where X dominates; Squares: solution where R dominates.

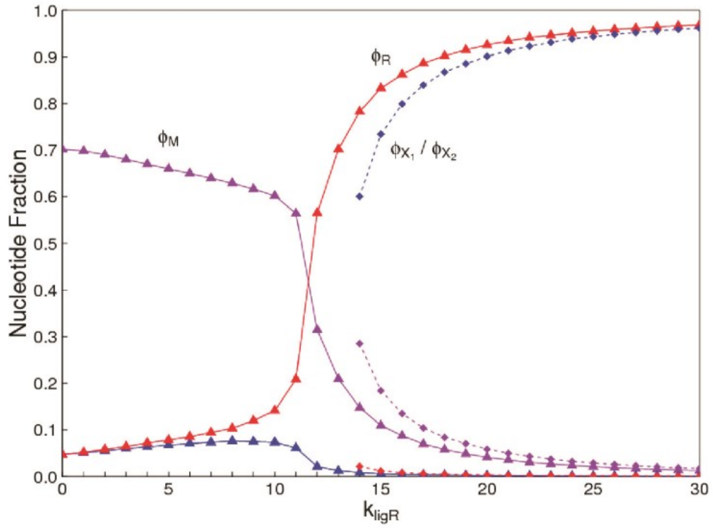


Figure 8. Possible solutions in the case where there are three kinds of nucleotides, R, X_1, and X_2, with equal monomer frequency and $k_{ligX_1} = k_{ligX_2} = 0.9k_{ligR}$. Triangles: solution where R dominates; Squares: two equivalent solutions where either X_1 or X_2 dominates.

3.3. Regioselectivity

If we assume a complete symmetry between 3′-5′ bonds and 2′-5′ bonds, then the regioselectivity model behaves similarly to the chirality model. There is a symmetry-breaking phase transition qualitatively similar to Figure 2, in which one or other of the two bond types dominates when k_{lig} is high. We do not show this case, because there are many ways in which these two bond types might differ in practice, and there is no reason to suppose perfect symmetry. In Figure 9, we suppose the major difference in the two bond types is that the uniform 3′-5′ oligomers have a higher ligation rate than the uniform 2′-5′ oligomers ($k_{lig2} = 0.9k_{lig3}$). It is seen that there is a solution where 3′-5′ strands dominate that is possible over the full range of ligation rates, and a solution where 2′-5′ strands dominate that is possible only at high ligation rates. The other rate parameters, k_{pol} and k_{hyd}, are assumed to be equal in this example, but are likely to be different in an experimental system. Which of the bond types dominates will be a function of the rates of all these parameters for both bond types.

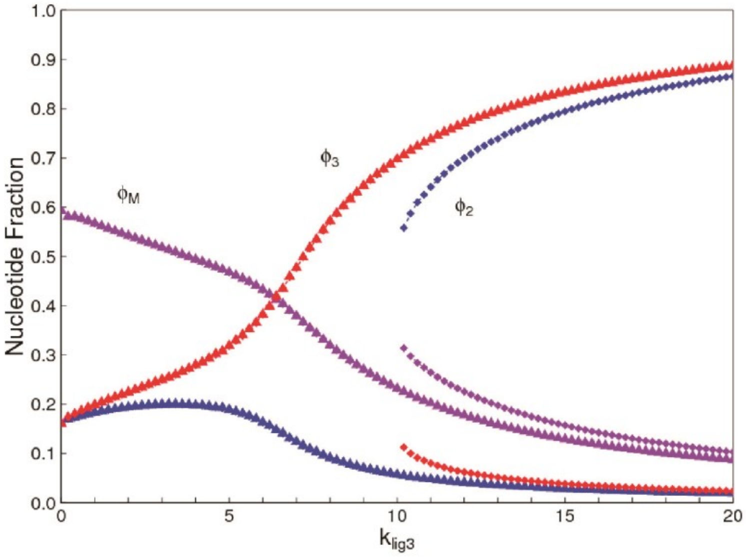


Figure 9. Concentration of nucleotides in uniform 3′-5′ and 2′-5′ strands and in mixed strands for the case where $k_{lig2} = 0.9k_{lig3}$. Red points: uniform 3′-5′ strands; Blue points: uniform 2′-5′ strands; Violet points: mixed strands. Triangles show the solution where 3′-5′ strands dominate and squares show the solution where 2′-5′ strands dominate.

4. Discussion

In the introduction, we reviewed the experimental evidence that oligomers with uniform chirality, monomer composition, or bond type are effective templates for the growth of complementary strands with the same chirality, monomer composition, or bond type, and that oligomers that are mixed in any of these properties are less effective templates than uniform oligomers. The computational models studied in this paper incorporate these features, and show that, when template-directed ligation is fast compared to random polymerization without a template, we expect that uniform biopolymers will emerge, and that they are homochiral, use a restricted monomer set, and have a highly regioselective backbone. We thus argue that these essential properties of biological nucleic acids have emerged as a result of the importance of template-directed reactions in the early stages of evolution. Although models for the emergence of homochirality have been widely studied previously, the link between chirality, monomer selection, and regioselectivity has not been emphasized in the past. The present theory therefore contributes by showing in a simple way why these three problems are very similar. These results lead us to expect that RNA with the three uniform properties seen in biology could emerge from a prebiotic chemical mixture *before* the origin of life, in the sense that the proposed mechanism requires only the existence of short oligomers that can undergo template-directed replication, but it does not require the existence of specific sequences that would act as catalysts (ribozymes).

The mechanism we discuss here depends on the ability of nucleic acids to be templates; hence, it applies to RNA and polymers with similar structures, but not to other kinds of biomolecules that are also chiral in today's organisms but cannot be templates. The ability of nucleic acids to be templates is, of course, a prime argument in favour of some kind of RNA World scenario for the origin of life. In an origins scenario where RNA replication arises early, and where ribozymes are the initial biological catalysts, it is possible to explain the transfer of chirality from RNA to other biomolecules by the fact that chiral RNAs catalyze the synthesis of the other molecules. In particular, if protein synthesis depended on chiral tRNAs and rRNAs, then amino acids of the appropriate chirality could be charged onto tRNAs, and protein sequences of uniform chirality could be synthesized. This would be true, even if the mixture of single amino acids were completely racemic, because amino acids of the wrong enantiomer would not be recognized by the ribozymes. On the other hand, if the amino acids were already chiral for some other reason at the time of the invention of protein synthesis, the early ribosome would simply make use of the existing supply of chiral amino acids.

Related to this is the possibility that the transfer of chirality was in the other direction: from amino acids to sugars and nucleotides. Several studies have shown [31,32] that chiral amino acids can catalyze the synthesis of chiral glyceraldehye and other sugars. There is also evidence for a moderate enantiomeric excess in amino acids found in meteorites [30], which is presumed to originate by an abiotic mechanism outside the Earth, such as chirally biased degradation under the influence of circularly polarized light. We have previously considered the way that a small extraterrestial chiral bias can contribute to scenarios involving the emergence of homochirality [10]. Understanding the direction of chiral transfer will require a more detailed picture of the full network of reactions involving biomolecules of all kinds at the time of the origin of life. The template-based mechanism presented here is not in conflict with other studies showing transfer of chirality from one molecule to another, whichever the direction of transfer, and it is quite possible that more than one mechanism jointly contributed to the creation of a biology in which several kinds of chiral biopolymer are inter-dependent. Nevertheless, we maintain that templating must have been an essential part of this process. Life requires the copying of sequence information that occurs via the template-directed synthesis of complementary strands of nucleic acids. A mixture of small molecules (whether chirally biased or not) that did not possess a mechanism of sequence replication would not yet constitute life, in our view. Templating is a straightforward mechanism that 'purifies' biopolymers at the level of chirality, chemical composition, and regioselectivity, as we have shown. The point of this paper is to show that templating alone can provide a mechanism for the emergence of all these properties, even in absence of other sources of asymmetry. If other sources, or other mechanisms, exist, then templating magnifies these effects to a much greater extent.

The mechanism for the emergence of RNA discussed in the present paper is an example of what we have termed *chemical evolution* [33]. The main distinctions between chemical evolution and biological evolution, as defined in [33], are (1) that, for chemical evolution, the major part of the sequence diversity on which natural selection acts is provided by random polymer synthesis, rather than by mutations occurring during replication of previously existing sequences; and (2) that selection is acting on physicochemical properties possessed by all short oligomers (like ligation and hydrolysis rates) rather than on encoded function of long sequences (such as ribozymes) that are only possessed by a small number of sequences in a large sequence space. If RNA can emerge directly from a prebiotic mixture by chemical evolution, as envisaged here and previously [15,16], then it is not necessary to consider more complex pathways requiring the evolution of biological function in some pre-RNA polymer and the subsequent transfer of sequences and function to RNA, as has sometimes been proposed [34].

The monomer selection problem is probably the most complex and least understood of the three problems considered in this paper. It is clear that the real world is much more complicated than the simple model we have studied. For example, hybrid duplexes of RNA, DNA, and TNA were studied experimentally [17], and it was shown that the details of the nucleotide sequences and backbone structures significantly affect the melting temperatures (T_m) of helices, and presumably also affect the rates of polymerization, hydrolysis, and ligation. Table 2 of [17] illustrates how complex this problem is. There are 14 cases where the hybrid duplex has a lower T_m than the uniform duplexes of both polymers, 8 cases where the hybrid has a T_m in between the two uniform duplexes, and 2 cases where the hybrid has a T_m higher than both the uniform duplexes. The latter two cases are surprising, but the effect only works one way round: if the sequences are switched, the hybrid has a lower T_m than both uniform duplexes. Although the thermodynamic details are rather complex, there is a general tendency to favour uniform duplexes, which is the essential point that we assume in our model. The case where both uniform duplexes are good templates and the hybrids are weak corresponds to the case studied here. There is an approximate symmetry between the two uniform polymers that is broken when the template-directed reaction is fast. If the hybrid is intermediate between the two uniform duplexes, then there will be a straightforward selection for the better of the two uniform systems. The monomer selection problem does not require symmetry breaking in that case.

We hope to continue this work to consider important details not included in the present paper, including differences in thermodynamic properties and rate constants between different base sequences with the same backbone, and adding separate steps for helix formation and melting, rather than combining them into a single effective ligation step. The simple model presented here is sufficient to demonstrate the similarity between the problems of chirality, monomer selection, and regioselectivy, and to establish the relevance of symmetry breaking transitions in all three cases.

Acknowledgments: This work is supported by a Discovery grant (1244111) from the Natural Sciences and Engineering Research Council of Canada.

Author Contributions: P.G.H. and A.S.T. conceived and designed the experiments; A.S.T. and K.S. performed the simulations and analyzed the data; P.G.H. wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Here we consider the distribution of lengths of oligomers and the effect of template-directed reactions on this distribution. In most models of polymerization in which rates of linking and breaking monomers are independent on the lengths of the oligomers, e.g., [35], the equilibrium distribution of lengths is exponential, and it is often simple enough to be calculated exactly. If there is no templating (k_{lig} = 0), the model in this paper is almost the same as that discussed in [35], and the length distribution can be calculated in the same way. Let C_n be the total concentration of oligomers of length n. At equilibrium, the rates of polymerization and hydrolysis must balance for each pair of lengths n and m:

$$k_{pol}C_nC_m = k_{hyd}C_{n+m} \tag{A1}$$

This means that the solution is of the form

$$C_n = C_1(KC_1)^{n-1} \tag{A2}$$

where C_1 is the equilibrium concentration of single monomers, and $K = k_{pol}/k_{hyd}$. The total nucleotide concentration C_{tot} is fixed; hence

$$\sum_{n=1}^{l\max} nC_n = C_{tot} \tag{A3}$$

Substituting (A2) into (A3), we obtain the following equation, from which the solution for C_1 can be found numerically.

$$\frac{C_1\left(1 + l_{\max}(KC_1)^{l\max+1} - (l_{\max}+1)(KC_1)^{l\max}\right)}{(1-KC_1)^2} = C_{tot} \tag{A4}$$

To illustrate the shape of the strand length distribution, it is useful to plot the fraction of nucleotides in strands of length n—this is $f_n = nC_n/C_{tot}$. Figure A1a shows the theoretical curve calculated from Equations (A2) to (A4), together with simulation results. The fact that these are equal confirms that our simulation results are accurate in the case we can calculate analytically. Figure A1b shows the length distribution for k_{lig} = 20, which is the largest rate considered in Figure 4. The addition of templating shifts the length distribution towards longer lengths. However, the length distribution always decays exponentially for long lengths, even when k_{lig} is high and when $l_{max} = \infty$. The shapes of the length distributions in similar models to ours that include templating have also been considered by other authors [36,37].

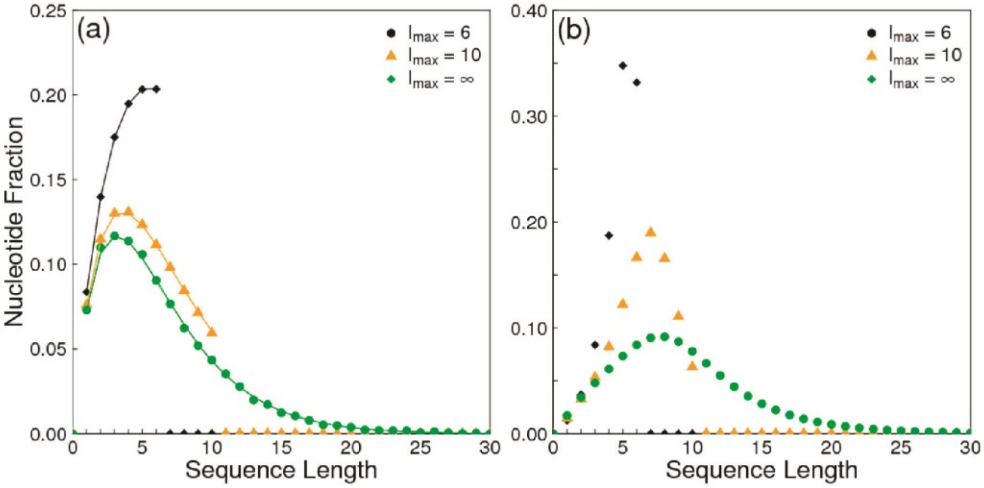


Figure A1. Fraction of nucleotides, f_n, in strands of length n. (**a**) No templating, k_{lig} = 0. Points = simulation. Lines = analytical calculation from Equations (A2) to (A4). (**b**) High templating rate, k_{lig} = 20. Points = simulations.

References

1. McKay, C.P. What is life—And how do we search for it in other worlds? *PLoS Biol.* **2004**, *2*. [CrossRef] [PubMed]
2. Meringer, M.; Cleaves, H.J.; Freeland, S.J. Beyond terrestrial biology: Charting the chemical universe of α-amino acid structures. *J. Chem. Inf. Model.* **2013**, *53*, 2851–2862. [CrossRef] [PubMed]
3. Benner, S.A.; Kim, H.J.; Kim, M.J.; Tocardo, A. Planetary organic chemistry and the origins of biomolecules. *Cold Spring Harb. Perspect. Biol.* **2010**, *2*. [CrossRef] [PubMed]
4. Cleaves, H.J.; Bada, J.L. The prebiotic chemistry of alternative nucleic acids. In *Genesis—In the Beginning*; Seckbach, J., Ed.; Springer: Berlin/Heidelberg, Germany, 2012; Volume 22, pp. 3–33.
5. Cafferty, B.J.; Hud, N.V. Was a pyrimidine-pyrimidine base pair the ancestor of Watson-Crick base pairs? Insights from a systematic approach to the origin of RNA. *Isr. J. Chem.* **2015**, *55*, 891–905. [CrossRef]
6. Plasson, R.; Kondepudi, D.K.; Bersini, H.; Commeyras, A.; Kouichi, A. Emergence of homochirality in far-from-equilibrium systems: Mechanisms and role in prebiotic chemistry. *Chirality* **2007**, *19*, 589–600. [CrossRef] [PubMed]
7. Frank, F.C. On spontaneous asymmetric synthesis. *Biochim. Biophys. Acta* **1953**, *11*, 459–463. [CrossRef]
8. Islas, J.R.; Lavabre, D.; Grevy, J.M.; Lamoneda, R.H.; Cabrera, H.R.; Micheau, J.C.; Buhse, T. Mirror-symmetry breaking in the Soai reaction: A kinetic understanding. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 13743–13748. [CrossRef] [PubMed]
9. Sandars, P.G.H. A toy model for the generation of homochirality during polymerization. *Orig. Life Evol. Biosph.* **2003**, *33*, 575–587. [CrossRef] [PubMed]

10. Wu, M.; Walker, S.I.; Higgs, P.G. Autocatalytic replication and homochirality in biopolymers: Is homochirality a requirement for life or a result of it? *Astrobiology* **2012**, *12*, 818–829. [CrossRef] [PubMed]
11. Bolli, M.; Micura, M.; Eschenmoser, A. Pyranolsyl-RNA: Chiroselective self-assembly of base sequences by ligative oligomerization of tetranucleotide-2′,3′-cyclophosphates (with a commentary concerning the origin of biomolecular homochirality). *Chem. Biol.* **1997**, *4*, 309–320. [CrossRef]
12. Joyce, G.F.; Visser, G.M.; van Boeckel, C.A.A.; van Boom, J.H.; Orgel, L.E.; van Westrenen, J. Chiral selection in poly(C)-directed synthesis of oligo(G). *Nature* **1984**, *310*, 602–604. [CrossRef] [PubMed]
13. Gavette, J.V.; Stoop, M.; Hud, N.V.; Krishnamurthy, R. RNA-DNA chimeras in the context of an RNA world transition to an RNA/DNA world. *Angew. Chem.* **2016**, *55*, 1–7. [CrossRef] [PubMed]
14. Eschenmoser, A. Chemical etiology of nucleic acid structure. *Science* **1999**, *284*, 2118–2124. [CrossRef] [PubMed]
15. Krishnamurthy, R. RNA as an emergent entity: An understanding gained through studying its non-functional alternatives. *Synlett* **2014**, *25*, 1511–1517. [CrossRef]
16. Krishnamurthy, R. On the emergence of RNA. *Isr. J. Chem.* **2015**, *55*, 837–850. [CrossRef]
17. Schöning, K.U.; Scholz, P.; Guntha, S.; Wu, X.; Krishnamurthy, R.; Eschenmoser, A. Chemical etiology of nucleic acid structure: The α-threofuranosyl-(3′-2′) oligonucleotide system. *Science* **2000**, *290*, 1347–1351.
18. Leu, K.; Obermayer, B.; Rajamani, S.; Gerland, U.; Chen, I.A. The prebiotic evolutionary advantage of transferring information from RNA to DNA. *Nucleic Acids Res.* **2011**, *39*, 8135–8147. [CrossRef] [PubMed]
19. Malyshev, D.A.; Romesberg, F.E. The expanded genetic alphabet. *Angew. Chem.* **2015**, *54*, 11930–11944. [CrossRef] [PubMed]
20. Benner, S.A.; Karalkar, N.B.; Hoshika, S.; Laos, R.; Shaw, R.W.; Matsuura, M.; Fajardo, D.; Moussatche, P. Alternative Watson-Crick synthetic genetic systems. *Cold Spring Harb. Perspect. Biol.* **2016**. [CrossRef] [PubMed]
21. Leu, K.; Kervio, E.; Obermayer, B.; Turk-MacLeod, R.; Yuan, C.; Luevano, J.-M.; Chen, E.; Gerland, U.; Richert, C.; Chen, I.A. Cascade of reduced speed and accuracy after errors in enzyme-free copying of nucleic acid sequences. *J. Am. Chem. Soc.* **2013**, *135*, 354–366. [CrossRef] [PubMed]
22. Bridson, P.K.; Orgel, L.E. Catalysis of accurate poly(C)-directed synthesis of 3′-5′-linked oligoguanylates by Zn^{2+}. *J. Mol. Biol.* **1980**, *144*, 567–577. [CrossRef]
23. Inoue, T.; Orgel, L.E. Oligomerization of (guanosine 5′-phosphor)-2-methylimidazolide on poly(C). *J. Mol. Biol.* **1982**, *162*, 201–217. [CrossRef]
24. Kozlov, I.A.; Politis, P.K.; Van Aerschot, A.; Busson, R.; Herdewijn, P.; Orgel, L.E. Nonenzymatic synthesis of RNA and DNA oligomers on hexitol nucleic acid templates: The importance of the A structure. *J. Am. Chem. Soc.* **1999**, *121*, 2653–2656. [CrossRef] [PubMed]
25. Rohatgi, R.; Bartel, D.P.; Szostak, J.W. Monenzymatic, template-directed ligation of oligoribonucleotides is highly regioselective for the formation of 3′-5′ phosphodiester bonds. *J. Am. Chem. Soc.* **1996**, *118*, 3340–3344. [CrossRef] [PubMed]
26. Engelhart, A.E.; Powner, M.W.; Szostak, J.W. Functional RNAs exhibit tolerance for non-heriTable 2′-5′ versus 3′-5′ backbone heterogeneity. *Nat. Chem.* **2013**, *5*, 390–394. [CrossRef] [PubMed]
27. Shen, F.; Luo, Z.; Liu, H.; Wang, R.; Zhang, S.; Gan, J.; Sheng, J. Structural insights into RNA duplexes with multiple 2′-5′-linkages. *Nucleic Acids Res.* **2017**, *45*, 3537–3546. [CrossRef] [PubMed]
28. Usher, D.A.; McHale, A.H. Hydrolytic stability of helical RNA: A selective advantage for the natural 3′-5′ bond. *Proc. Natl. Acad. Sci. USA* **1976**, *73*, 1149–1153. [CrossRef] [PubMed]
29. Mariani, A.; Sutherland, J.D. Non-enzymatic RNA backbone proofreading through energy-dissipative recycling. *Angew. Chem.* **2017**, *56*, 6563–6566. [CrossRef] [PubMed]
30. Glavin, D.P.; Dworkin, J.P. Enrichment of the Amino Acid L-Isovaline by Aqueous Alteration on Cl and CM Meteorite Parent Bodies. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 5487–5492. [CrossRef] [PubMed]
31. Breslow, R.; Chen, Z.-L. L-amino acids catalyze the formation of an excess of D-glyceralidehyde, and thus of other *D* sugars under credible prebiotic conditions. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 5723–5725. [CrossRef] [PubMed]
32. Hein, J.E.; Blackmond, D.G. On the origin of single chirality of amino acids and sugars in biogenesis. *Acc. Chem. Res.* **2012**, *45*, 2045–2054. [CrossRef] [PubMed]
33. Higgs, P.G. Chemical evolution and the evolutionary definition of life. *J. Mol. Evol.* **2017**, *84*, 225–235. [CrossRef] [PubMed]

34. Hud, N.V.; Cafferty, B.J.; Krishnamurthy, R.; Williams, L.D. The origin of RNA and "my grandfather's axe". *Chem. Biol.* **2013**, *20*, 466–474. [CrossRef] [PubMed]
35. Higgs, P.G. The effects of limited diffusion and wet-dry cycling on reversible polymerization reactions: Implications for prebiotic synthesis of nucleic acids. *Life* **2016**, *6*, 24. [CrossRef] [PubMed]
36. Derr, J.; Manapat, M.L.; Rajamani, S.; Leu, K.; Xulvi-Brunet, R.; Joseph, I.; Nowak, M.A.; Chen, I.A. Prebiotically plausible mechanisms increase compositional diversity of nucleic acid sequences. *Nucleic Acids Res.* **2012**, *40*, 4711–4722. [PubMed]
37. Tkachenko, A.V.; Maslov, S. Spontaneous emergence of autocatalytic information coding polymers. *J. Chem. Phys.* **2015**, *143*. [CrossRef] [PubMed]

Chapter 4: Can the RNA world Function without cytidine?

The contents of this chapter were published in *Molecular Biology and Evolution* in September 2019. The manuscript, figures, and tables in this chapter are used with permission from the publisher, Oxford University Press.

Ralph Pudritz provided the initial inspiration and idea for the project. Paul Higgs and I contributed to the design of the models and the writing of the manuscript. Ralph Pudritz helped edit the final manuscript. I wrote the computer programs to determine the statistics of RNA secondary structures, accumulate the data, and make the corresponding figures. Paul Higgs wrote the two-step model of RNA replication which incorporates G-U wobble base pairing.

Tupper, A. S., Pudritz, R. E., & Higgs, P. G. (2020). Can the RNA World Still Function without Cytidine?. *Molecular Biology and Evolution*, *37*(1), 71-83.

Can the RNA World Still Function without Cytidine?

Andrew S. Tupper,[1] Ralph E. Pudritz,[2] and Paul G. Higgs*[,2]

[1]Origins Institute and Department of Biochemistry and Biomedical Sciences, McMaster University, Hamilton, Ontario, Canada

[2]Origins Institute and Department of Physics and Astronomy, McMaster University, Hamilton, Ontario, Canada

***Corresponding author:** E-mail: higgsp@mcmaster.ca.

Associate editor: Sergei Kosakovsky Pond

Abstract

Most scenarios for the origin of life assume that RNA played a key role in both catalysis and information storage. The A, U, G, and C nucleobases in modern RNA all participate in secondary structure formation and replication. However, the rapid deamination of C to U and the absence of C in meteorite samples suggest that prebiotic RNA may have been deficient in cytosine. Here, we assess the ability of RNA sequences formed from a three-letter AUG alphabet to perform both structural and genetic roles in comparison to sequences formed from the AUGC alphabet. Despite forming less thermodynamically stable helices, the AUG alphabet can find a broad range of structures and thus appears sufficient for catalysis in the RNA World. However, in the AUG case, longer sequences are required to form structures with an equivalent complexity. Replication in the AUG alphabet requires GU pairing. Sequence fidelity in the AUG alphabet is low whenever G's are present in the sequence. We find that AUG sequences evolve to AU sequences if GU pairing is rare, and to RU sequences if GU pairing is common (R denotes A or G). It is not possible to conserve a G at a specific site in either case. These problems do not rule out the possibility of an RNA World based on AUG, but they show that it wouldbe significantly more difficult than with a four-base alphabet.

Key words: **RNA World, cytosine deamination, RNA secondary structure, error threshold, ribozyme.**

Article

Introduction

The RNA World theory for the origin of life postulates that the first life, or a very early stage of life, utilized RNA to store genetic information and catalyze chemical reactions (Robertson and Joyce 2012; Higgs and Lehman 2015). Skeptics of this theory have cited the difficulty of forming RNA prebiotically and have proposed simpler polymers deemed pre-RNA and a corresponding pre-RNA World which predates the RNA world (Lazcano and Miller 1996; Hud et al. 2013). One difficulty of forming RNA prebiotically is the instability of cytosine nucleobases. Cytosine undergoes a deamination reaction to form uracil with a half-life on the order of ~100 years at 37 °C, as opposed to the ~2,500 years half-life for adenine and guanine, and ~250,000 years half-life for uracil (Levy and Miller 1998). Cytosine nucleobases are also absent from samples of carbonaceous chrondrites, whereas many other organic compounds, including adenine, guanine, and uracil, are found (Pearce and Pudritz 2015). The instability of cytosine and its absence from meteorites suggests that prior to biological innovation the first genetic polymers were deficient in cytosine. C to U deamination also occurs in modern organisms and is one of the "fastest" and most frequent forms of mutation (Lewis et al. 2016). This is not a problem, because all mutations are "slow" on the time scale of replication of modern organisms. However, if we are relying on slow processes of prebiotic nucleotide synthesis, or the gradual accumulation of nucleotides delivered from meteorites and/or dust, then the shorter lifetime of C becomes important. Therefore, we ask whether it is possible that an AUG alphabet could have operated prior to the AUGC alphabet in modern RNA.

Previous experimental studies show that sequences formed from three-letter and two-letter alphabets can have enzymatic activity. A ligase ribozyme derived from the AUG alphabet was shown to catalyze its reaction 10^5-fold faster than the uncatalyzed reaction (Rogers and Joyce 1999). Similar ligase ribozymes have also been discovered in the DUG and DU alphabets, where D is the nonstandard 2-6 diaminopurine nucleobase (Reader and Joyce 2002). These alternative alphabets have some enzymatic activity, although it is not yet clear how well they could catalyze the diverse chemical reactions required by early life in comparison to the standard four-letter alphabet.

In addition to enzymatic activity, the first genetic polymer must be able to efficiently transfer sequence and structure information upon replication. In canonical RNA, Watson–Crick base pairing of A to U and G to C allows information to be transferred between strands during the two stage replication process; plus to minus, followed by minus to plus. Mispairing of bases leads to errors in sequence replication. G-U pairing is of particular interest because, in replication of AUGC sequences, G-U pairing is the dominant form of error under both nonenzymatic (Leu et al. 2011) and enzymatic conditions (Johnston et al. 2001; Attwater et al. 2013). However, for an AUG alphabet, G-U pairing is essential for replication of G nucleobases, as G lacks a Watson–Crick complement. Thus, the ability of RNA to form G-U pairs is advantageous in allowing flexible structure formation but is a

Mol. Biol. Evol. 37(1):71–83 doi:10.1093/molbev/msz200 Advance Access publication September 4, 2019

hindrance to accurate replication, and this conflict is particularly relevant in the AUG alphabet.

The aim of this article is to assess the ability of an AUG alphabet to perform both the structural and genetic roles necessary for early life. RNAs fold to stem-loop secondary structures by base pairing between matching regions in different parts of the sequence. The ViennaRNA folding routine (Lorenz et al. 2011) can be used to find the minimum free energy (MFE) secondary structure for a given sequence. If base-pairing energies are weak, if there are few matching bases, or if the temperature is too high, the MFE structure may be unfolded (no base pairing) or something very simple like a single hairpin loop. To be functional structurally, we assume that the sequence must form a stable stem-loop secondary structure similar to those of known ribozymes and biological RNAs, that is, it should form a structure with several different stems and loops, and the structure might be branched, or composed of several domains.

The statistics of structures formed by AUGC sequences, and also by two-letter sequences (either GC or AU), have been considered previously (Fontana et al. 1993) using the ViennaRNA folding routine. The method makes use of experimentally measured thermodynamic parameters (Turner and Mathews 2010), and it produces structures that are similar to those of real RNAs, even though three dimensional information is not used (Higgs 2000). We thus assume that this method is useful as a means to survey the statistical properties of structures formed by large numbers of sequences, even though it does not give fully accurate secondary structure predictions for every real sequence. Statistical properties of structures of AUG sequences have not been studied previously. Here, we ask whether the structures formed by AUG sequences are qualitatively similar to those of AUGC sequences. If the AUG alphabet is able to form a variety of relatively complex secondary structures then it may be sufficient in the structural role for the origin of life.

To test the genetic role of the AUG alphabet, we developed a computational model of RNA replication that includes essential features that are usually ignored in models of nucleic acid evolution. Typical models of gene sequence evolution, such as those used in molecular phylogenetics (Higgs and Attwood 2005; Posada 2008), use a four-letter model for replication of double-stranded DNA in which accurate replication of a functional (plus) strand produces another plus strand. However, replication of single-stranded RNA is a two-step process in which a complementary (minus) strand is copied from a plus strand, and a new plus strand is copied from a minus strand. Theoretical evolutionary models for RNA replication (Eigen et al. 1988; Reidys et al. 2001; Kun et al. 2005; Takeuchi et al. 2005; Takeuchi and Hogeweg 2012; Szilagyi et al. 2014) often ignore the minus strands and treat the two steps as a single step, although minus strands can also be included (Shay et al. 2015). None of these previous models captures the importance of G-U pairing in AUG sequences. For the present study, we used a novel model that describes the two steps separately and includes parameters for rates of G-U pairing relative to Watson–Crick pairing and to other types of mismatches.

For the AUG alphabet to be viable genetically, it must allow the functionality of molecules to be inherited during replication. Early models of RNA replication, such as the quasispecies theory (Eigen et al. 1988), demonstrate that replication accuracy is a key factor for survival of replicating RNAs. The simplest case considers a single master sequence replicating at a high rate r_0, in competition with all possible mutant sequences that replicate at a lower rate r_1. The sequence fidelity, Q, is the probability that the master sequence is replicated without sequence error. The net rate of accurate replication is Qr_0, and the theory shows that the master sequence only survives if this rate is faster than r_1. The concentration of master sequences falls to zero at the error threshold, where $Q = r_1/r_0$. Subsequent theories assume that RNA function depends on secondary structure, and consider the case where all sequences folding to the same specified secondary structure replicate at the same rate, while all sequences folding to a different structure replicate more slowly. In this case, we can define a structural fidelity, which is the probability that a sequence with the correct structure gives rise to an offspring sequence with the same structure. There is a phenotypic (or structural) error threshold in these models, where the fidelity becomes too low to maintain the functional structure (Reidys et al. 2001; Kun et al. 2005; Takeuchi et al. 2005; Szilagyi et al. 2014).

Here, using our two-step model for RNA replication, we determine the average sequence fidelity and structure fidelity for sequences in the AUG and AUGC alphabets. We show that sequence fidelity is an unavoidable problem in the AUG alphabet, unless G bases are excluded, or are interchangeable with A bases.

Results

Comparison of Secondary Structures Formed with Different Nucleotide Alphabets

In this section, we assess the ability of AUG sequences to fold to the kinds of structures that would be required in the RNA World. We show that a wide range of structures is possible in AUG sequences, although folding is slightly more difficult than for AUGC sequence.

Following (Fontana et al. 1993), we define the stickiness, S, of an alphabet as the fraction of two-base combinations that form an allowed pair in the secondary structure. The stickiness is lower for alphabets with larger numbers of nucleotides (see table 1). The other important property is the strength of hydrogen bonding and stacking interactions of base pairs in helices. Table 1 shows the mean nearest-neighbor free energy ΔG_{nn} for each alphabet using the "Turner 2004" thermodynamic model (Turner and Mathews 2010) assuming all possible nearest-neighbor pairs occur randomly. Alphabets that include GC pairs have stronger stacking (more negative ΔG_{nn}) than alphabets that exclude GC pairs. The AUG alphabet is the weakest-pairing alphabet (least negative ΔG_{nn}). It is weaker than the AU alphabet because it includes GU pairs, which are less stable than AU pairs in most configurations.

We generated 10^6 random sequences for each length L in the range 1–100 nucleotides and determined the MFE

Table 1. Comparison of Properties of Alphabets with Differing Numbers of Nucleotides.

Alphabet	Allowed Pairs	Stickiness S	ΔG_{nn} (kcal/mol)
AUGC	AU, UA, GC, CG, GU, UG	3/8	−1.59
AUG	AU, UA, GU, UG	4/9	−0.76
AU	AU, UA	1/2	−1.05
GC	GC, CG	1/2	−3.10

structure of each sequence using the ViennaRNA folding software (Lorenz et al. 2011). Figure 1*a* shows the probability that the MFE structure was a folded structure of any kind (i.e., at least one base pair present in the MFE structure). This probability tends to 1 for all alphabets if the sequences are sufficiently long. For short sequences, the folding probability is strongly dependent on ΔG_{nn}. The alphabets with strongest base pairs (GC and AUGC) have high probabilities of folding to secondary structures even when sequences are short, whereas the more weakly pairing alphabets (AU and AUG) require longer sequences before folding becomes common. With 20-mers, for example, almost all GC sequences are folded in some way, but only 6% of AUG sequences are folded. With 50-mers, almost all sequences in the GC, AUGC, and AU alphabets are folded, whereas only about 80% of AUG sequences are folded.

Figure 1*b–d* shows the probabilities of formation of structures of increasing complexity. In figure 1*b*, we consider all structures that contain at least two hairpin loops; in figure 1*c*, we consider all structures with at least two separate domains (see Materials and Methods); and in figure 1*d*, we consider all structures with at least one multibranched loop. These probabilities are always lower for the AUG alphabet than any of the others. Multibranched loops (as in fig. 1*d*) are particularly rare in the AUG alphabet, even for the longest sequences we considered ($L = 100$). This does not mean that complex structures are impossible in the AUG alphabet. It merely shows that there are fewer AUG sequences that fold to complex structures, and that longer sequences will typically be required for forming structures of a given degree of complexity.

In figure 2, we consider the probability of formation of specific secondary structures, rather than the combined probabilities of structures of different types. For any structure with a specified base-pairing pattern, we define P_{str} as the probability that a random sequence folds to an MFE structure with exactly this pattern, and P_{comp} as the probability that a random sequence is compatible with the structure (i.e., it has allowed base pairs in all the required places). These probabilities are calculated in Materials and Methods. A compatible sequence does not necessarily fold to the correct structure because an alternative structure may be more stable.

Figure 2 shows the calculated P_{str} for two series of structures, one with three helices and one with four. The number of base pairs per helix, $n_{pairs/helix}$, is the same in each helix. The structure is shown for $n_{pairs/helix} = 3$ in each panel of figure 2. In each series, when $n_{pairs/helix}$ is increased, we keep the number of bases in the single-stranded regions constant. The total length of the sequence is $L = 11 + 6n_{pairs/helix}$ for the series of structures with three helices, and $L = 16 + 8n_{pairs/helix}$ for the series with four helices. The detection limits are also shown in figure 2 (see Materials and Methods). When no symbol is shown, the structure is too rare to be detected in the sample we analyzed.

For each alphabet there is a minimum $n_{pairs/helix}$ for which each structure is detected in our sample. This minimum length is shorter for strongly pairing alphabets than weakly pairing alphabets. The three-helix structure in figure 2*a* is detectable for $n_{pairs/helix} \geq 3$ in the GC and AUGC alphabets (meaning $L \geq 29$), whereas it is only detectable for $n_{pairs/helix} \geq 7$ in the AU and AUG alphabets ($L \geq 53$). Similarly, the clover-leaf structure in figure 2*b* is detectable if $n_{pairs/helix} \geq 3$ in the AUGC alphabet ($L \geq 40$), whereas it is only detectable for $n_{pairs/helix} \geq 6$ in the AUG alphabet ($L \geq 64$).

Similar curves for two other example structures are shown in figure 3. These structures have several helices but no multibranched loop, whereas the structures in figure 2 contain a multibranched loop. These results in figures 2 and 3 show similar behavior for all the structures. Longer sequences are required in the AUG alphabet in order to form complex structures, and complex structures with short helices are so rare that we cannot detect them. The requirement for longer sequences in the AUG alphabet relative to the AUGC alphabet is a significant disadvantage, because longer sequences are more difficult to make by chemical means and they require a higher replication accuracy per base in order to replicate without error.

It can also be seen in figures 2 and 3 that P_{str} approaches the limit of P_{comp} as $n_{pairs/helix}$ increases, because when there are many pairs in a helix, it is almost certain that the helix is stable and that there is no other structure that is more stable. As P_{comp} increases with the stickiness, and as stickiness increases as we move from four- to three- to two-letter alphabets, we find that P_{str} is lowest in the four-letter alphabet. However, the number of possible sequences that fold to the given structure is $N_{str} = B^L P_{str}$, where B is the number of bases in the alphabet. Figure 2*c* and *d* shows that N_{str} increases for larger alphabets. Thus, there are many more possible sequences that form a given structure in the larger alphabets. Formation of the correct secondary structure is a minimal criterion for ribozyme function, but it is likely that not all sequences with the correct secondary structure are functional, and that not all function equally well. The larger the number of sequences with the correct secondary structure, the greater will be the diversity of sequences with high enzymatic activity, and the higher the level of activity of the best ones is likely to be. This is another advantage of the four-letter alphabet over three- and two-letter alphabets.

From this section, we conclude that it is possible to form structures of a wide range of types with all four of the alphabets considered. Therefore, from structural considerations, we do not rule out the possibility that AUG sequences could have been possible in an early RNA World. Nevertheless, the AUG alphabet has several disadvantages: 1) complex structures with short helices are not possible, 2) longer sequences are required for complex structures, making abiotic synthesis more difficult and error-free replication less

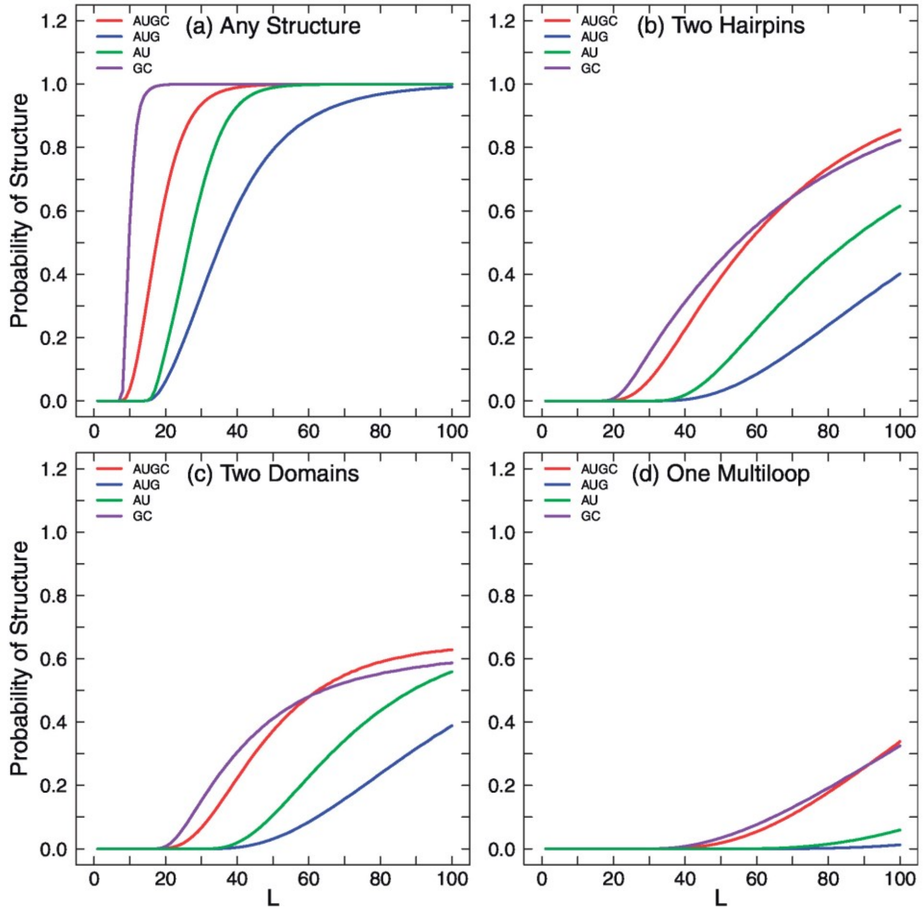


Fig. 1. (*a*) Probability of forming a folded structure of any kind, (*b*) probability that the minimum free energy (MFE) contains at least two hairpins, (*c*) probability that the MFE contains at least two separate domains, and (*d*) probability that the MFE contains at least one multiloop.

likely, 3) the diversity of functional sequences for a given structure is lower, and 4) the enzymatic activity of the best sequences is likely to be lower.

The Fidelity of Sequence Replication

We now assess the ability of AUG sequences to replicate in comparison to AUGC sequences. We are interested in understanding the conditions in which long sequences can be accurately replicated. We therefore want to determine the fidelity of sequence replication per base, and the overall fidelity of replication of the whole sequence.

When we look at alignments of RNAs with conserved function, we often find that the secondary structure is conserved, and that there are also particular sites in the molecule where the sequence is conserved. However, large amounts of sequence variation is possible at many other sites, and the number of sites that need to be conserved may be much smaller than the total length of the molecule. In this article, therefore, we consider a functional sequence to be one possessing a specified secondary structure and a specified set of conserved bases at particular positions. We consider replication as a two-step process from plus strand to minus, then from minus strand to plus. Beginning with a functional sequence that possesses the correct structure and required conserved bases, we define the sequence fidelity, Q_{seq}, as the probability that an offspring sequence formed by the two-step replication process still possesses the required conserved bases, and we define the structural fidelity, Q_{str}, as the probability that the offspring still folds to the correct secondary structure.

In this section, we consider the sequence fidelity. We define the one-step probability matrix p_{ij} as the probability that base

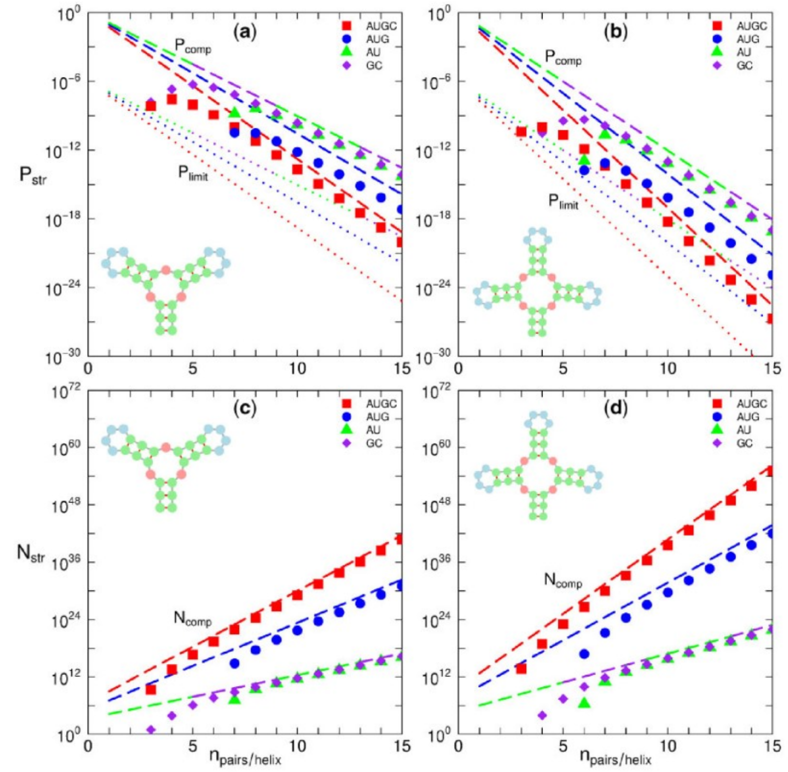


Fig. 2. (*a*, *b*) Data points show the probabilities P_{str} that a random sequence folds to the structures shown, as a function of the number of pairs $n_{pairs/helix}$ in each helix. The structures are illustrated with $n_{pairs/helix} = 3$. Dashed lines show the probability that a random sequence is compatible with the structure, and dotted lines show the detection limits. (*c*, *d*) Data points show the number N_{str} of random sequences that fold to the same structures. Dashed lines show the number of compatible sequences.

i in the template leads to base *j* in the complementary strand (where *i* and *j* = A, U, G, or C). For a two-step cycle, the probability that base *i* in the initial plus strand becomes base *j* in the descendant plus strand is $q_{ij} = \sum_k p_{ik}\ p_{kj}$. The diagonal elements q_{ii} are the fidelities of replication of each base. Consider a sequence with numbers n_A, n_U, n_G, and n_C sites of each base that are required to be conserved. The sequence fidelity is $Q_{seq}(n_A, n_U, n_G, n_C) = q_{AA}^{n_A} q_{UU}^{n_U} q_{GG}^{n_G} q_{CC}^{n_C}$. When this fidelity is low, sequences cannot be maintained by selection in a population. The minimum fidelity required for replication depends on the replication rates of functional and nonfunctional sequences, as described above. We do not wish to consider specific values of the replication rates here, we simply wish to demonstrate that the fidelity of sequences containing G bases is extremely low in the AUG alphabet.

Accurate replication in each of the two steps means that a base gives rise to its complementary base in the opposing strand (A → U, G → C, etc.). In experiments (Johnston et al. 2001; Leu et al. 2011; Attwater et al. 2013), it is observed that the most likely complementary base is the Watson–Crick base, and that all the other bases can be added to the complementary strand with smaller probabilities. It is also observed that the most significant probabilities of the non-Watson–Crick bases are for the "wobble" pairs G → U and U → G. Here, we consider a model where the relative rates of addition of Watson–Crick, wobble, and mismatch bases are 1 : *w* : *v*, where $v \le w \le 1$. "Mismatch" refers to any combination of base pairs that is not a Watson–Crick or a wobble pair. For simplicity, we suppose all mismatch combinations occur at the same relative rate *v*. If sequences are repeatedly replicated under this mutational model without any selection acting, the frequencies of the bases in the sequences will converge to steady-state frequencies f_i. The steady-state frequencies and the fidelities of each base are given in Materials and Methods.

For the AUGC alphabet, A and C are equivalent, and U and G are equivalent in our model. Therefore, the fidelities and base frequencies are only shown for A and U in figure 4. We calculate quantities as a function of the wobble rate *w*,

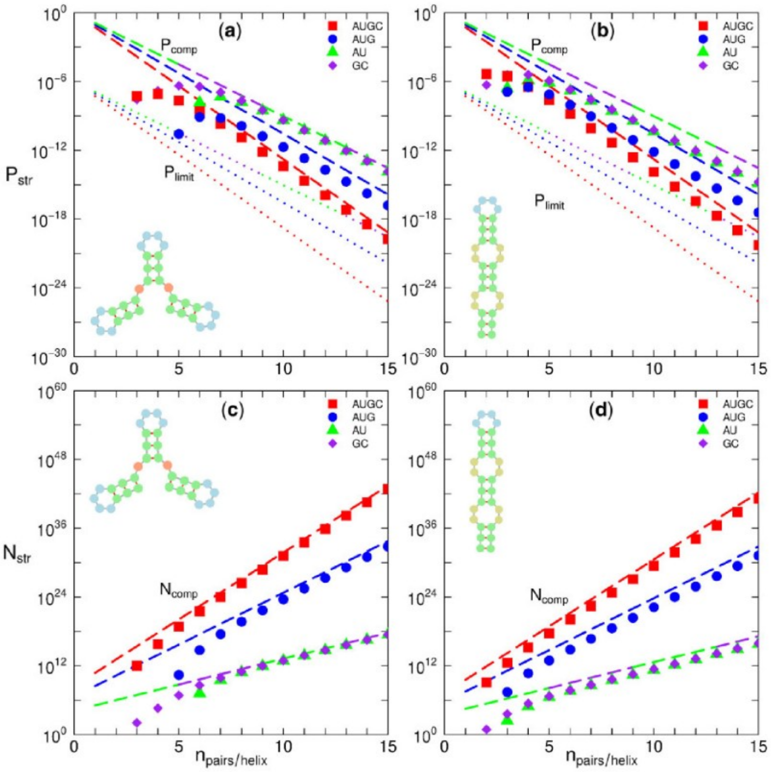


Fig. 3. (*a*, *b*) Data points show the probabilities P_{str} that a random sequence folds to the structures shown, as a function of the number of pairs $n_{pairs/helix}$ in each helix. The structures are illustrated with $n_{pairs/helix} = 3$. Dashed lines show the probability that a random sequence is compatible with the structure, and dotted lines show the detection limits. (*c*, *d*) Data points show the number N_{str} of random sequences that fold to the same structures. Dashed lines show the number of compatible sequences.

assuming that the mismatch rate v varies in proportion to w with a fixed ratio $r = v/w$. We chose $r = 0.25$ in this example, following out expectation that this ratio will be substantially <1. However, the value of r does not change the qualitative conclusions below. The fidelities q_{UU} and q_{GG} are equal and are slightly greater than q_{AA} and q_{CC}. The stationary frequencies f_U and f_G are slightly higher than f_A and f_C. This occurs because, when $w > v$, U and G bases are incorporated into sequences at a slightly higher rate than A and C, even though the four bases are assumed to be available as monomers at equal frequencies. For the accurate replication of long sequences, we require both v and w to be small. In the limit where v and w tend to zero with fixed ratio $r = v/w$, figure 4a shows that $f_i \rightarrow 0.25$ and figure 4c shows that $q_{ii} \rightarrow 1$ for all bases. The four-letter AUGC alphabet is "well-behaved," in the sense that it has high fidelity and unbiased base composition when the error rate is low.

In contrast, the three-letter AUG alphabet is not well-behaved in the above sense. Figure 4b and d shows the base frequencies and fidelities in the AUG alphabet as a function of w, keeping the ratio $r = 0.25$. There is a very large difference in the stationary frequencies f_A, f_U, and f_G. Even though we have assumed that the three bases are of equal concentration in the monomer mixture, the rate of incorporation into sequences is significantly different. The rate of incorporation of G depends on the rate w. When w and v both tend to zero, f_G tends to zero, meaning that the G bases are excluded from sequences that evolve under this mutational model. The AUG alphabet becomes a two-letter AU alphabet under conditions where the error rate is small.

Figure 4d also shows that the fidelities are significantly different for the three bases. Although q_{AA} and q_{UU} tend to 1 when the error rate is small, q_{GG} is very low for all values of w and actually decreases when w tends to zero. In Materials and Methods, it is shown that $q_{GG} = r^2/(1+2r)^2$ in this limit, which is always very small for all possible values of $r \leq 1$. This means that G bases cannot be maintained in a conserved position in the sequence in the AUG alphabet. Another case worth considering is where either of the two purines, A or G,

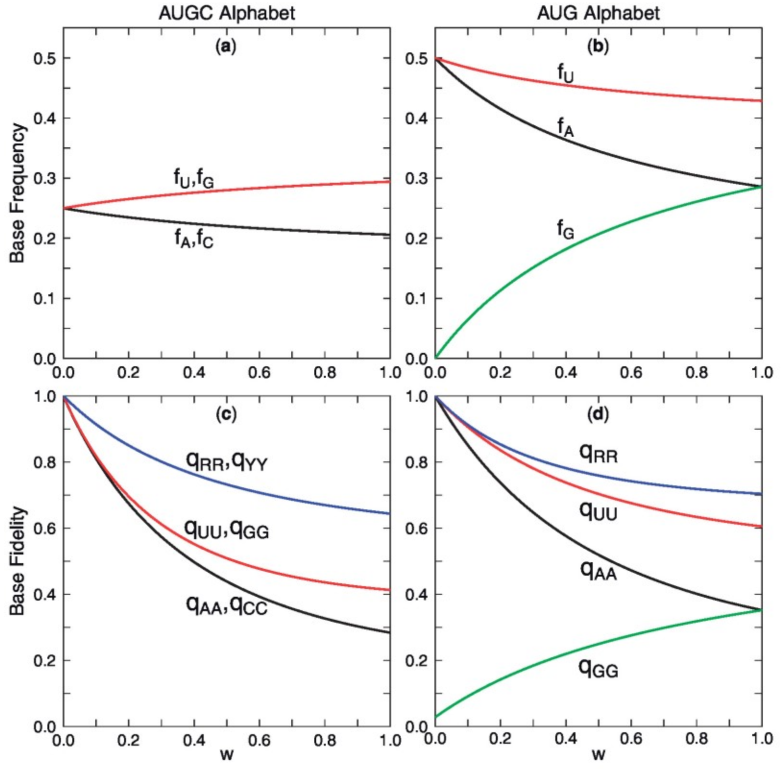


Fig. 4. Stationary base frequencies (*a*, *b*), and single-base fidelities (*c*, *d*), as a function of wobble rate, *w*, for the AUGC alphabet (left) and the AUG alphabet (right). The ratio $r = v/w = 0.25$ in all cases.

would do equally well in a sequence, and it is simply necessary to distinguish U from a purine (R denotes a purine, A or G). Figure 4*c* and *d* shows that the purine fidelity q_{RR} is high in both alphabets, as we discuss further below.

As an example, in figure 5, we consider several sequences in which there are $N = 8$ conserved sites in total, and the number of conserved A, U, G, and C sites is specified. Three examples were chosen with different numbers of A and G bases. In the AUGC alphabet, the sequence fidelity Q is only very slightly different for the three sequences, because q_{AA} and q_{GG} are almost equal for these parameters. All three sequences have a high fidelity when the error rate is low; hence these can be maintained in the population by selection if $Q > Q_{\min}$. In contrast, in the AUG alphabet, the sequence fidelity is very strongly dependent on the number of G's in the sequence. Addition of even a single G, makes the fidelity extremely low. The sequence AAAGUUUU, with a single G, has Q <5% over the whole range of *w*, and the sequence with two G's is hardly visible on this scale. Thus, sequences with even a single required G are almost bound to be beyond the error threshold, whatever the value of *w* and whatever the replication rate ratio r_1/r_0.

We have seen that the AUG alphabet becomes a two-letter AU alphabet when *w* is small, and that only sequences without G's have a high fidelity. However, if A and G bases are alternatives as regards function and structure, the AUG alphabet is effectively a two-letter RU alphabet. In this case, the ratio of A to G in the sequence depends on *w*, and G's are still present in the sequence interchangeably with A's. The fidelity for a single R base, q_{RR}, is given in the Materials and Methods, and plotted in figure 4*d*. This tends to 1 when the error rate is small. The sequence fidelity in the RU alphabet is $Q_{\mathrm{seq}}(n_R, n_U) = q_{RR}^{n_R} q_{UU}^{n_U}$. In figure 5*b*, all three AUG sequences reduce to RRRRUUUU in the RU alphabet. The RU sequence fidelity is high (blue line). Thus, the AUG alphabet could work when *w* is high, but only if A and G bases are interchangeable. It would not be possible to specify either an A or G separately. In a similar way, using Y = U or C to denote a pyrimidine, the AUGC alphabet might reduce to a two-letter RY alphabet, if U and C were interchangeable. The fidelity in the RY alphabet would be $Q_{\mathrm{seq}}(n_R, n_Y) = q_{RR}^{n_R} q_{YY}^{n_Y}$. This is shown for the sequence RRRRYYYY in figure 5*a*.

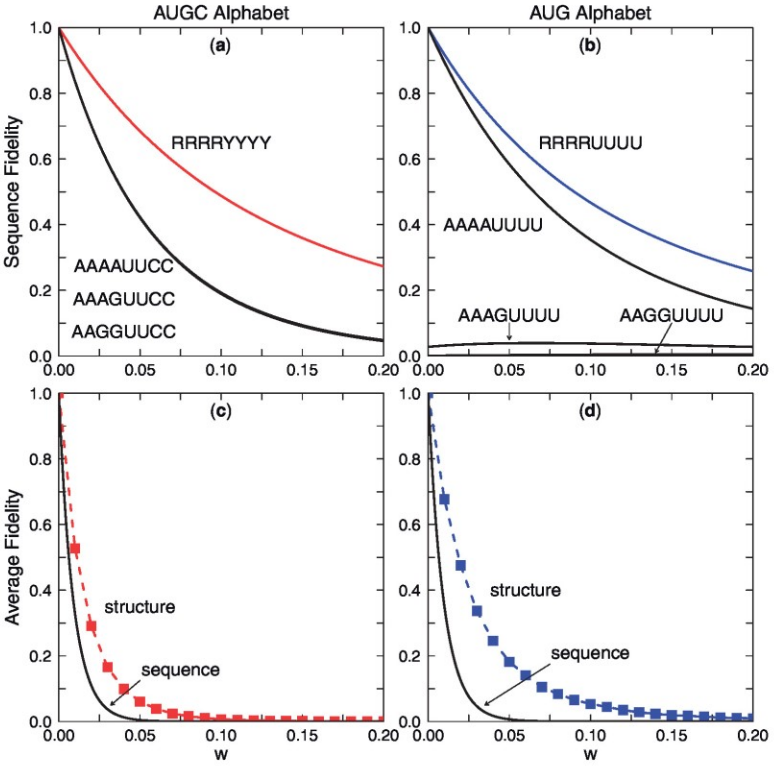


Fig. 5. The sequence fidelity of three example sequences with varying composition of conserved bases is shown in the AUGC and AUG alphabets as a function of *w*, with $r = v/w = 0.25$. The bottom two plots show average structure fidelities for sequences of length 50 containing 10 bonds. (*a*) In the AUGC alphabet, the sequences have almost the same fidelity. (*b*) In the AUG alphabet, the fidelity depends strongly on the number of G's in the sequences, and is very low, unless G's are completely absent. At the purine–pyrimidine level, fidelities are high if we work with RU or RY sequences. (*c, d*) The average sequence fidelity and structure fidelity are shown in the 2 alphabets for sequences of length 50.

From this section, we conclude that accurate three-letter replication is not possible. Either G's are excluded, or G's are used interchangeably with A's. Thus it is not possible to specify a G alone as a required base for function if there are no C's present.

Structural Fidelity

We now consider the structural fidelity, Q_{str}, which is the probability that an offspring strand after the two-step replication has the same secondary structure as the parent strand. Rather than calculate this for many individual structures, we determined the average structural fidelity of typical sequences/structures that evolve in a given alphabet. We considered 10^6 sequences of length 50 nucleotides replicating independently such that their base frequencies reach their stationary frequencies as defined previously. We then determined the MFE structure of each of these. Sequences whose MFE structure contained exactly ten base pairs were retained, in order to give a representative sample of alternative structures with a typical number of base pairs in comparison to the sequence length. These sequences were folded and replicated one further time in order to determine the probability that the offspring has the same structure as its parent. This probability is the average structural fidelity of typical structures of length 50 containing 10 base pairs.

In figure 5c and *d*, the average structure fidelity of the AUG and AUGC alphabets are shown as a function of w with fixed $v/w = 0.25$. In the limit of $w \rightarrow 0$, structure fidelities of both alphabets approach 1, meaning that structure can be accurately passed on in both alphabets. As *w* increases, structure fidelities of both alphabets decrease exponentially, with the AUG alphabet having an advantage over the AUGC alphabet for all values of *w*. The average sequence fidelities of the same sequences are also shown assuming that all 50 sites have required bases.

When the GU pairing rate *w* is small, both alphabets have high sequence and structure fidelity. The fact that the average sequence fidelity of the AUG alphabet in figure 5*d* tends to 1, when $w \rightarrow 0$ may seem surprising, given that we have shown in figure 5*b* that sequence fidelity is very low, even if there is only one G in the sequence. The reason is that typical sequences do not contain G's when w is small! This is because G's are excluded by the mutation process, as is shown by our mutational model above. Thus, the AUG alphabet effectively becomes a two-letter AU alphabet. The two-letter AU alphabet is well behaved, with a high fidelity if the error rate is small.

The concept of structural fidelity arises in "neutral network" models of RNA replication in which the function depends only on the secondary structure and there is a large set of sequences of equal fitness that all fold to this same structure (Reidys et al. 2001; Kun et al. 2005; Takeuchi et al. 2005; Szilagyi et al. 2014). To assume that the sequence is not important is an oversimplification, but it is nevertheless useful in explaining why RNAs such as tRNAs and rRNAs can evolve substantially in sequence while maintaining the same structure (Savill et al. 2001). In reality, some elements of sequence will be important too, in addition to maintaining the structure, but this is still far from the other oversimplified case where there is only a single master sequence that is functional, and every mutation is deleterious. Here, we have measured the two types of fidelity separately in order to distinguish them. We conclude that, if a structure occurs with appreciable frequency in the AUG alphabet, then its structural fidelity will be reasonably high, and comparable to the structural fidelity in the AUGC alphabet. Structural fidelity is therefore no more of a problem for the AUG case than for the AUGC case. The serious problem for the AUG case is the sequence fidelity, as we showed in the previous section, where it is difficult to maintain even very small numbers of conserved bases.

Discussion

We have asked the question of whether the RNA World could be possible in absence of C bases. Combining the conclusions from our studies of structure and replication, we conclude that this possibility is not entirely ruled out, but there are several factors that make an AUG alphabet less suitable for an RNA World than an AUGC alphabet.

The GU pairing rate plays an important role in the analysis in this article. It is worth remembering that GU pairing is a large source of errors in experimental studies of RNA replication (Johnston et al. 2001; Leu et al. 2011; Attwater et al. 2013) when A, U, G, and C are all present. The fact that GT mispairing in DNA is less frequent than GU mispairing in RNA is one major advantage of DNA over RNA as a genetic polymer (Leu et al. 2011), and this is consistent with the idea that there was a genetic take-over of information by DNA from RNA. The large GU error rate is an important factor that would limit the fidelity of replication and the maximum sequence lengths that are sustainable in the usual AUGC RNA World. This article shows that this problem would be substantially greater if the RNA World depended on the AUG alphabet. It is perhaps intuitive that something goes wrong with replication of G's in an AUG alphabet, but it would not be obvious what would happen without doing a quantitative analysis. Our article makes it clear that two things happen when replication happens with the three-letter alphabet. If the wobble rate is low, the G's disappear and it becomes a two-letter AU alphabet. If the wobble rate is high, it becomes an RU alphabet.

Details of nonenzymatic replication are complicated. Leu et al. (2013) show that the occurrence of a first mismatch slows down the rate of subsequent additions and increases the probability of a second error occurring. The net result of this is to slow down the overall replication rate but to reduce the frequency of errors in sequences that are completed. It also increases the diversity of sequences generated (Derr et al. 2012). Speed and accuracy of primer also depend on the presence of downstream binding oligonucleotides on the other side of the base that is being added (Kervio et al. 2010; Prywes et al. 2016), and the possibility of primer extension via addition of dimers and trimers, rather than single nucleotides, has also been studied (Sosson et al. 2019).

From structure formation considerations, we showed that secondary structures with short helices are difficult to form in the AUG alphabet, because the base-pairing interactions are weaker. If all helices in the structure are quite long, then a range of structures is possible with AUG sequences that is comparable to that with AUGC. The computational folding algorithm we have used cannot tell us whether these structures are functional, but we see no reason to suppose that there could not be at least some AUG sequences with stable structures that are also functional. The shapes of these structures might be a little different from those of AUGC sequences. It is important however, that to reach a structure of a given degree of complexity (i.e., a given number of loops, branches, or domains), we need a longer sequence in the AUG case because each helix has to be longer. This poses greater problems for prebiotic synthesis, where equilibrium sequence concentration is likely to decrease exponentially with length (Higgs 2016) and for replication, where the per-base error rate has to decrease inversely with the length in order to beat the error threshold (Eigen et al. 1988).

One limitation of the above argument is that it is temperature dependent. We have only considered a temperature of 37 °C, which is the default temperature of ViennaRNA. Colder temperatures would increase the thermodynamic stability of smaller helices in the AU and AUG alphabets, and would probably make it easier to find complex structures with shorter sequences.

Accurate replication of both sequence and structure is an essential requirement for the RNA World, whatever alphabet is used. We have shown that structural fidelity can be high in cases where sequence replication is accurate, for both AUG and AUGC alphabets. Figure 5c and *d* shows mean structural fidelities for typical sequences/structures that evolve in the given alphabet. The sets of structures will be different, as is shown by the first half of this article, but the structural fidelities of the structures that are formed are comparable. In fact the average structural fidelity in the AUG case is somewhat higher than the AUGC case because the stickiness S is higher, so the likelihood of remaining compatible after a mutation is higher. We conclude that structural fidelity is not a particularly difficult problem for the AUG alphabet.

The serious issue for the AUG alphabet is that it is not possible to replicate sequences containing G's with high sequence fidelity. Thus it is not possible to use a G in a position where it is a required base. If the wobble rate *w* is low, then replication of AUG sequences leads to the exclusion of G's. Thus the AUG alphabet is really just an AU alphabet. If *w* is high, then G's get incorporated as alternatives to A's. If a purine, either A or G, is equally good in the sequence, then this is a viable possibility, and we have an RU alphabet. However, in this case the G is not really doing anything, and we might as well just have an AU alphabet. The conclusion is that G's are not much help in the RNA World unless C's are also present.

Therefore, if cytidine was not present at the time of the origin of life, G nucleobases could not play any critical role in ribozymes and life would have to start with an AU alphabet. If a source of cytidine were later discovered, for example, if life invented the pathway of cytidine synthesis, the AU/AUG alphabet could evolve into the canonical AUGC alphabet. There would be an evolutionary advantage to incorporating the additional GC pair because it would allow greater versatility of structure formation, greater diversity of sequences folding to any specified structure, and improved sequence fidelity.

The other possibility is that we are being misled by the low frequencies of C in meteorites and in equilibrium thermodynamics calculations (Pearce and Pudritz 2015). The half-life of C is much shorter than for the other bases because of deamination of C to U (Levy and Miller 1998), but it could still be very long compared with the time of replication of an RNA. Cytidine deamination is relevant only if we are relying on gradual accumulation of nucleotides from an external source, such as delivery by meteorites. In that case, the C will not accumulate. If there is a source of continued nucleotide synthesis on Earth (as there must be at some point fairly early in the history of life, or even prior to the origin of life), then C can be created alongside other nucleotides, and the RNA World can make use of it in a replication process that is must faster than the deamination. As we noted in the introduction, C to U deamination in modern organisms (Lewis et al. 2016) is not a problem, because synthesis and replication is fast. Although it is possible that the supply of organic molecules for the origin of life may have come from meteorites (Pearce et al. 2017), the results in the present article concerning the difficulties of an AUG alphabet might also suggest that there must have been a reliable and renewable abiotic source of all four nucleotides on Earth prior to the origin of life. In this case, the RNA World would have begun with four bases from the outset.

We have concentrated on the differences between the three-letter and four-letter alphabet in this article; however, there are also arguments in favor of two-letter alphabets over four-letter alphabets. Hänle and Richert (2018) showed that nonenzymatic replication of GC sequences in the presence of G and C monomers is much less error prone than replication of AUGC sequences in the presence of all four monomers. This makes sense, as the only possible errors in the two-base system are GG and CC, neither of which is a frequent kind of mismatch, whereas many other types of errors are also possible in the four-base system. However, if a sequence of only G's and C's were replicated in the presence of all four bases, the other error types would also occur. So it is not clear whether there is much advantage for using only two letters in the sequence if the diversity of available bases is greater than two. We do not have a clear idea of what kinds of nucleotides and similar molecules were available prebiotically. The mixture could well have included a very high diversity of molecules differing in both backbone unit and pairing unit. Selection of the genetic alphabet must have occurred in such a mixture, and the best set of monomers would depend on their frequency in the mixture as well as the accuracy and speed of their replication. We have previously considered simulations of nonenzymatic replication in mixtures of ribonucleotides and alternative nucleotides, showing that the system moves to a stable state in which the monomers used are almost all of one backbone type (Tupper et al. 2017). This is similar to the way that either left- or right-handed homochiral molecules emerge in a situation that is initially racemic. Returning to the issues in the current article, the two-letter GC alphabet and the two-letter AU alphabet would both be feasible from the folding point of view and from the replication point of view. For two-letter alphabets, the frequency of structures (P_{str}) with large numbers of base pairs is higher than for four-letter alphabets because the stickiness is higher (e.g., fig. 2*a*), but the number of possible sequences (N_{str}) is much lower (e.g., fig. 2*c*). The stronger pairing of GC versus AU makes it easier to find shorter sequences with complex secondary structures, which is an argument favoring GC over AU. Of course, this argument is to no avail if there was no C present in the mixture.

This article focuses on the choice of nucleotide alphabet for the RNA World and the relative instability of cytosine. Another key stability problem is the rate of replication relative to hydrolysis of strands. If RNA replication is to be maintained, then the rate of polymerization (either by a polymerase/ligase or by nonenzymatically) must be faster than the rate of hydrolysis of the backbone. So either there must be a polymerase ribozyme with a fast rate or we must find conditions where nonenzymatic replication is fast. Clearly, a much more detailed evolutionary model could be considered that incorporates hydrolysis explicitly, including the possibility of different rates of hydrolysis for nucleotides in single-stranded and helical regions. These questions are important, but they are essentially independent of the questions about the nucleotide alphabet that are studied here. The first catalysts have to arise within a chemical system that is abiotic. The question of which nucleotides are supplied abiotically is a key one. If suitable monomers are not supplied by chemistry, we need a ribozyme to make them. It does not seem reasonable to propose a scenario for the origin of life where we require two different kinds of catalysts to appear simultaneously. The simplest scenario is where the first catalysts catalyze replication, using a set of monomers that is already present abiotically. The issue of which nucleotides were available and whether sequences composed of these nucleotides can be functional is therefore important. On the other hand, if nonenzymatic replication was fast enough, then polymerase ribozymes may not have been necessary initially, and the first catalysts could have been nucleotide synthetases, as discussed in our previous work (Kim and Higgs 2016; Higgs 2019). The question would then arise as to what kind of nucleotides were used by the nonenzymatic replication system within which the synthetases evolved. This article addresses important issues concerning the choice of nucleotide alphabet that are relevant whether the initial replication process was via a polymerase ribozyme or nonenzymatic. The conclusions regarding the ability of AUG sequences to fold and the problems associated with replication in the AUG alphabet are equally relevant whether replication is ribozyme catalyzed or nonenzymatic.

Materials and Methods

Definition of Secondary Structures and Domains

A secondary structure is a set of base pairs formed within a given sequence, such that the paired nucleotides are allowed pairs in a given alphabet (as in table 1), and such that all the specified pairs can be formed at the same time. If pseudoknots are excluded, as is usually the case with RNA folding algorithms (Higgs 2000), secondary structures can be represented by nested sets of brackets. For example, the notation ((((. . ..)))) represents a hairpin loop with a stem of four-base pairs and five unpaired bases in the loop. A domain is defined as a region of structure enclosed by a base pair that is not nested within any other base pair. The structure (((.(((. . ..))).(((. . ..))).))) is a single domain containing two hairpin loops within a multibranched loop. The structure .(((. . ..))).(((. . ..))). has two separate domains, each of which is a single hairpin.

Method of Estimation of Structural Frequencies

We wish to calculate the probability P_{str} that a random sequence folds to a particular structure denoted by a given bracket notation. A minimal criterion for folding is that a sequence should be *compatible* with the defined structure, that is, there must be an allowed base pair at each pair of sites indicated by the bracket notation. If a sequence is not compatible, it cannot form the correct structure; therefore, it is only necessary to fold compatible sequences in order to calculate P_{str}. It is possible to generate compatible sequences by choosing one of the allowed base pairs at random for each pair of brackets in the structure, and choosing one single base at random for each unpaired site. This method generates all compatible sequences with equal probability, assuming that the frequency of each of the nucleotides in the alphabet is equal.

To determine P_{str}, we used the relationship $P_{str} = P_{str\ |\ comp} P_{comp}$, where P_{comp} is the probability that a random sequence is compatible, and $P_{str\ |\ comp}$ is the conditional probability that a sequence folds to the specified structure, given that it is compatible. P_{comp} can be calculated exactly from the stickiness: $P_{comp} = S^{n_{pairs}}$, where n_{pairs} is the number of specified pairs in the bracket notation. For each specified structure, we generated 10^6 compatible sequences. $P_{str\ |\ comp}$ is the fraction of these sequences that folded to the correct structure according to the ViennaRNA package. If at least one correctly folding sequence was found, our estimate of the frequency of the structure is $P_{str} = P_{str\ |\ comp}\ P_{comp}$. Where no correct structure was found, P_{str} is below our limit of detection, that is, $P_{str} < P_{limit}$, where $P_{limit} = 10^{-6} P_{comp}$. By making use of compatible sequences, it requires fewer calculations to get an accurate estimate of P_{str}. If we simply folded 10^6 random sequences without requiring them to be compatible, the detection limit would be 10^{-6}, and most structures would be undetectable in a sample of this size.

A Two-Step Model of RNA Replication

The probability matrix for a single step (from plus to minus strand) is the probability p_{ij} that base i in the template gives rise to base j in the complementary sequence. We suppose that the relative rates of addition of Watson–Crick, wobble, and mismatch bases are 1 : w : v. The rates are also proportional to the concentrations, ϕ_i, of the four nucleotides in the surrounding mixture. The p_{ij} matrix is obtained by normalizing the rates of the possible pairs for each template base so that $\sum_j p_{ij} = 1$. With row and column ordering A:U:G:C, we have

$$p = \begin{bmatrix} \dfrac{v\phi_A}{\phi_U + v(\phi_A + \phi_G + \phi_C)} & \dfrac{\phi_U}{\phi_U + v(\phi_A + \phi_G + \phi_C)} & \dfrac{v\phi_G}{\phi_U + v(\phi_A + \phi_G + \phi_C)} & \dfrac{v\phi_C}{\phi_U + v(\phi_A + \phi_G + \phi_C)} \\ \dfrac{\phi_A}{\phi_A + w\phi_G + v(\phi_U + \phi_C)} & \dfrac{v\phi_U}{\phi_A + w\phi_G + v(\phi_U + \phi_C)} & \dfrac{w\phi_G}{\phi_A + w\phi_G + v(\phi_U + \phi_C)} & \dfrac{v\phi_C}{\phi_A + w\phi_G + v(\phi_U + \phi_C)} \\ \dfrac{v\phi_A}{\phi_C + w\phi_U + v(\phi_A + \phi_G)} & \dfrac{w\phi_U}{\phi_C + w\phi_U + v(\phi_A + \phi_G)} & \dfrac{v\phi_G}{\phi_C + w\phi_U + v(\phi_A + \phi_G)} & \dfrac{\phi_C}{\phi_C + w\phi_U + v(\phi_A + \phi_G)} \\ \dfrac{v\phi_A}{\phi_G + v(\phi_A + \phi_U + \phi_C)} & \dfrac{v\phi_U}{\phi_G + v(\phi_A + \phi_U + \phi_C)} & \dfrac{\phi_G}{\phi_G + v(\phi_A + \phi_U + \phi_C)} & \dfrac{v\phi_C}{\phi_G + v(\phi_A + \phi_U + \phi_C)} \end{bmatrix}.$$

For the AUGC alphabet, if we assume equal concentrations of the four bases in the mixture of monomers, we have

$$p = \begin{pmatrix} \frac{v}{1+3v} & \frac{1}{1+3v} & \frac{v}{1+3v} & \frac{v}{1+3v} \\ \frac{1}{1+w+2v} & \frac{v}{1+w+2v} & \frac{w}{1+w+2v} & \frac{v}{1+w+2v} \\ \frac{v}{1+w+2v} & \frac{w}{1+w+2v} & \frac{v}{1+w+2v} & \frac{1}{1+w+2v} \\ \frac{v}{1+3v} & \frac{v}{1+3v} & \frac{1}{1+3v} & \frac{v}{1+3v} \end{pmatrix}.$$

For a complete two-step replication cycle, the probability that base i in the initial plus strand becomes base j in the descendant plus strand is $q_{ij} = \sum_k p_{ik}p_{kj}$. The diagonal elements q_{ii} are the fidelities of replication of each base i. In this case, we have

$$q_{AA} = q_{CC} = \frac{1+v^2}{(1+3v)(1+w+2v)} + \frac{2v^2}{(1+3v)^2},$$

$$q_{UU} = q_{GG} = \frac{1+v^2}{(1+3v)(1+w+2v)} + \frac{v^2+w^2}{(1+w+2v)^2}.$$

If sequences are repeatedly replicated under this mutational model without any selection acting, the frequencies of the bases in the sequences will converge to the steady-state frequencies f_i that satisfy $f_i = \sum_j f_j q_{ji}$. In this case, we have

$$f_A = f_C = \frac{1+3v}{4+2w+10v},$$

$$f_U = f_G = \frac{1+w+2v}{4+2w+10v}.$$

For the AUG case with equal concentrations of A, U, and G in the monomer mixture, but no C, the replication probability matrix reduces to a 3 × 3 matrix:

$$p = \begin{pmatrix} \frac{v}{1+2v} & \frac{1}{1+2v} & \frac{v}{1+2v} \\ \frac{1}{1+w+v} & \frac{v}{1+w+v} & \frac{w}{1+w+v} \\ \frac{v}{w+2v} & \frac{w}{w+2v} & \frac{v}{w+2v} \end{pmatrix}.$$

The key difference from the AUGC case is that there is no unit rate for Watson–Crick pairing in the denominator of the G row of the matrix. This makes the behavior of the model significantly different from the four-base model. The fidelities and the steady-state frequencies are

$$q_{AA} = \frac{1}{(1+2v)(1+w+v)} + \frac{v^2}{(1+2v)^2} + \frac{v^2}{(1+2v)(w+2v)},$$
$$q_{UU} = \frac{1}{(1+2v)(1+w+v)} + \frac{v^2}{(1+w+v)^2} + \frac{w^2}{(1+w+v)(w+2v)},$$
$$q_{GG} = \frac{v^2}{(w+2v)(1+2v)} + \frac{w^2}{(w+2v)(1+w+v)} + \frac{v^2}{(w+2v)^2},$$

$$f_A = \frac{1+2v}{2+2w+5v},$$
$$f_U = \frac{1+w+v}{2+2w+5v},$$
$$f_G = \frac{w+2v}{2+2w+5v}.$$

If we combine A and G bases as purines, R = A or G, in either the AUG or AUGC alphabet, then

$$f_R = f_A + f_G,$$

$$q_{RR} = f_A(q_{AA} + q_{AG}) + f_G(q_{GG} + q_{GA}).$$

If we combine U and C bases as pyrimidines, Y = U or C, in the AUGC alphabet, then

$$f_Y = f_U + f_C,$$

$$q_{YY} = f_U(q_{UU} + q_{UC}) + f_C(q_{CC} + q_{CU}).$$

Acknowledgments

This work was supported by Discovery Grant 2017-05911 from the Natural Sciences and Engineering Research Council of Canada and made use of computational resources supplied by Compute Canada.

References

Attwater J, Wochner A, Holliger P. 2013. In-ice evolution of RNA polymerase ribozyme activity. *Nat Chem*. 5(12):1011–1018.

Derr J, Manapat ML, Rajamani S, Leu K, Xulvi-Brunet R, Joseph I, Nowak MA, Chen IA. 2012. Prebiotically plausible mechanisms increase compositional diversity of nucleic acid sequences. *Nucl Acids Res*. 40(10):4711–4722.

Eigen M, McCaskill J, Schuster P. 1988. Molecular quasi-species. *J Phys Chem*. 92(24):6881–6891.

Fontana W, Konings DA, Stadler PF, Schuster P. 1993. Statistics of RNA secondary structures. *Biopolymers* 33(9):1389–1404.

Hänle E, Richert C. 2018. Enzyme-free replication with two or four bases. *Angew Chem Int Ed Engl*. 57(29):8911–8915.

Higgs PG. 2000. RNA secondary structure: physical and computational aspects. *Quart Rev Biophys*. 33(3):199–253.

Higgs PG. 2016. The effect of limited diffusion and wet-dry cycling on reversible polymerization reactions: implications for prebiotic synthesis of nucleic acids. *Life* 6(2):24.

Higgs PG. 2019. Three ways to make an RNA sequence: steps from chemistry to the RNA World. In: Kolb VM, editor. Handbook of astrobiology. Boca Raton, FL: CRC Press. pp295–408.

Higgs PG, Attwood TK. 2005. Bioinformatics and molecular evolution. Malden (MA): Blackwell Publishing. pp58–76.

Higgs PG, Lehman N. 2015. The RNA World: molecular co-operation at the origin of life. *Nat Rev Genet*. 16(1):7–17.

Hud NV, Cafferty BJ, Krishnamurthy R, Williams LD. 2013. The origin of RNA and "my grandfather's axe." *Chem Biol*. 20(4):466–474.

Johnston WK, Unrau PJ, Lawrence MS, Glasner ME, Bartel DP. 2001. RNA-catalyzed RNA polymerization: accurate and general RNA-templated primer extension. *Science* 292(5520):1319–1325.

Kervio E, Hochgesand A, Steiner UE, Richert C. 2010. Templating efficiency of naked DNA. *Proc Natl Acad Sci U S A*. 107(27):12074–12079.

Kim YE, Higgs PG. 2016. Co-operation between polymerases and nucleotide synthetases in the RNA World. *PLoS Comput Biol*. 12(11):e1005161.

Kun A, Santos M, Szathmary E. 2005. Real ribozymes suggest a relaxed error threshold. *Nat Genet.* 37(9):1008–1011.

Lazcano A, Miller SL. 1996. The origin and early evolution of life: prebiotic chemistry, the pre-RNA World, and time. *Cell* 85(6):793–798.

Leu K, Kervio E, Obermayer B, Turk-MacLeod RM, Yuan C, Luevano J-M, Chen E, Gerland U, Richert C, Chen IA. 2013. Cascade of reduced speed and accuracy after errors in enzyme-free copying of nucleic acid sequences. *J Am Chem Soc.* 135(1):354–366.

Leu K, Obermayer B, Rajamani S, Gerland U, Chen IA. 2011. The prebiotic evolutionary advantage of transferring genetic information from RNA to DNA. *Nucleic Acids Res.* 39(18):8135–8147.

Levy M, Miller SL. 1998. The stability of the RNA bases: implications for the origin of life. *Proc Natl Acad Sci U S A.* 95(14):7933–7938.

Lewis CA Jr, Crayle J, Zhou S, Swanstrom R, Wolfenden R. 2016. Cytosine deamination and the precipitous decline of spontaneous mutation during Earth's history. *Proc Natl Acad Sci U S A.* 113(29):8194–8199.

Lorenz R, Bernhart SH, Zu Siederdissen CH, Tafer H, Flamm C, Stadler PF, Hofacker IL. 2011. ViennaRNA package 2.0. *Algorithms Mol Biol.* 6(1):26.

Pearce BK, Pudritz RE. 2015. Seeding the pregenetic Earth: meteoritic abundances of nucleobases and potential reaction pathways. *Astrophys J.* 807(1):85.

Pearce BKD, Pudritz RE, Semenov DA, Henning TK. 2017. Origin of the RNA World: the fate of nucleobases in warm little ponds. *Proc Natl Acad Sci U S A.* 114(43):11327–11332.

Posada D. 2008. jModelTest: phylogenetic model averaging. *Mol Biol Evol.* 25(7):1253–1256.

Prywes N, Blain JC, Del Frate F, Szostak JW. 2016. Nonenzymatic copying of RNA templates containing all four letters is catalyzed by activated nucleotides. *eLife* 5:e17756.

Reader JS, Joyce GF. 2002. A ribozyme composed of only two different nucleotides. *Nature* 420(6917):841.

Reidys C, Forst CV, Schuster P. 2001. Replication and mutation on neutral networks. *Bull Math Biol.* 63(1):57–94.

Robertson MP, Joyce GF. 2012. The origins of the RNA World. *Cold Spring Harbor Perspect Biol.* 4(5):a003608.

Rogers J, Joyce GF. 1999. A ribozyme that lacks cytidine. *Nature* 402(6759):323.

Savill NJ, Hoyle DC, Higgs PG. 2001. RNA sequence evolution with secondary structure constraints: comparison of substitution rate models using maximum likelihood methods. *Genetics* 157(1):399–411.

Shay JA, Huynh C, Higgs PG. 2015. The origin and spread of a cooperative replicase in a prebiotic chemical system. *J Theor Biol.* 364:249–259.

Sosson N, Pfeffer D, Richert C. 2019. Enzyme-free ligation of dimers and trimers to RNA primers. *Nucleic Acids Res.* 47(8):3836–3845.

Szilagyi A, Kun A, Szathmary E. 2014. Local neutral networks help maintain inaccurately replicating ribozymes. *PLoS One* 9(10):e109987.

Takeuchi N, Hogeweg P. 2012. Evolutionary dynamics of RNA-like replicator systems: a bioinformatic approach to the origin of life. *Phys Life Rev.* 9(3):219–263.

Takeuchi N, Poorthuis PH, Hogeweg P. 2005. Phenotypic error threshold; additivity and epistasis in RNA evolution. *BMC Evol Biol.* 5(1):9.

Tupper AS, Shi K, Higgs PG. 2017. The role of templating in the emergence of RNA from the prebiotic chemical mixture. *Life* 7(4):41.

Turner DH, Mathews DH. 2010. NNDB: the nearest neighbor parameter database for predicting stability of nucleic acid secondary structure. *Nucleic Acids Res.* 38(Suppl 1):D280–D282.

Chapter 5: Rolling-circle Replication in an RNA world

The contents of this chapter are unpublished and constitute a manuscript in preparation. Paul Higgs and I both contributed to the design of the replication schemes and the writing of the manuscript.

Rolling-Circle and Strand-displacement Mechanisms for Non-enzymatic Replication in the RNA World

Abstract

It is likely that RNA replication began non-enzymatically, and that polymerases were later selected to speed up the process. We consider replication mechanisms in modern viruses and ask which of these is possible non-enzymatically, using mathematical models and experimental data found in the literature to estimate rates of RNA synthesis and replication. Replication via alternating plus and minus strands is found in some single-stranded RNA viruses. However, if this occurred non-enzymatically it would lead to double-stranded RNA that would not separate. With some form of environmental cycling, such as temperature, salinity, or pH cycling, double-stranded RNA can be melted to form single-stranded RNA, although re-annealing of existing strands would then occur much faster than synthesis of new strands. We show that re-annealing blocks this form of replication almost entirely. Other kinds of viruses synthesize linear double strands from single strands and then make new single strands from double strands via strand-displacement. This does not require environmental cycling and is not blocked by re-annealing. However, under non-enzymatic conditions, we expect the incomplete new strand to be almost always displaced by the tail end of the old strand through toehold-mediated displacement. A third kind of replication in viruses and viroids is rolling-circle replication which occurs via strand-displacement on a circular template. Rolling-circle replication does not require environmental cycling and is not prevented by toehold-mediated displacement. Rolling-circle replication is therefore expected to occur non-enzymatically and is a likely starting point for the evolution of polymerase-catalysed replication.

Introduction

Current origin of life research is founded on the idea of an RNA world and the ability of RNA polymers to act as both genes and catalysts (Robertson and Joyce 2012; Higgs and Lehman 2015). In the RNA World, it is envisaged that RNA replication was catalyzed by RNA polymerase ribozymes. Although there are no natural polymerase ribozymes in biology today, there has been continued progress in development of such ribozymes in the laboratory (Attwater et al. 2013; Horning and Joyce 2016; Attwater et al. 2018; Wachowius and Holliger 2019). The best polymerases to date are able to synthesize strands of length ~200 nucleotides, although they are still not at the level of self-replication. Before the origin of polymerases, it seems likely that non-enzymatic template-directed replication was occurring, at least for oligomer sequences, and eventually for strands that were long enough to be able to function as ribozymes. Non-enzymatic replication would generate a mixture of short sequences that functioned as templates without necessarily having any other encoded function. Natural selection would operate on the physicochemical properties of such sequences, such as melting temperatures and stabilities against hydrolysis or UV damage. We have termed this phase Chemical Evolution (Higgs 2017), emphasizing that this is an important step between random chemical synthesis without replication, and fully-fledged Biological Evolution of long strands with a function, such as being a ribozyme. If non-enzymatic replication were occurring, then strands arising in the mixture that had polymerase ability could be selected to improve the efficiency of the non-enzymatic replication process that was already there.

Non-enzymatic replication of nucleic acid polymers in the laboratory has been observed for very short templates containing 4 and 6 bases (von Kiedrowski 1986; Zielinski and Orgel 1987; Achilles and von Kiedrowski 1993; Sievers and Von Kiedrowski 1994), and 24-mers formed by ligation of 12-mer oligomers has also been observed (Edeleva et al. 2019). Single-step synthesis of much longer templates formed by ligation of several 32-mers has also been achieved in the lab (He et al. 2017; He et al. 2019). Naturally occurring ribozymes contain 40 - 2600 bases (Doudna and Cech 2002), and polymerase ribozymes contain ~150-200 bases (Attwater et al. 2013; Horning and Joyce 2016; Attwater et al. 2018; Wachowius and Holliger 2019). Repeated replication of strands of this length has not yet been observed in cases where synthesis occurs one nucleotide at a time.

The focus of this paper is to understand the mechanism by which non-enzymatic replication might have occurred in the RNA World. Important clues are available by examining the mechanisms of protein-catalyzed RNA replication that occur in modern RNA viruses and viroids. We will consider which of these mechanisms might have occurred non-enzymatically prior to the origin of ribozymes and protein catalysts. If the catalyzed mechanism is possible non-enzymatically as well, then a smooth transition from non-enzymatic to catalyzed replication could have occurred.

The conceptually-simplest form of RNA replication, shown as Scheme I in Figure 1, is via alternating plus and minus strands, where each strand is a template for the other. Some RNA viruses, including the well-studied Qβ bacteriophage, do replicate in this way. Importantly, the Qβ polymerase protein has a mechanism of separating the new strand from the template while it is being synthesized (Takeshita and Tomita 2012). The two strands leave the protein polymerase through separate channels, thus avoiding the formation of double stranded RNA. We have included in Scheme I the possibility of annealing of two single strands to form a double strand. The double strand is a dead end that cannot be replicated in this scheme. In our calculations below, we show that there is a critical strand concentration C^* at which annealing becomes dominant. This reaction scheme gives exponential growth for strand concentrations below C^* and replication is inhibited when C^* is reached.

Although Scheme I works for some viruses, it would be impossible in the absence of a polymerase because a double stranded helix would be bound to form during the template-directed strand synthesis. Therefore, we consider Scheme II, where the synthesis step leads to double stranded RNA formation, and a melting step is required to separate the strands. Long double strands are extremely stable; hence the separation of the strands will hardly ever occur if the reaction conditions are constant (based on the kinetics of annealing from Rauzan et al. 2013 and stacking energies from Xia et al. 1998). This is commonly referred to as the product inhibition problem of RNA replication (Szostak 2012). To overcome the stability of double stranded RNA, many groups have proposed some form of environmental cycling, such as temperature (Kreysing et al. 2015; He et al. 2017; Edeleva et al. 2019; He et al. 2019), salinity (Lathe 2004; Lathe 2005; Ianeselli et al. 2019), or pH cycling (Mariani et al. 2018) to drive the melting of double stranded RNA and allow for continuous RNA replication. In the synthesis phase, the single strands are used as templates to form new double strands. In the subsequent melting phase, the double strands are forced to melt by changing the reaction conditions, thereby freeing up the single strand templates. By cycling the reaction conditions between the two phases, replication of RNA polymers may be possible. We consider the kinetics of Scheme II in this paper, both at constant reaction conditions, and with cycling. Exponential replication is possible if cycling occurs, but only for low strand concentrations $C < C^*$. Once C^* is reached, reannealing of existing strands occurs faster than the synthesis of new strands. Below, we estimate that reannealing results in product inhibition at an extremely low strand concentration.

An alternative approach to non-enzymatic RNA replication is to consider the possibility of a strand displacement mechanism, in which a double strand is used as a template to synthesize a new single strand. By coupling the synthesis of a new strand to the displacement of an old strand, the product inhibition problem is avoided (Bartel 1999; Müller 2006; Cheng and Unrau 2010; Zhou et al. 2019). Strand displacement occurs in both ssRNA and dsRNA viruses (Hulo et al. 2011). In ssRNA viruses, the virus capsids contain single strands, but these are converted to double strands inside the host cell, and strand-displacement then creates further single strands, as in Figure 1, Scheme III. In this paper, we consider the possibility of non-enzymatic replication via strand-displacement, and we argue that this is more likely than non-enzymatic replication via melting (Scheme II). We note that in Scheme III, melting is not required; hence Scheme III is possible at constant reaction conditions and does not require cycling. Reannealing is still relevant in Scheme III, but we show that Scheme III allows exponential replication even when reannealing is very fast. In this latter case, single strands made by strand-displacement are converted almost immediately to new double strands, but this does not inhibit replication because continued synthesis is possible from the double strands.

However, non-enzymatic replication via strand-displacement faces one further problem that we have not yet considered. If a new complementary strand is partially complete, as in Figure 2(a), it is possible for the growing strand to come off the template, at which point the two tails can move in either direction rather easily because there is little energetic barrier to overcome (Radding et al. 1977; Green and Tibbetts 1981; Srinivas et al. 2013). If the branch point migrates to the left, the growing strand is lost. This is known as toehold-mediated displacement. This is not a problem during virus replication because the virus polymerase ensures the new strand remains bound to the template. Whereas, in recent experiments on non-enzymatic strand-displacement (Zhou et al. 2019), toehold-mediated displacement is prevented by using a tail strand which is non-complementary to the template. However, if there is no mechanism of preventing toehold-mediated displacement, our calculations below show that strand synthesis through displacement will almost always be prevented before the new strand is complete.

At this point, we take one further clue from looking at virus replication mechanisms. The rolling-circle mechanism shown in Figure 2(b), is common to viroids, satellite RNAs, and some viruses (Flores et al. 2011; Flores et al. 2014). The mechanism of primer extension is the same as in strand displacement, except for the fact that it occurs on a circular template. Synthesis of the complementary strand can continue indefinitely, and new product strands are created when the tail is cleaved. Cleavage would occur eventually by random hydrolysis, however, some viroids, satellite RNAs, and viruses replicating by this mechanism contain self-cleaving ribozyme sequences that ensure cleavage occurs rapidly at a specific position. The linear strand formed after cleavage is then able to circularize and begin a new round of rolling circle replication. The key difference between linear strand-displacement and rolling-circle replication is that flipping of the two tails in the circular case does not lead to the loss of the growing strand (Figure 2b).

While both linear strand-displacement and rolling-circle mechanisms of RNA replication are clearly viable in modern viruses and viroids which utilize protein polymerases, they have received relatively little attention under non-enzymatic conditions. The aim of this paper is to show (i) that some form of strand-displacement mechanism is likely to be more viable in the RNA World than a mechanism relying on melting, and (ii) that the rolling-circle mechanism is likely to be more viable than strand-displacement on a linear template. As the rolling-circle mechanism should be possible non-enzymatically, a smooth pathway of evolution from chemical evolution to the RNA world is possible by selecting for polymerase sequences that increase the rate of the non-enzymatic rolling-circle mechanism, without requiring a change in the mechanism.

Results

Comparison of Three Reaction Schemes for RNA Replication

In this section, we give simple rate equations for the three reaction schemes in Figure 1. R_{syn} is the rate of synthesis of a complementary strand from a single strand. In Scheme I, the complementary strand is assumed to be separated immediately. In Schemes II and III, the complementary strand forms a double strand with the template. R_{dis} is the rate of strand synthesis via strand-displacement, in which case the old complementary strand becomes a single strand. R_{melt} is the rate of melting of double strands to single strands, and R_{ann} is the rate of annealing of complementary plus and minus strands to form double strands. C_P and C_M are the concentration of plus and minus single strands, and C_D is the concentration of double strands.

Scheme I:
$$\frac{dC_P}{dt} = R_{syn}C_M - R_{ann}C_PC_M$$
$$\frac{dC_M}{dt} = R_{syn}C_P - R_{ann}C_PC_M$$
$$\frac{dC_D}{dt} = R_{ann}C_PC_M.$$

It is clear that the single strands grow exponentially in proportion to $\exp(R_{syn}t)$ if annealing is negligible. However, there is a concentration of single strands $C^* = \frac{R_{syn}}{R_{ann}}$ where the annealing term balances the synthesis. Growth of single strands is blocked entirely by annealing when the concentration approaches this limit. If both plus and minus strand concentrations begin at C_{init}, the exact solution of the above equations is

$$C_P = C_M = \frac{C^* b\exp(R_{syn}t)}{1+b\exp(R_{syn}t)},$$

where $\beta = \frac{C_{init}/C^*}{1-C_{init}/C^*}$. The concentration of single strands tends to the limit of $C = C^*$ at long times. For today's viruses, synthesis is rapid and annealing is unlikely to be a problem because strands are always infecting new hosts. However, for prebiotic non-enzymatic replication, synthesis would be slow, and annealing would be very fast relative to synthesis, meaning that C^* would be exceedingly small. We give numerical examples later in the paper in the section on Rate Estimates.

We now consider Scheme II, where synthesis leads to formation of a double strand. It is necessary to include melting here because this is the only way that single strands can be reformed.

Scheme II: $\frac{dC_P}{dt} = -R_{syn}C_P + R_{melt}C_D - R_{ann}C_P C_M$

$$\frac{dC_M}{dt} = -R_{syn}C_M + R_{melt}C_D - R_{ann}C_P C_M$$

$$\frac{dC_D}{dt} = R_{syn}(C_P + C_M) - R_{melt}C_D + R_{ann}C_P C_M$$

If there is no annealing ($R_{ann} = 0$) both single and double strands grow exponentially as $\exp(l_{melt}t)$ where

$$l_{melt} = \frac{(R_{syn}+R_{melt})}{2}\left(-1 + \sqrt{1 + \frac{4R_{syn}R_{melt}}{(R_{syn}+R_{melt})^2}}\right).$$

If $R_{melt} \ll R_{syn}$, then $l_{melt} \approx R_{melt}$, *i.e.* the growth rate is limited by melting. In reality, melting will be extremely slow because the reaction conditions must favor stable helix formation in order to get the templating reaction to work (a numerical example is given below). Hence, we expect non-enzymatic replication by Scheme II to be impossible in constant reaction conditions because of slow melting, even if there is no annealing. If annealing is added too, the situation is even worse, because annealing will become dominant for a very low concentration C^*, as in Scheme I. If the reaction conditions cycle through synthesis and melting phases, as discussed in the Introduction, then Scheme II is more viable than with constant conditions, although annealing limits strand concentrations to very small values. A numerical example of Scheme II with temperature cycling is given later in this paper.

We now consider Scheme III, in which replication can occur from double strands via strand-displacement. In this case, we ignore melting because it is extremely slow and it is not required for exponential growth if strand displacement is possible.

Scheme III: $\frac{dC_P}{dt} = -R_{syn}C_P + \frac{1}{2}R_{dis}C_D - R_{ann}C_P C_M$

$$\frac{dC_M}{dt} = -R_{syn}C_M + \frac{1}{2}R_{dis}C_D - R_{ann}C_P C_M$$

$$\frac{dC_D}{dt} = R_{syn}(C_P + C_M) + R_{ann}C_P C_M$$

If annealing is negligible ($R_{ann} = 0$) the single and double strand concentrations grow exponentially as $\exp(\lambda_{dis}t)$, where

$$\lambda_{dis} = \frac{R_{syn}}{2}\left(-1 + \sqrt{1 + \frac{4R_{dis}}{R_{syn}}}\right).$$

During this exponential growth, the two single-strand concentrations are equal, and are proportional to the double strand concentration:

$$\frac{C_P}{C_D} = \frac{\frac{1}{2}R_{dis}}{\lambda_{dis}+R_{syn}} \approx \frac{R_{dis}}{2R_{syn}}.$$

If the annealing rate is non-zero, then there comes a point where the rate of production of double strands by annealing is equal to that by synthesis: $R_{syn}C_P \approx R_{ann}C_P^2$. This occurs when $C_P \approx C^* = \frac{R_{syn}}{R_{ann}}$ and $C_D \approx 2R_{syn}^2/(R_{ann}R_{dis})$. When the concentrations are greater than this, annealing dominates, and the single strands are converted almost immediately to double strands. However, even in this limit, double strands continue to grow exponentially in proportion to $exp(1/2R_{dis}t)$. In other words, exponential growth of double strands is still possible even when annealing is fast, and even when the concentration of strands becomes larger than C^*. For higher concentrations, single strands become negligible in comparison to double strands, but they still increase.

Rates of Template-Directed RNA Strand Synthesis

In order to give quantitative examples of the behavior of these reaction schemes, we need to consider the processes at a more detailed level in order to determine the expected rates of strand synthesis. As defined previously, R_{syn} is the rate of synthesis of a complementary strand on a single-stranded template. We will assume that synthesis starts by annealing of a primer, which is then repeatedly extended by single nucleotides (monomers) to produce a new complementary strand (Fig. 3).

If the primer is of length l and the total template length is $l + L$, such that L nucleotides must be added to complete the strand, then the mean time for synthesis of the complementary strand is $1/R_{syn} = 1/k_{on} + L/k_{ext}$, where k_{on} is the net rate of primer binding and k_{ext} is the rate of primer extension. For simplicity, we will assume throughout this discussion that the annealing of primers is not rate limiting, such that $R_{syn} \approx k_{ext}/L$. If primers were rare prebiotically, then the net rate of synthesis would be even slower. We are therefore considering a best-case scenario, which is still problematic due to product inhibition.

In the introduction to this paper, we argued that replication from circular templates via the rolling-circle mechanism may be important in the RNA World. Therefore, we also wish to estimate R_{syn} for circular templates. We expect the process of primer extension to be very similar for circular and linear templates. For primer binding, there may be a slight advantage for the circular template because a primer could bind anywhere on the circle rather than just at the beginning of the linear strand. However, if a primer could be extended in both directions, then a primer could also bind in the middle of a linear strand. In any case, we are assuming that primer extension is the limiting factor, not primer binding. Thus, we conclude that $R_{syn} \approx k_{ext}/L$ for circular templates, equal to the rate for linear templates of the same length.

We now calculate the rate, R_{dis}, of synthesizing single strands via strand-displacement. In this case, there is a significant difference between linear and circular templates, as shown in Fig. 4. We will summarize the key points here and provide the details of the calculation in the Materials and Methods section. We assume that a primer can bind to the end of one of the strands of the duplex at a net rate of k'_{on} due to partial unzipping of the duplex, and that this is not the rate limiting step. We assume that there is a rate k'_{ext} of extension of the primer by one nucleotide whilst the old complementary strand is being displaced. This rate is expected to be slower than the rate k_{ext} of primer extension on a single strand, but not so slow as to prevent synthesis entirely. During strand displacement from a linear strand, it is possible for the two tails to flip as in Fig 2(a), in which case the growing strand is lost, and synthesis has to begin again from a new primer. This tail flipping occurs at a rate of k_{flip}, which is expected to be fast compared to k'_{ext}. The probability that one monomer is added before the flipping occurs is $P_{ext} = \left(\frac{k'_{ext}}{k'_{ext}+k_{flip}}\right)$. Tail flipping can occur at any of the intermediate states of synthesis of the new strand (as in Figure 4). For complete synthesis of the growing strand, there must be L extensions before the flipping occurs once. The rate of complete strand displacement is therefore proportional to P^{L}_{ext}. The calculation in the Materials and Methods section shows that $R_{dis} \approx k'_{on}P^{L}_{ext}$. As we expect $P_{ext} \ll 1$, we expect R_{dis} to be vanishingly small for linear templates more than just a few nucleotides long unless there is a polymerase, or some other mechanism, that prevents the tail flipping.

The situation is more optimistic for strand-displacement from a circular template, as occurs in the rolling-circle mechanism. For starters, no primer annealing step is required as one (or both) ends of the annealed strand could act as a primer. Here, we estimate that $R_{dis} = \frac{k'_{ext}}{2L}$, where the factor of 2 comes from the assumption that growth occurs from one end of the strand, and that the growing end is annealed on the template roughly half the time. Since the rate of synthesis is independent of tail flipping, the rolling-circle mechanism is more likely to occur non-enzymatically than strand-displacement on a linear template. For rolling-circle, the rate of synthesis is therefore inversely proportion to length, the same as in the case of strand synthesis on a single strand template.

Rate Estimates and Numerical Examples of RNA Strand Replication

In this section, we use experimentally measured rate constants from the literature to estimate the rates of the processes in the three reaction schemes. We then show numerical examples of concentrations as a function of time in several cases.

The rate constant for RNA primer annealing, k_{pri}, has been measured to be in the range $2 - 24\ mM^{-1}\ s^{-1}$ for tetramers to octamers, depending on sequence and temperature (Craig et al. 1971; Rauzan et al. 2013). Similar values of $0.3 - 11\ mM^{-1}\ s^{-1}$ have been also been reported for DNA hexamers (Williams et al. 1989). This is roughly $10^{10}\ M^{-1}\ h^{-1}$. Assuming a primer concentration, C_{pri}, of $10^{-6}\ M$, we would therefore expect $k_{on} \approx 10^{4}\ h^{-1}$, meaning that annealing of a primer would take on average less than a second to complete. As long as the resulting duplex is sufficiently long, the primer is stable on the timescale of RNA synthesis.

In comparison, the rates of primer extension observed by Leu et al. 2011 and Walton et al. 2019 using activated nucleotides are around 1 h^{-1} and 20 h^{-1} respectively at mM concentrations of nucleotides. For the examples here, we will take $k_{ext} \approx 1\ h^{-1}$ as an estimate of what might be achieved by non-enzymatic replication under prebiotic conditions. This is much slower than the rate of primer annealing, so our simplification that synthesis is rate limited by extension appears justified. For comparison, we will also consider overly optimistic rates of non-enzymatic extension based on known ribozymes and enzymes. As of now, the fastest ribozyme polymerase extends a primer at a rate of roughly $72\ h^{-1}$ (Horning and Joyce 2016), which for convenience, we will round up to $100\ h^{-1}$. Whereas protein polymerases can extend a primer multiple times per second, which will be taken as roughly $10{,}000\ h^{-1}$ (Schwartz and Quake 2009; Olsen et al. 2013). In the examples here, we will consider a sequence of length $L = 100$ nucleotides; therefore, the rate of strand synthesis on a single strand template under non-enzymatic conditions is $R_{syn} \approx \frac{k_{ext}}{L} \approx 0.01\ h^{-1}$.

The annealing rate constant for octamers and shorter oligomers was estimated above to be $10^{10}\ M^{-1}\ h^{-1}$. This rate is known to increase slowly with sequence length, appearing to scale as the square-root of polymer length (Wetmur and Davidson 1968). For a sequence of length 100, the annealing rate constant should be at least $10^{10}\ M^{-1}\ h^{-1}$. The single strand concentration at which annealing dominates strand synthesis is therefore $C^* = R_{syn}/R_{ann} = 10^{-12}\ M$. Thus, we expect annealing to prevent replication in scheme I at a very low strand concentration. Figure 5 shows the single strand concentration as a function of time, with the estimated values of $R_{syn} = 0.01\ h^{-1}$ and $R_{ann} = 10^{10}\ M^{-1} h^{-1}$ under non-enzymatic conditions. For comparison, we also show the optimistic cases of R_{syn} based on the extension rates of known ribozyme and protein polymerases. Even when non-enzymatic synthesis is as fast as a modern protein polymerase, the strand concentrations still saturate at very low values. For reference, the volume of a typical bacteria cell is roughly 1 μm^3; therefore a concentration of 1 strand per cell is roughly $10^{-9}\ M$. The critical concentration, C^*, is therefore unreasonably small based on the experimentally measured annealing and extension rates. Thus, annealing is a serious inhibiting factor on replication by Scheme I.

It should be remembered that Scheme I is only possible if there is a mechanism of separation of strands at the same time as synthesis, and this seems unlikely for non-enzymatic replication. Therefore, it is necessary to consider Scheme II, where synthesis leads to double-strand formation. Replication is then dependent on double-strand melting, which is predicted to be very slow for long double strands. For example, at $37^{o}C$ an alternating AU sequence of 40 nucleotides would anneal to its complement with a standard free energy of $\Delta G^{o}_{helix} \approx -40\ kcal/mol$ (Turner and Mathews 2009). Assuming an annealing rate of $10^{10} M\ h^{-1}$, and a thermal energy RT of $0.616\ kcal\ M^{-1}$ at $37°C$, we would therefore predict a melting rate of $k_{off} \approx 10^{10} \exp\left(-\frac{\Delta G^{o}_{helix}}{RT}\right) h^{-1} \approx 6 * 10^{-19} h^{-1}$. In other words, the average time required to melt this duplex would be $\approx 10^{18}$ hours, which is much longer than the current age of the universe at $\approx 10^{14}$ hours. Longer sequences would take even longer to melt and adding GC pairs would only make the helix more stable and even slower to melt. For a sequence of length 100, the melting rate would be vanishingly small if the reaction conditions were held constant. In order for Scheme II to be viable, there needs to be cycling between a synthesis phase and a melting phase.

The cycling case of Scheme II can be solved exactly by assuming that no melting is possible in the synthesis phase, and that no synthesis is possible in the melting phase. Suppose the concentration of double strands at the end of one synthesis phase is C_{end} and the synthesis has gone to completion, such that no single strands remain. In the melting phase, these double strands are rapidly separated, so the concentration of each of the two single strands after melting is C_{end}. Thus, the initial value of the single strand concentrations at the beginning of a synthesis phase, which we call C_{init}, are equal to the final value of the double strand concentration at the end of the previous synthesis phase. The differential equations for the synthesis phase are:

Scheme II (synthesis phase): $\frac{dC_P}{dt} = -R_{syn}C_P - R_{ann}C_PC_M$

$$\frac{dC_M}{dt} = -R_{syn}C_M - R_{ann}C_PC_M$$

$$\frac{dC_D}{dt} = R_{syn}(C_P + C_M) + R_{ann}C_PC_M$$

These have an exact solution:

$$C_P(t) = C_M(t) = \frac{C^*\alpha\exp(-R_{syn}t)}{1-\alpha\exp(-R_{syn}t)}$$

$$C_D(t) = C^*\left(\frac{1}{1-\alpha} - \frac{1}{1-\alpha\exp(-R_{syn}t)} + \ln\left(\frac{1-\alpha\exp(-R_{syn}t)}{1-\alpha}\right)\right)$$

where $\alpha = \frac{C_{init}/C^*}{1+C_{init}/C^*}$. At the end of the synthesis phase, the double strand concentration is

$$C_{end} = C^*\left(\frac{\alpha}{1-\alpha} + \ln(\frac{1}{1-\alpha})\right) = C_{init} + C^*\ln(1 + \frac{C_{init}}{C^*}).$$

If the initial concentration $C_{init} \ll C^*$, then $C_{end} \approx 2C_{init}$, *i.e.* the concentration doubles each cycle at low concentrations where annealing is not important. If the initial concentration $C_{init} \gg C^*$, then annealing dominates, and most single stranded polymers find their complement to form double stranded polymers. However, as time progresses within a cycle, the single strand concentration eventually drops below C^* and then synthesis dominates. From the previous equation, we find that above the critical concentration, $C_{end} \approx C_{init} + C^*\ln(C_{init}/C^*)$, *i.e.* the concentration each cycle increases by roughly C^* due to the log scaling on C_{init}. For example, when $(C_{init}/C^*) = 10^3$, $C_{end} \approx C_{init} + 7C^*$. Increasing (C_{init}/C^*) to 10^6 results in only marginally increased strand growth: $C_{end} \approx C_{init} + 14C^*$.

Figure 6 shows the double strand concentration as a function of time for a single cycle assuming the three different R_{syn} rates for non-enzymatic, ribozyme, and protein catalysis. The initial concentration is $C_{init} = 10^{-9}M$, which corresponds to roughly one strand in the volume of a bacterial cell. For the fastest synthesis rate, the concentration almost doubles by the end of the cycle. For the non-enzymatic synthesis rate, there is almost no increase since annealing dominates.

In order to show the expected increase in double strand concentration over many cycles, we simply plot a point at the concentration C_{end} at the end of each cycle (see Figure 7). We begin at a very low concentration in order to illustrate the possibility of exponential growth that occurs when $C_{init} \ll C^*$. When $C_{init} \sim C^*$, the curve switches to slower than exponential growth. Hence replication by this kind of cycling becomes extremely slow once the concentration reaches C^*. Note that the time in Fig 7 is in number of cycles, not in hours. In order to allow the synthesis phase to go to completion, the length of the phase must be long compared to the inverse of the synthesis rate. For example, assuming the non-enzymatic synthesis rate of $R_{syn} = 0.01\ h^{-1}$, a cycle time in excess of $100\ h^{-1}$ is required for synthesis to complete. As long as the synthesis phase is sufficiently long, the multiplication factor achieved per cycle is independent of the cycle time. If the cycle time is faster (*e.g.* a daily cycle), then the reaction does not go to completion with this choice of R_{syn}, and the multiplication factor achieved per cycle is less. The case shown in Fig. 7 is a best possible case, and our conclusion is that cycling is not very effective even in the best case.

Lastly, we will consider Scheme III and show that it allows continued exponential growth at high strand concentrations, in contrast to the other schemes, which are all severely limited by annealing. For Scheme III, we need to estimate the rate of synthesis with strand displacement. From the arguments in Materials and Methods, we estimated $k'_{ext} \approx k_{ext}/K_{stack}$. For strand displacement from rolling-circle, we estimated $R_{dis} \approx k'_{ext}/2L$, which means that $R_{dis} \approx R_{syn}/2K_{stack}$. Stacking free energies depend on base sequence, but typical values would give $2K_{stack} \approx 100$. We have used $R_{syn} = 0.01\ h^{-1}$ in the previous non-enzymatic examples, which means that $R_{dis} = 10^{-4}\ h^{-1}$. Figure 8 shows exponential replication with scheme III with the three different values of R_{syn} and assuming $R_{dis} = R_{syn}/100$. Even in the non-enzymatic case, exponential growth is observed even at micromolar to millimolar concentrations. Whereas faster synthesis rates simply change the time scale for replication. At low concentrations, growth is exponential at rate λ_{dis}. At concentrations above C^*, where annealing is rapid, growth is still exponential at a slower rate $R_{dis}/2$. The essential point is that when strand displacement occurs, exponential growth is still possible even when annealing dominates. The growth rate is slower at high concentrations because $R_{dis}/2 < \lambda_{dis}$. In fact, when $R_{dis} \ll R_{syn}$, we can show that $\lambda_{dis} \approx R_{dis}$. Thus, the effect of annealing in this scheme is simply to slow growth by a factor of two, rather than to inhibit exponential growth altogether, as in the previous schemes.

Discussion

In this paper we considered several mechanisms of RNA replication and asked which of these could occur under non-enzymatic conditions relevant to the origin of life. Based on currently known reaction rates, the rolling-circle mechanism common to viroids is uniquely qualified for non-enzymatic replication in the RNA world. We are not the first to consider the importance of viroids, rolling-circle replication, or circular RNA's at the origin of life. Viroids have already been argued to be important to the RNA world due to their simplicity and encoding of ribozymes (Diener 1989; Flores et al. 2014; Diener 2016). Rolling-circle replication has also been considered in theoretical models of polymerase ribozyme replication in an RNA world (Ma et al. 2013). And circular RNA chromosomes have been considered an important bridge between the RNA world and DNA world (Soslau 2018). In this paper, we build on these ideas and argue that rolling-circle replication should be possible under non-enzymatic conditions, whereas the other mechanisms fail due to product inhibition and toehold-mediated displacement. As such, the rolling-circle mechanism is a likely starting place for non-enzymatic replication of RNA which predated the emergence of polymerase ribozymes.

The rolling-circle mechanism described here relies on primer extension coupled to strand-displacement to synthesize a new RNA polymer. Although this is expected to be slower than synthesis on a single strand without strand displacement, *i.e.* $R_{dis} \ll R_{syn}$, it is not expected to be so slow as to prevent synthesis entirely. In this paper, we argued based on thermodynamics that it may be ~100 times slower. Recent experiments show that with the addition of oligomers, the rate of primer extension with displacement on a double strand can be comparable to primer extension on a single strand (Zhou et al. 2019), *i.e.* $R_{dis} \approx R_{syn}$. In which case, our assumption here may be overly pessimistic, with the addition of oligomers further bolstering the exponential rate of replication. Furthermore, when comparing the rates of synthesis, it is important to consider the rate of template hydrolysis. The hydrolysis rate of a double stranded RNA template is predicted to be roughly 1000x slower than the hydrolysis of a single stranded RNA template (Rohatgi et al. 1996; Soukup and Breaker 1999), which may also compensate for the slower rate of RNA synthesis.

For rolling-circle replication to begin in an RNA world there needs to be a process which is generating circular templates, and the template itself must encode for a self-cleaving ribozyme. Rolling-circle replication therefore has more prerequisites than the other mechanisms described in this paper. This, however, does not make it an unlikely route for the origin of life. For starters, any process which is randomly polymerizing RNA monomers into single stranded RNA or double stranded RNA polymers will inevitably generate circular polymers. The ligation of the 3' end of an RNA polymer to the 5' end of the same polymer is an intramolecular reaction as opposed to polymerization which is an intermolecular reaction. Since the former is concentration independent, it likely occurs at a much faster rate than polymerization. Furthermore, the circular templates required for rolling-circle replication need not be exceedingly long. Viroids and satellite RNAs have typical sizes of ~220-400 bases (Flores et al. 1997; Collins et al. 1998; Flores et al. 2012). Whereas rolling-circle synthesis of ssDNA has been shown for much shorter templates containing 13-74 bases (Fire and Xu 1995; Liu et al. 1996; Frieden et al. 1999). Circular templates capable of undergoing rolling-circle replication may therefore be abundant under the conditions relevant to the origin of life

In our calculation of rolling circle replication, we assumed that there is a self-cleaving ribozyme that cuts the growing strand at a particular bond. Due to the abundance of these ribozymes, this does not appear to be restrictive. Self-cleaving ribozymes are already found in viroids and some viruses (Ferré-D'Amaré and Scott 2010), and have been detected in the genomes from every domain of life (Perreault et al. 2011; Hammann et al. 2012). Since these ribozymes have emerged at least a few times in natural evolution, their emergence in an RNA world seems plausible. Self-cleavage ribozymes also tend to be very short sequences and are thus likely to be generated within a pool of chemically synthesized random sequences. For instance, the hammerhead ribozyme is ~50 nucleotides long and has relatively few restrictions on catalytic activity (Ruffner et al. 1990). The required level of catalytic activity is also not restrictive. To be effective, the self-cleaving ribozyme needs to operate on a timescale which is faster than hydrolysis, such that cleavage consistently occurs in a single location. Assuming each bond in single stranded RNA hydrolyzes at a rate of $\sim 10^{-7}\ h^{-1}$ (Li and Breaker 1999), a polymer of 100 bases would break at a net rate of $\sim 10^{-5}\ h^{-1}$. In comparison, hammerhead ribozymes cleave at a rate of $\sim 10^{2}\ h^{-1}$ (Hertel et al. 1994), a rate which is $\sim 10^{7}$-fold faster than necessary. Even a poor hammerhead ribozyme is likely sufficient at this stage of the origin of life.

A further benefit of the rolling-circle mechanism is that it becomes primer-independent at high strand concentrations. When annealing of strands is fast, as is expected under conditions which favor RNA synthesis, single-stranded RNA is immediately converted to double stranded RNA which can undergo further rolling-circle synthesis without the need of a primer. This may be of great importance to the origin of life as primers are expected to be rare under prebiotically relevant conditions. For instance, if we assume that random polymerization of RNA monomers is generating 10-mers at a concentration of $1\mu M$, then the concentration of a sequence-specific 10-mer primer would be $\sim 1pM$ as there are 4^{10} possible 10-mers. Primers would also be very slow to regenerate from random polymerization when consumed, which would severely limit the rate of replication. Further compounding the problem is that primer-length polymers are incompatible with known protocells as the lipid bilayer prevents diffusion of oligomers across the membrane (Mansy and Szostak 2009; Szostak 2012). Therefore, even if the environment is providing an abundance of sequence specific primers, they are unable to pass into a protocell. For the rolling-circle mechanism, these problems are avoided, and replication can proceed with only monomers as a food source. Since monomers are able to diffuse into and out of protocells, rolling-circle replication is also compatible with protocells.

Based on the findings in this paper, as well as the supporting literature, we propose that the first non-enzymatic RNA replication occurred through a rolling-circle mechanism. The properties of which should allow for exponential growth of RNA polymers under plausible prebiotic conditions while avoiding the problems of product inhibition, toehold-mediated displacement, and primer dependence. Given sufficient time, a polymer replicating by this mechanism could explore sequence space and discover more complex ribozymes. The discovery of polymerase ribozymes would enhance the rate of replication without necessarily changing the mechanism of replication.

Materials and Methods

RNA synthesis by strand-displacement on a linear double strand

Details of the calculation of the strand displacement rate R_{dis} are given in this section. When primer extension occurs by strand displacement, one nucleotide must be peeled back from the displaced strand for every new monomer that is added. If the equilibrium constant for pealing back a single nucleotide is K_{stack}^{-1}, then the net rate of primer extension would be:

$$k_{ext}' = \left(\frac{C_{nucl}K_{nucl}/K_{stack}}{1 + C_{nucl}K_{nucl}/K_{stack}}\right)k_{lig}$$

We note that the thermodynamic association constant for monomer binding, K_{nucl}, has been measured to be somewhere between $2\,M^{-1}$ and $66\,M^{-1}$ depending on the nucleobase for a monomer bound to the end of primer (Izgu *et al*., 2015). In this case, we have $C_{nucl}K_{nucl}/K_{stack} \ll 1$, as long as the monomer concentration is not too high. Hence $k'_{ext} \approx C_{nucl}K_{nucl}k_{lig}/K_{stack}$, whereas the extension rate on a single strand would be $k_{ext} \approx C_{nucl}K_{nucl}k_{lig}$. Hence $k_{ext}' \approx k_{ext}/K_{stack}$.

During primer extension, it is possible for the tail of the displaced strand to flip up, as in Figure 2(a) which can lead to the removal of the growing strand. In the first step, a branch point is formed. After which, the branch point migrates left or right through a random walk process. If the random walk reaches the left boundary before the right, then the primer is removed. If the random walk reaches the right boundary before the left, then the primer is not removed during this branch formation. However subsequent branch formation events can still result in primer removal. We suppose that tail flipping results in primer removal at a net rate of k_{flip}. Based on the biophysical analysis of the process (Srinivas et al. 2013), the net rate of tail flipping is roughly:

$$k_{flip} = \frac{1}{l}\left(\frac{1}{k_{first}} + \frac{l-1}{k_{bm}}\right)^{-1}$$

Where k_{first} is the rate at which the branch point forms, k_{bm} is the rate of branch migration, and l is the length of the primer. Based on experimentally measured rates, $k_{first} \cong 1.1 * 10^{6}\ h^{-1}$ and $k_{bm} \cong 6.5 * 10^{7}\ h^{-1}$ (AEL model of Srinivas et al. 2013). For a primer of length 20, $k_{flip} \approx 4 * 10^{4}\ h^{-1}$. Increasing the primer length to 100 would decrease k_{flip} to roughly $4 * 10^{3}\ h^{-1}$. For simplicity, we assume that the net rate of flipping is constant throughout synthesis.

The linear template can exist in several states as shown in Figure 4. In state L, with concentration C_L, there is a fully synthesized complementary strand and no primer. In state 0, with concentration C_0, a primer is bound, but there is not yet any monomer extension. In state *n*, with concentration C_n, the primer is bound and has been extended by *n* nucleotides, where $1 \leq n \leq L-1$. After L extensions, the growing strand is complete, the old strand is displaced, and the template is once again in state L, which is the initial state. However, if flipping of the tail occurs at any of these intermediate stages, the incomplete new strand is lost, and the template returns to the initial state without synthesis of a new strand. From Fig 4, we find the following equations for the rates of change of the concentrations.

$$\frac{dC_0}{dt} = k_{on}'C_L - \left(k_{ext}' + k_{flip}\right)C_0$$

$$\frac{dC_n}{dt} = k_{ext}'C_{n-1} - \left(k_{ext}' + k_{flip}\right)C_n \quad \text{for } 1 \leq n \leq L-1$$

$$\frac{dC_L}{dt} = -k'_{on}C_L + k'_{ext}C_{L-1} + \sum_{n=0}^{L-1} k_{flip}C_n$$

In the stationary state, all the double strand concentrations are constant, and we obtain:

$$C_n = C_L\left(\frac{k'_{on}}{k'_{ext}+k_{flip}}\right)P_{ext}^{n} \quad \text{for } 0 \le n \le L-1$$

Here, $P_{ext} = \left(\frac{k'_{ext}}{k'_{ext}+k_{flip}}\right)$ is the probability that a single nucleotide is ligated to the primer before the flip occurs. Extension of the primer by n nucleotides requires n nucleotides to be added without a flip occurring. Therefore, the concentrations decrease exponentially in proportion to P_{ext}^{n}. The total concentration of double strands in all these states is

$$C_{tot} = C_L + \sum_{n=0}^{L-1} C_n = C_L(1 + \left(\frac{k'_{on}}{k_{flip}}\right)(1 - P_{ext}^{L})).$$

Complete single strands are being created by displacement whenever the final nucleotide is added to state L-1. The quantity we are trying to calculate in this section is R_{dis}, which we define to be the rate of synthesis of single strands per double strand template:

$$R_{dis} = \frac{k'_{ext}C_{L-1}}{C_{tot}} = \left(\frac{k_{flip}k'_{on}}{k_{flip}+k'_{on}(1-P_{ext}^{L})}\right)P_{ext}^{L}.$$

If flipping is fast compared to k'_{on}, $R_{dis} \approx k'_{on}P_{ext}^{L}$. This rate is extremely small because of the factor P_{ext}^{L}. For the estimated parameter values, $k_{flip} \gg k'_{ext}$, $P_{ext} \ll 1$, and therefore, R_{dis} is vanishingly small. Even if $k_{flip} = k'_{ext}$, we still have a factor of $(1/2)^{L}$ in the rate, therefore synthesis via strand displacement is still extremely slow unless the sequence is very short. The only way that strand synthesis from a linear template can occur at an appreciable rate is if $Lk_{flip} \ll k'_{ext}$, such that the flipping reaction hardly ever occurs even once during the synthesis of the whole sequence. In this limit, $P_{ext}^{L} \approx 1 - Lk_{flip}/k'_{ext}$, and

$$R_{dis} \approx \left(\frac{k_{flip}k'_{on}}{k_{flip}+k'_{on}\left(L\frac{k_{flip}}{k'_{ext}}\right)}\right)\left(1 - \frac{L\,k_{flip}}{k'_{ext}}\right) \approx \left(\frac{k'_{on}k'_{ext}}{k'_{ext}+Lk'_{on}}\right).$$

If we further assume that the primer addition rate k'_{on} is fast compared to extension rate, this equation simplifies to $R_{dis} \approx k'_{ext}/L$, which is comparable to the rate of synthesis from a single strand, $R_{syn} \approx k_{ext}/L$.

Synthesis by strand-displacement on a circular double strand

For circular templates, RNA synthesis can occur via a rolling-circle mechanism (Fig. 2 and Fig 4). In this mechanism, a break is present in one of the strands of the circle. One end of this strand acts as a primer to be repeatedly extended while displacing the tail end of the same strand. The slow annealing of a primer to a linear double strand (k'_{on}) is thus avoided. During rolling-circle replication in viruses and viroids, when the tail cleaves itself free from the circle due to the presence of a hairpin, hammerhead, or similar self-cleaving ribozyme within the strand (Ferré-D'Amaré and Scott 2010; Flores et al. 2011). For our calculation, we assume that this self-cleavage occurs immediately when the tail gets to the full length of the template.

The model for calculation of the rate of rolling-circle replication in shown in Fig 4. State 0 is a circular double strand with a nick in one strand but no additional strand growth. State n has n nucleotides added to the growing strand, with $1 \leq n \leq L$. We will assume that growth is only possible when the 3' end is down and the 5' end is in the tail. We also assume that flipping of the tail backwards and forwards occurs rapidly in comparison to the extension reaction, such that the growing end is annealed to the template half the time. The rate of growth is therefore $k'_{ext}/2$ when $n \geq 1$, and k'_{ext} when $n = 0$. For simplicity, we assume that growth cannot occur beyond length L because cleavage occurs rapidly. Hence the equations for the concentrations in this model are

$$\frac{dC_0}{dt} = k_{cleave}C_L - k'_{ext}C_0$$

$$\frac{dC_1}{dt} = k'_{ext}C_0 - \frac{k'_{ext}}{2}C_1$$

$$\frac{dC_n}{dt} = \frac{k'_{ext}}{2}C_{n-1} - \frac{k'_{ext}}{2}C_n, \quad \text{for } 2 \leq n \leq L-1$$

$$\frac{dC_L}{dt} = \frac{k'_{ext}}{2}C_{L-1} - \ k_{cleave}C_L.$$

In the stationary state

$$C_n = 2C_0, \quad \text{for } 1 \leq n \leq \ L-1,$$

$$C_L = \frac{k'_{ext}}{k_{cleave}}C_0,$$

and the total double strand concentration is

$$C_{tot} = C_0\left(1 + 2(L-1) + \frac{k'_{ext}}{k_{cleave}}\right).$$

The rate of production of single strands per double stranded template is therefore:

$$R_{dis} = k_{cleave}\left(\frac{C_L}{C_{tot}}\right) = \frac{k'_{ext}}{2L-1+\frac{k'_{ext}}{k_{cleave}}}.$$

When cleavage is rapid compared to extension, and $L \gg 1$, this is simply $R_{dis} = \frac{k'_{ext}}{2L}$. If rolling-circle replication were catalyzed by a polymerase that prevented the tail flipping, the factor of two would be removed (and the extension rate k'_{ext} might be faster as well), but the polymerase is not essential in the rolling-circle case.

Acknowledgements

This research was funded by the Natural Sciences and Engineering Research Council of Canada, Discovery Grant number 2017-05911.

References

Achilles T, von Kiedrowski G. 1993. A Self-Replicating System from Three Starting Materials. Angew Chemie Int Ed English. 32(8):1198–1201. doi:10.1002/anie.199311981.

Attwater J, Raguram A, Morgunov AS, Gianni E, Holliger P. 2018. Ribozyme-catalysed RNA synthesis using triplet building blocks. Elife. 7. doi:10.7554/eLife.35255.

Attwater J, Wochner A, Holliger P. 2013. In-ice evolution of RNA polymerase ribozyme activity. Nat Chem. 5(12):1011–1018. doi:10.1038/nchem.1781.

Bartel DP. 1999. Re-creating an RNA Replicase. RNA World, Second Ed Nat Mod RNA Suggest a Prebiotic RNA World. 37:143–162.

Cheng LKL, Unrau PJ. 2010. Closing the circle: replicating RNA with RNA. Cold Spring Harb Perspect Biol. 2(10). doi:10.1101/cshperspect.a002204.

Collins RF, Gellatly DL, Sehgal OP, Abouhaidar MG. 1998. Self-Cleaving Circular RNA Associated with Rice Yellow Mottle Virus Is the Smallest Viroid-like RNA. Virology. 241(2):269–275. doi:10.1006/viro.1997.8962.

Craig ME, Crothers DM, Doty P. 1971. Relaxation kinetics of dimer formation by self complementary oligonucleotides. J Mol Biol. doi:10.1016/0022-2836(71)90434-7.

Diener TO. 1989. Circular RNAs: Relics of precellular evolution?

Diener TO. 2016. Viroids: “living fossils” of primordial RNAs? Biol Direct. 11(1). doi:10.1186/s13062-016-0116-7.

Doudna JA, Cech TR. 2002. The chemical repertoire of natural ribozymes. Nature. 418(6894):222–228. doi:10.1038/418222a.

Edeleva E, Salditt A, Stamp J, Schwintek P, Boekhoven J, Braun D. 2019. Continuous nonenzymatic cross-replication of DNA strands with in situ activated DNA oligonucleotides. Chem Sci. 10(22):5807–5814. doi:10.1039/c9sc00770a.

Ferré-D’Amaré AR, Scott WG. 2010. Small self-cleaving ribozymes. Cold Spring Harb Perspect Biol. 2(10). doi:10.1101/cshperspect.a003574.

Fire A, Xu SQ. 1995. Rolling replication of short DNA circles. Proc Natl Acad Sci. 92(10):4641–4645. doi:10.1073/pnas.92.10.4641.

Flores R, Gago-Zachert S, Serra P, Sanjuán R, Elena SF. 2014. Viroids: Survivors from the RNA World? Annu Rev Microbiol. 68(1):395–414. doi:10.1146/annurev-micro-091313-103416.

Flores R, Grubb D, Elleuch A, Nohales MÁ, Delgado S, Gago S. 2011. Rolling-circle replication of viroids, viroid-like satellite RNAs and hepatitis delta virus: Variations on a theme. RNA Biol. 8(2):200–206. doi:10.4161/rna.8.2.14238.

Flores R, Ruiz-Ruiz S, Serra P. 2012. Viroids and hepatitis delta virus. Semin Liver Dis. 32(3):201–210. doi:10.1055/s-0032-1323624.

Flores R, Di Serio F, Hernández C. 1997. Viroids: The noncoding genomes. Semin Virol.

8(1):65–73. doi:10.1006/smvy.1997.0107.

Frieden M, Pedroso E, Kool ET. 1999. Tightening the belt on polymerases: Evaluating the physical constraints on enzyme substrate size. Angew Chemie - Int Ed. doi:10.1002/(SICI)1521-3773(19991216)38:24<3654::AID-ANIE3654>3.0.CO;2-S.

Green C, Tibbetts C. 1981. Reassociation rate limited displacement of DNA strands by branch migration. Nucleic Acids Res. 9(8):1905–1918. doi:10.1093/nar/9.8.1905.

Hammann C, Luptak A, Perreault J, De La Peña M. 2012. The ubiquitous hammerhead ribozyme. Rna. 18(5):871–885. doi:10.1261/rna.031401.111.

He C, Gállego I, Laughlin B, Grover MA, Hud N V. 2017. A viscous solvent enables information transfer from gene-length nucleic acids in a model prebiotic replication cycle. Nat Chem. 9(4):318–324. doi:10.1038/nchem.2628.

He C, Lozoya-Colinas A, Gállego I, Grover MA, Hud N V. 2019. Solvent viscosity facilitates replication and ribozyme catalysis from an RNA duplex in a model prebiotic process. Nucleic Acids Res. 47(13):6569–6577. doi:10.1093/nar/gkz496.

Hertel KJ, Uhlenbeck OC, Herschlag D. 1994. A Kinetic and Thermodynamic Framework for the Hammerhead Ribozyme Reaction. Biochemistry. 33(11):3374–3385. doi:10.1021/bi00177a031.

Higgs PG. 2017. Chemical Evolution and the Evolutionary Definition of Life. J Mol Evol. 84(5–6):225–235. doi:10.1007/s00239-017-9799-3.

Higgs PG, Lehman N. 2015. The RNA World: Molecular cooperation at the origins of life. Nat Rev Genet. 16(1):7–17. doi:10.1038/nrg3841.

Horning DP, Joyce GF. 2016. Amplification of RNA by an RNA polymerase ribozyme. Proc Natl Acad Sci U S A. 113(35):9786–9791. doi:10.1073/pnas.1610103113.

Hulo C, De Castro E, Masson P, Bougueleret L, Bairoch A, Xenarios I, Le Mercier P. 2011. ViralZone: A knowledge resource to understand virus diversity. Nucleic Acids Res. 39(SUPPL. 1). doi:10.1093/nar/gkq901.

Ianeselli A, Mast CB, Braun D. 2019. Periodic Melting of Oligonucleotides by Oscillating Salt Concentrations Triggered by Microscale Water Cycles Inside Heated Rock Pores. Angew Chemie - Int Ed. 58(37):13155–13160. doi:10.1002/anie.201907909.

von Kiedrowski G. 1986. A Self-Replicating Hexadeoxynucleotide. Angew Chemie Int Ed English. 25(10):932–935. doi:10.1002/anie.198609322.

Kreysing M, Keil L, Lanzmich S, Braun D. 2015. Heat flux across an open pore enables the continuous replication and selection of oligonucleotides towards increasing length. Nat Chem. 7(3):203–208. doi:10.1038/nchem.2155.

Lathe R. 2004. Fast tidal cycling and the origin of life. Icarus. 168(1):18–22. doi:10.1016/j.icarus.2003.10.018.

Lathe R. 2005. Tidal chain reaction and the origin of replicating biopolymers. Int J Astrobiol. 4(1):19–31. doi:10.1017/S1473550405002314.

Leu K, Obermayer B, Rajamani S, Gerland U, Chen IA. 2011. The prebiotic evolutionary advantage of transferring genetic information from RNA to DNA. Nucleic Acids Res. doi:10.1093/nar/gkr525.

Li Y, Breaker RR. 1999. Kinetics of RNA degradation by specific base catalysis of transesterification involving the 2γ-hydroxyl group. J Am Chem Soc. 121(23):5364–5372. doi:10.1021/ja990592p.

Liu D, Daubendiek SL, Zillman MA, Ryan K, Kool ET. 1996. Rolling Circle DNA Synthesis: Small Circular Oligonucleotides as Efficient Templates for DNA Polymerases. J Am Chem Soc. 118(7):1587–1594. doi:10.1021/ja952786k.

Ma W, Yu C, Zhang W. 2013. Circularity and self-cleavage as a strategy for the emergence of a chromosome in the RNA-based protocell. Biol Direct. 8(1). doi:10.1186/1745-6150-8-21.

Mansy SS, Szostak JW. 2009. Reconstructing the emergence of cellular life through the synthesis of model protocells. Cold Spring Harb Symp Quant Biol. 74:47–54. doi:10.1101/sqb.2009.74.014.

Mariani A, Bonfio C, Johnson CM, Sutherland JD. 2018. PH-Driven RNA Strand Separation under Prebiotically Plausible Conditions. Biochemistry. 57(45):6382–6386. doi:10.1021/acs.biochem.8b01080.

Müller UF. 2006. Re-creating an RNA world. Cell Mol Life Sci. 63(11):1278–1293. doi:10.1007/s00018-006-6047-1.

Olsen TJ, Choi Y, Sims PC, Gul OT, Corso BL, Dong C, Brown WA, Collins PG, Weiss GA. 2013. Electronic measurements of single-molecule processing by DNA polymerase i (Klenow fragment). J Am Chem Soc. 135(21):7855–7860. doi:10.1021/ja311603r.

Perreault J, Weinberg Z, Roth A, Popescu O, Chartrand P, Ferbeyre G, Breaker RR. 2011. Identification of Hammerhead Ribozymes in All Domains of Life Reveals Novel Structural Variations. PLoS Comput Biol. 7(5). doi:10.1371/journal.pcbi.1002031.

Radding CM, Beattie KL, Holloman WK, Wiegand RC. 1977. Uptake of homologous single-stranded fragments by superhelical DNA. IV. Branch migration. J Mol Biol. 116(4):825–839. doi:10.1016/0022-2836(77)90273-X.

Rauzan B, McMichael E, Cave R, Sevcik LR, Ostrosky K, Whitman E, Stegemann R, Sinclair AL, Serra MJ, Deckert AA. 2013. Kinetics and thermodynamics of DNA, RNA, and hybrid duplex formation. Biochemistry. 52(5):765–772. doi:10.1021/bi3013005.

Robertson MP, Joyce GF. 2012. The origins of the RNA World. Cold Spring Harb Perspect Biol. 4(5):1. doi:10.1101/cshperspect.a003608.

Rohatgi R, Bartel DP, Szostak JW. 1996. Nonenzymatic, template-directed ligation of oligoribonucleotides is highly regioselective for the formation of 3′-5′ phosphodiester bonds. J Am Chem Soc. 118(14):3340–3344. doi:10.1021/ja9537134.

Ruffner DE, Uhlenbeck OC, Stormo GD. 1990. Sequence Requirements of the Hammerhead RNA Self-Cleavage Reaction. Biochemistry. 29(47):10695–10702. doi:10.1021/bi00499a018.

Schwartz JJ, Quake SR. 2009. Single molecule measurement of the “speed limit” of DNA polymerase. Proc Natl Acad Sci U S A. doi:10.1073/pnas.0907404106.

Sievers D, Von Kiedrowski G. 1994. Self-replication of complementary nucleotide-based oligomers. Nature. 369(6477):221–224. doi:10.1038/369221a0.

Soslau G. 2018. Circular RNA (circRNA) was an important bridge in the switch from the RNA world to the DNA world. J Theor Biol. 447:32–40. doi:10.1016/j.jtbi.2018.03.021.

Soukup GA, Breaker RR. 1999. Relationship between internucleotide linkage geometry and the stability of RNA. RNA. doi:10.1017/S1355838299990891.

Srinivas N, Ouldridge TE, Šulc P, Schaeffer JM, Yurke B, Louis AA, Doye JPK, Winfree E. 2013. On the biophysics and kinetics of toehold-mediated DNA strand displacement. Nucleic Acids Res. 41(22):10641–10658. doi:10.1093/nar/gkt801.

Szostak JW. 2012. The eightfold path to non-enzymatic RNA replication. J Syst Chem. 3(1). doi:10.1186/1759-2208-3-2.

Takeshita D, Tomita K. 2012. Molecular basis for RNA polymerization by Qβ replicase. Nat Struct Mol Biol. 19(2):229–238. doi:10.1038/nsmb.2204.

Turner DH, Mathews DH. 2009. NNDB: The nearest neighbor parameter database for predicting stability of nucleic acid secondary structure. Nucleic Acids Res. 38(SUPPL.1). doi:10.1093/nar/gkp892.

Wachowius F, Holliger P. 2019. Non-Enzymatic Assembly of a Minimized RNA Polymerase Ribozyme. ChemSystemsChem. 1(1–2):12–15. doi:10.1002/syst.201900004.

Walton T, Pazienza L, Szostak JW. 2019. Template-Directed Catalysis of a Multistep Reaction Pathway for Nonenzymatic RNA Primer Extension. Biochemistry. 58(6):755–762. doi:10.1021/acs.biochem.8b01156.

Wetmur JG, Davidson N. 1968. Kinetics of renaturation of DNA. J Mol Biol. 31(3):349–370. doi:10.1016/0022-2836(68)90414-2.

Williams AP, Longfellow CE, Freier SM, Kierzek R, Turner DH. 1989. Laser Temperature-Jump, Spectroscopic, and Thermodynamic Study of Salt Effects on Duplex Formation by dGCATGC. Biochemistry. doi:10.1021/bi00436a025.

Xia T, SantaLucia J, Burkard ME, Kierzek R, Schroeder SJ, Jiao X, Cox C, Turner DH. 1998. Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson - Crick base pairs. Biochemistry. 37(42):14719–14735. doi:10.1021/bi9809425.

Zhou L, Kim SC, Ho KH, O’Flaherty DK, Giurgiu C, Wright TH, Szostak JW. 2019. Non-enzymatic primer extension with strand displacement. Elife. 8. doi:10.7554/eLife.51888.

Zielinski WS, Orgel LE. 1987. Autocatalytic synthesis of a tetranucleotide analogue. Nature. 327(6120):346–347. doi:10.1038/327346a0.

Figures

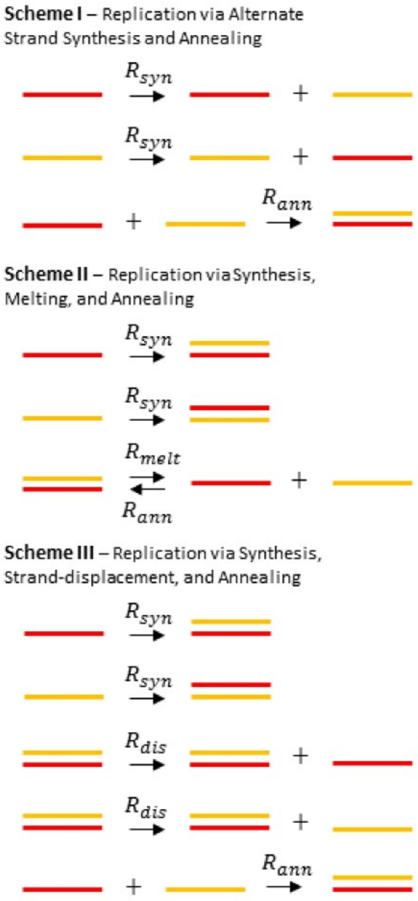

Scheme I – Replication via Alternate Strand Synthesis and Annealing
R_{syn}
+
R_{syn}
+
+
R_{ann}
Scheme II – Replication via Synthesis, Melting, and Annealing
R_{syn}
R_{syn}
R_{melt}
R_{ann}
+
Scheme III – Replication via Synthesis, Strand-displacement, and Annealing
R_{syn}
R_{syn}
R_{dis}
+
R_{dis}
+
+
R_{ann}

Fig 1 Three possible schemes for RNA replication. Red and orange strands are complementary to one another.

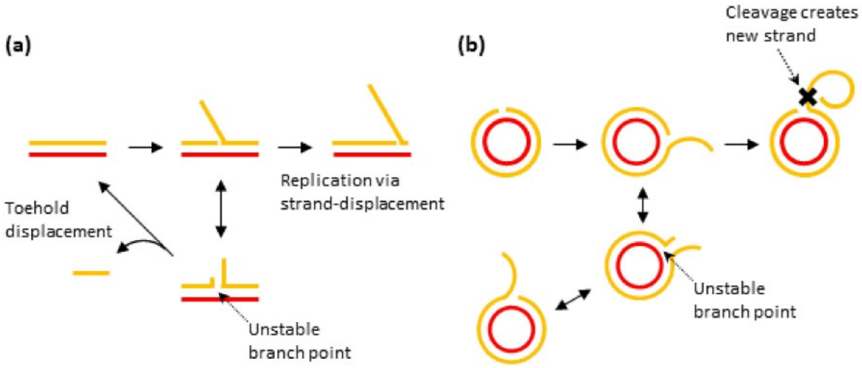


Fig 2. Toehold-mediated displacement is an important problem for the linear strand-displacement mechanism (a), but not for the rolling circle mechanism (b).

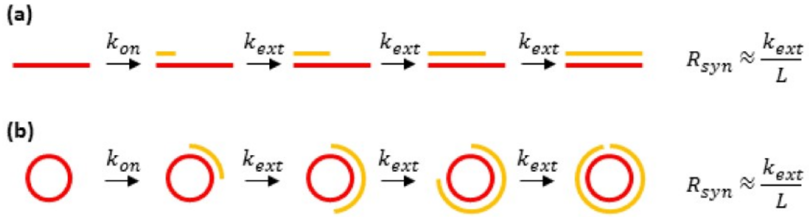


Fig 3: Mechanisms of RNA synthesis on linear (a) and circular (b) ssRNA templates. When primer extension is rate limiting, the net rate of RNA synthesis is inversely proportional to the number of extension reactions required (L).

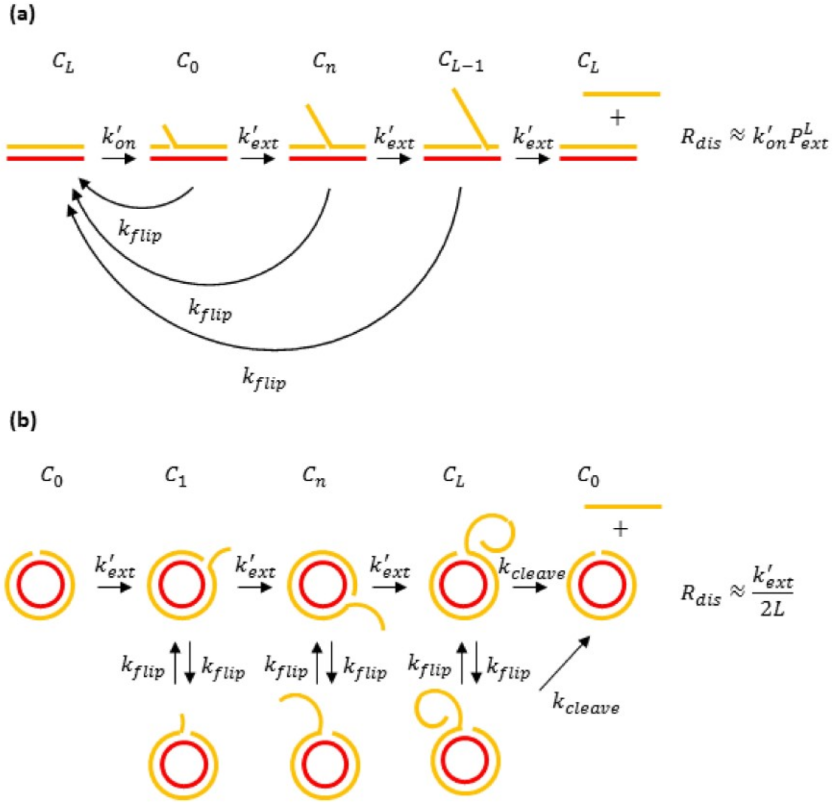


Fig 4: Mechanisms of RNA synthesis from a linear (a) or circular (b) dsRNA template. The extension rate in this case, k'_{ext}, is different from that of a ssRNA template, k_{ext}. For a linear template, the competing tail-displacement reaction, k_{flip}, results in RNA synthesis being proportional to $P^L_{ext} = \left(\frac{k'_{ext}}{k'_{ext}+k_{flip}}\right)^L$. The approximate rate shown is the limit in which $k'_{on} \ll k_{flip}$. In the case of rolling-circle, the net rate of RNA synthesis via displacement is $R_{dis} \approx \frac{k'_{ext}}{2L}$, comparable to that of RNA synthesis on a single stranded template: $R_{syn} \approx \frac{k_{ext}}{L}$.

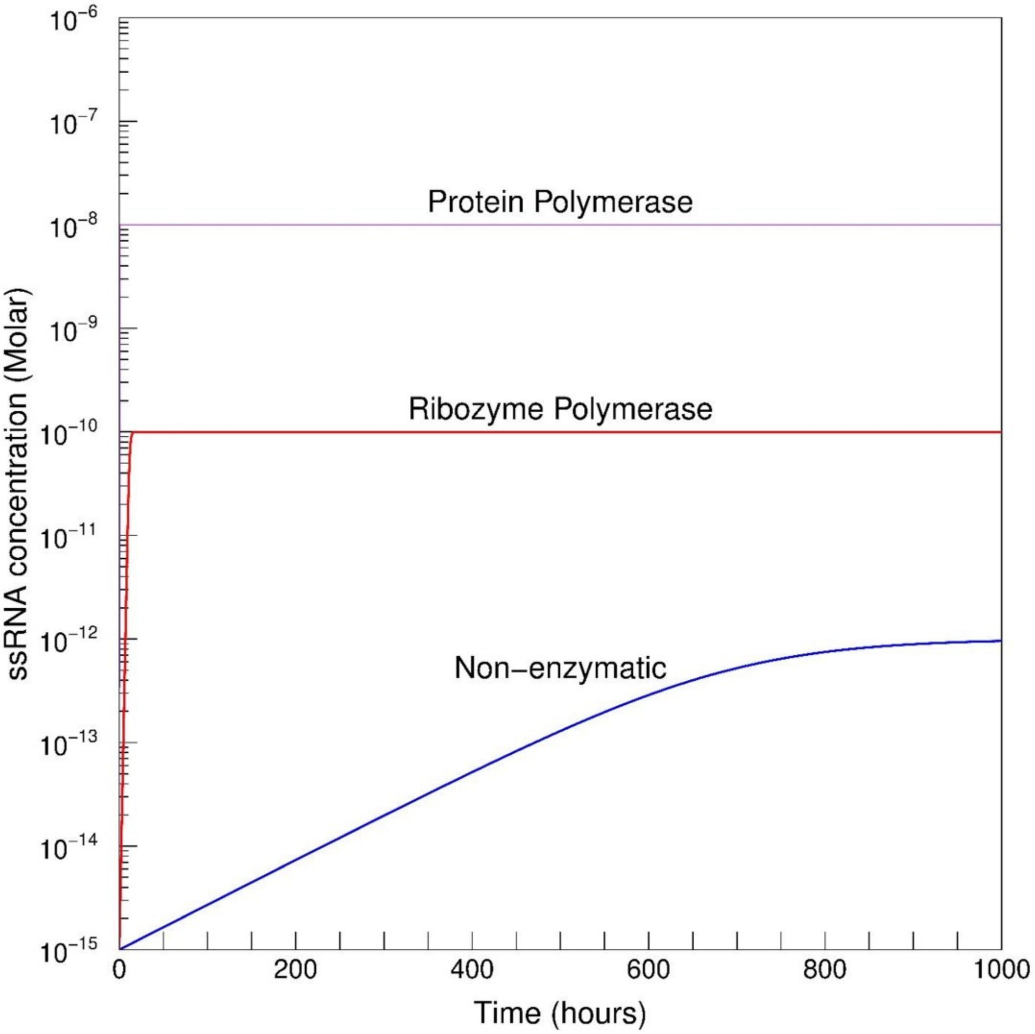


Fig 5 - Replication via Scheme I leads to saturation at $C_P = C^* = R_{syn}/R_{ann}$. Non-enzymatic curve assumes an $R_{syn} = 0.01\ h^{-1}$, ribozyme polymerase assumes an $R_{syn} = 1\ h^{-1}$, and protein polymerase assumes an $R_{syn} = 100\ h^{-1}$. All curves assume an $R_{ann} = 10^{10}\ M^{-1}h^{-1}$ and a C_{init} of $10^{-15}\ M$ for both plus and minus strands.

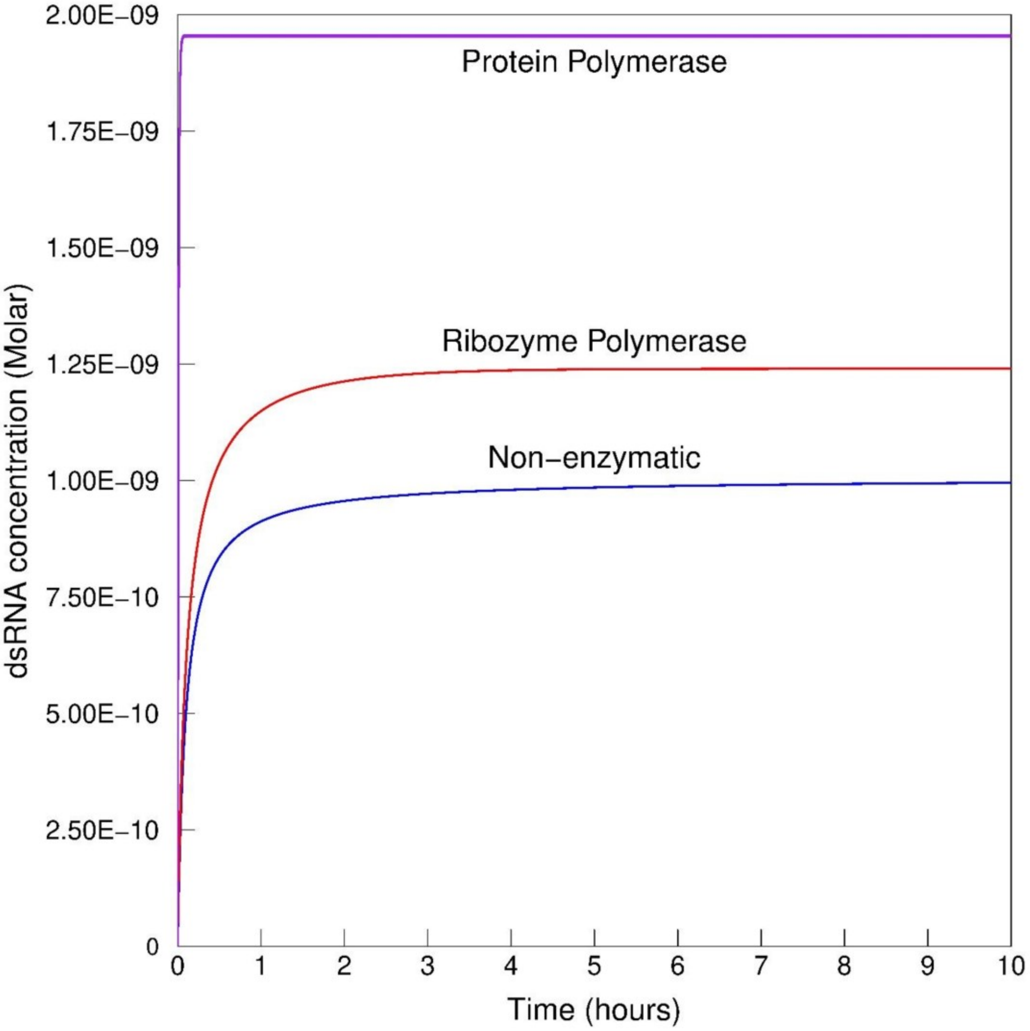


Fig 6 - A single cycle of synthesis with Scheme II. $C_{init} = 10^{-9}$ for all curves. Non-enzymatic curve assumes an $R_{syn} = 0.01\ h^{-1}$, ribozyme polymerase assumes an $R_{syn} = 1\ h^{-1}$, and protein polymerase assumes an $R_{syn} = 100\ h^{-1}$. If synthesis occurs at a rate of a modern protein polymerase, then the concentration nearly doubles. If synthesis occurs at a rate of a ribozyme polymerase, then the concentration increases by roughly 25%. Wherease for non-enzymatic synthesis rates the concentration returns to its initial value of C_{init} with no significant increase.

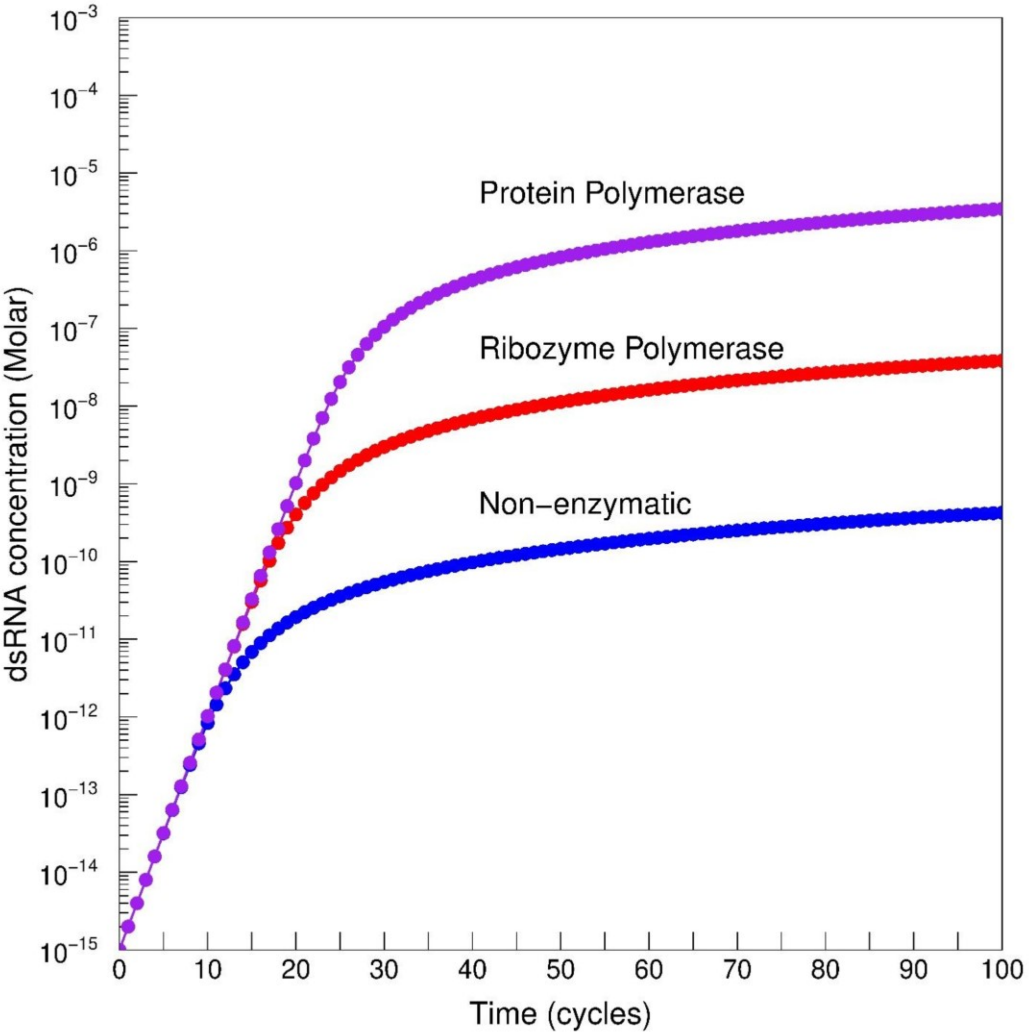


Fig 7 - Replication via Scheme II with cycling. The points show the double strand concentration at the end of each cycle. Concentrations double each cycle until annealing becomes dominant at the critical concentration C^*, after which growth is sub-exponential. Non-enzymatic curve assumes an $R_{syn} = 0.01\ h^{-1}$, ribozyme polymerase assumes an $R_{syn} = 1\ h^{-1}$, and protein polymerase assumes an $R_{syn} = 100\ h^{-1}$. For all curves $R_{ann} = 10^{10}\ M^{-1}h^{-1}$ and $C_{init} = 10^{-15}\ M$.

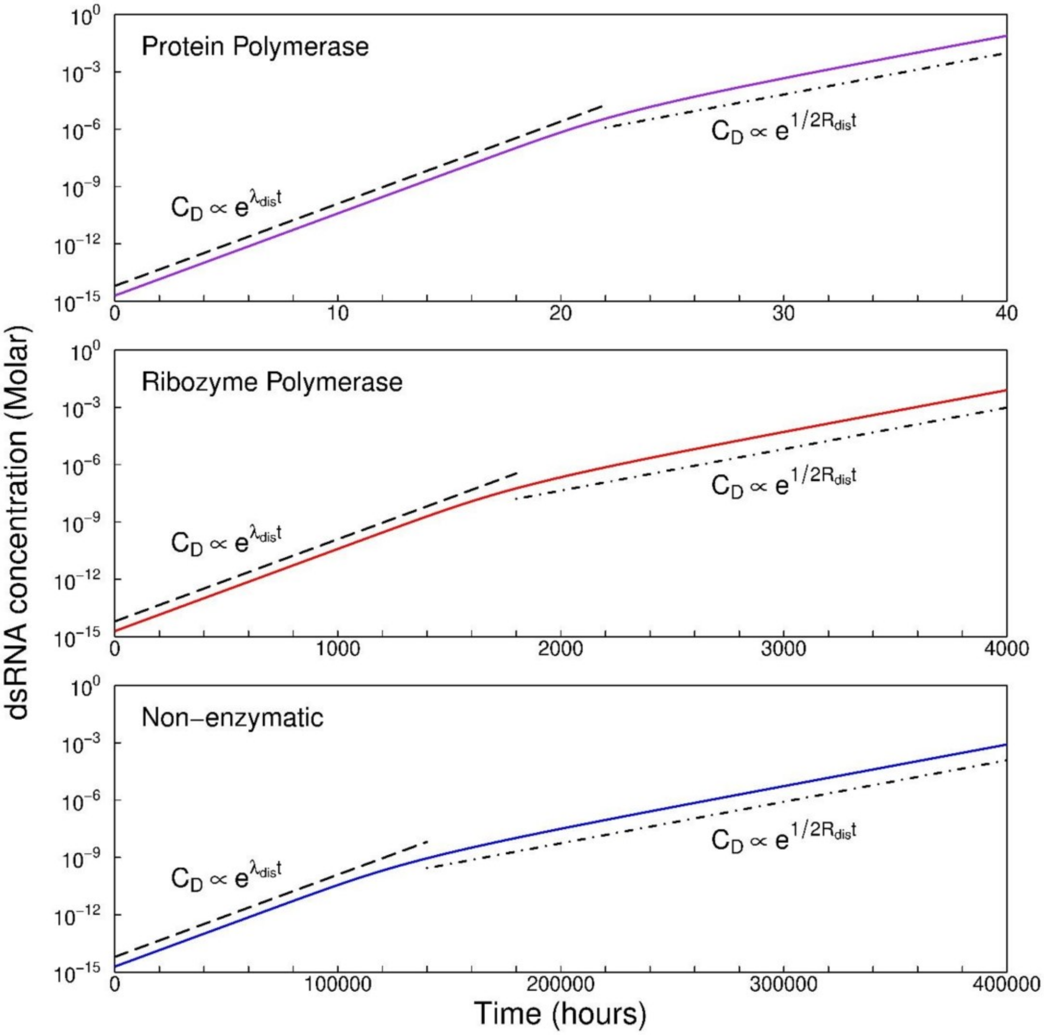


Fig 8 - Replication via Scheme III. Non-enzymatic curve assumes an $R_{syn} = 0.01\ h^{-1}$, ribozyme polymerase assumes an $R_{syn} = 1\ h^{-1}$, and protein polymerase assumes an $R_{syn} = 100\ h^{-1}$. For all curves $R_{dis} = R_{syn}/100$, $R_{ann} = 10^{10}\ M^{-1}h^{-1}$, and $C_{init} = 10^{-15}\ M$. The upper dashed curves indicate exponential growth proportional to $\exp(\lambda_{dis}t)$. The lower dashed curves indicate exponential growth proportional to $\exp(1/2R_{dis}t)$. Note that different time scales are used in each plot.

Chapter 6: Error-correction in an RNA world

The contents of this chapter are unpublished and constitute a manuscript in preparation. I developed the model, implemented the computer simulations, and wrote the manuscript. Paul Higgs provided guidance throughout the research proc

Non-enzymatic Rolling-circle Synthesis and the Possibility of Error Correction in the RNA world

Abstract:

The poor fidelity of non-enzymatic and ribozyme-catalyzed RNA synthesis poses a significant obstacle to the RNA world theory for the origin of life. When the fidelity of synthesis is poor, errors accumulate with each generation and an error catastrophe ensues, resulting in the loss of any encoded ribozymes. In this paper, we consider a simple model of non-enzymatic rolling-circle synthesis based on the current kinetic and thermodynamic data found in the literature. We find that there are two important limits for rolling-circle synthesis corresponding to slow and fast tail-flipping, which is analogous to the process of toehold-mediated strand displacement. When tail-flipping is slow, the 3' end of the primer remains annealed to the template and fidelity of synthesis is poor. This is the limit in which viroids and viruses appear to occupy due to their reliance on a protein polymerase. When no polymerase is present, the experimental data predicts fast tail-flipping. In which case, the template flips between a 3' end annealed state and a 5' end annealed state. In the fast limit, incorporated errors are removed through a thermodynamically driven error-correction mechanism. We investigate this error-correction and find that under mundane prebiotic conditions, the resulting fidelity of product sequences is sufficient to prevent an error catastrophe. Non-enzymatically replicating templates would therefore be free to discover complex ribozymes. Interestingly, the presence of this error-correction provides a selective advantage for ribozymes which further enhance the fidelity of primer extension, or alternatively, enhance the rate of error-correction through nuclease activity. The evolution of a complex polymerase ribozyme would therefore by driven by selection on fast replicating templates.

1. Introduction:

The RNA world hypothesis states that life began from a relatively simple RNA polymer which was capable of replication (Joyce 2002; Higgs and Lehman 2015). Initially, this replication would have occurred under non-enzymatic conditions, prior to the advent of ribozyme catalysts. At this stage, the non-enzymatically replicating templates would undergo natural selection based on their physiochemical properties, such as stability against hydrolysis or UV damage (Higgs 2017). Through replication, these templates would explore sequence space and discover ribozymes. Any ribozyme which enhanced the rate of replication could also be selected for, even if the initial rate enhancement were minor. Through further replication and stepwise refinement, a complex ribozyme such a polymerase could have emerged. The discovery of a polymerase ribozyme would have greatly enhanced the rate of replication and allowed for larger RNA genomes and the continued evolution of the RNA organism.

The fatal flaw to stories such as this one is the poor fidelity of RNA replication. Without accurate replication, any polymerase ribozyme discovered would be inevitably lost through the accumulation of mutations. To maintain complex ribozymes, and long RNA genomes, accurate replication is a necessity. Eigen's error threshold theory quantifies this relationship between the accuracy of replication and the maximum genome length by considering a population of master and mutant sequences. Master sequences contain all the information of the RNA organism, whereas mutants are created through inaccurate replication of the master sequences. According to Eigen's theory, the maximum genome length of an RNA organism is $L^* = \frac{\ln(\sigma)}{1-q}$, where σ is the rate superiority of the master sequence, and q is the per-base fidelity of replication (Eigen et al. 1988). While the value of σ is an unknown, it is typically assumed to be roughly e such that the numerator is unity. This corresponds to a relatively small replicative advantage of the master sequence over the mutant sequences.

Using Eigen's theory, we can determine the minimum fidelity required to maintain a genome of a given length. For instance, known polymerase ribozymes contain roughly 200 bases (Attwater et al. 2013; Horning and Joyce 2016; Attwater et al. 2018) and would therefore require a fidelity of $q = 0.995$ to be maintained. To leave the RNA world, an RNA genome of ~10,000 bases is required for protein synthesis (Jeffares et al. 1998), corresponding to a minimum fidelity of $q = 0.9999$. In contrast, the experimentally measured fidelity of RNA synthesis under non-enzymatic and ribozyme catalyzed conditions is at best $q = 0.99$ (see Table 1). This corresponds to a maximum genome length of roughly 100 bases, which is insufficient to maintain a polymerase ribozyme, let alone an RNA organism capable of primitive protein synthesis. With this fidelity, emerging life would be stuck at the non-enzymatic stage of evolution since maintaining complex ribozymes is not possible. This problem of poor RNA fidelity is commonly referred to as Eigen's paradox (Kun et al. 2015) or the catch-22 of the origin of life (Smith 1983).

The phenotypic error threshold theory builds on Eigen's theory and incorporates the possibility of neutral mutations (Takeuchi et al. 2005). With neutral mutations, the maximum genome length scales as: $L^* = \frac{\ln(\sigma)}{-ln(q+(1-q)\lambda)}$, where λ is the probability that a point mutation is neutral. Based on the naturally occurring hairpin and *Neurospora* VS ribozymes, λ values of 0.22 and 0.26 have been inferred respectively (Kun et al. 2005). Incorporating neutral mutations relaxes the fidelity constraints slightly. For instance, to maintain a polymerase ribozyme of 200 bases, a minimum fidelity of 0.993 is required, as opposed to 0.995 without neutral mutations. Similarly, to maintain a genome of 10,000 base pairs would require a fidelity of 0.99986 as opposed to 0.9999. This, however, is assuming that the probability of neutral mutation for complex polymerase ribozymes or an RNA organism is the same as the relatively simple hairpin or *Neurospora* VS ribozymes. Furthermore, even when σ is optimistically increased to ~350, a fidelity of 0.999 is still required for the last RNA organism (Kun et al. 2015). Therefore, there is still a disconnect between the fidelity of RNA synthesis observed in experiments and that which is required for replication in an RNA world.

In experiments on non-enzymatic synthesis of RNA, it is observed that the fidelity of full-length product sequences tends to be higher than incomplete fragments. This is due to the stalling of primer extension after an error is incorporated. Since sequences which have incorporated an error are slower to replicate, the error threshold is relaxed. With primer stalling, the maximum genome length is expected to scale as: $L^* = \frac{\ln(\sigma+(\sigma-1)(1-q)S)}{1-q}$, where S is the stalling factor (Rajamani et al. 2010). Under non-enzymatic conditions $S \approx 100$ (Rajamani et al. 2010), resulting in a minimum fidelity of 0.993 to maintain a polymerase ribozyme. Whereas the fidelity to maintain the last RNA organism remains unchanged at 0.9999. While the effect of stalling increases the average fidelity of product sequences, it also drastically decreases the rate of synthesis. For long polymers, the incorporation of at least one error is more or less guaranteed. The problem with primer stalling is that further primer extension after the error has abysmal fidelity, $q \approx 0.38$ (Leu et al. 2013). In some cases, the most likely extension is no longer the Watson-Crick complement, i.e. the extension fidelity becomes worse than random chance. If the primer is unable to correct itself, then the net rate of product synthesis becomes orders of magnitude slower. This could be disastrous given the fast rate of template degradation from hydrolysis.

In this paper, we extend our previous work and consider a more realistic model of rolling-circle synthesis based on the experimentally known kinetics and thermodynamics found in the literature. Remarkably different results are found depending on the net rate of "tail flipping", which is analogous to the process of toehold-mediated displacement. When the rate of tail flipping is much faster than primer extension, as is expected under non-enzymatic conditions, rolling-circle synthesis undergoes a thermodynamically driven error correction cycle. This error correction results in product sequences with average fidelities in excess of 0.9999. Eigen's paradox is thus avoided entirely when tail-flipping is fast. At the other extreme, when tail-flipping is prevented, such as due to the presence of a processive polymerase, no error correction is observed. Therefore, the same model which predicts error correction under non-enzymatic conditions also predicts the poor fidelity of rolling-circle replication by viroids and viruses.

2. Results:

2.1 Model of non-enzymatic rolling-circle synthesis

In this section, we describe the model of rolling-circle synthesis utilized in this paper. This model is intentionally minimalistic and designed to show the importance of tail-flipping during rolling-circle synthesis. Despite its simplicity, the model is consistent with the current understanding of reaction kinetics and thermodynamics, and references to the experimental values will be discussed throughout this section. For ease of discussion, the two limits of slow and fast tail-flipping are illustrated in Figure 1.

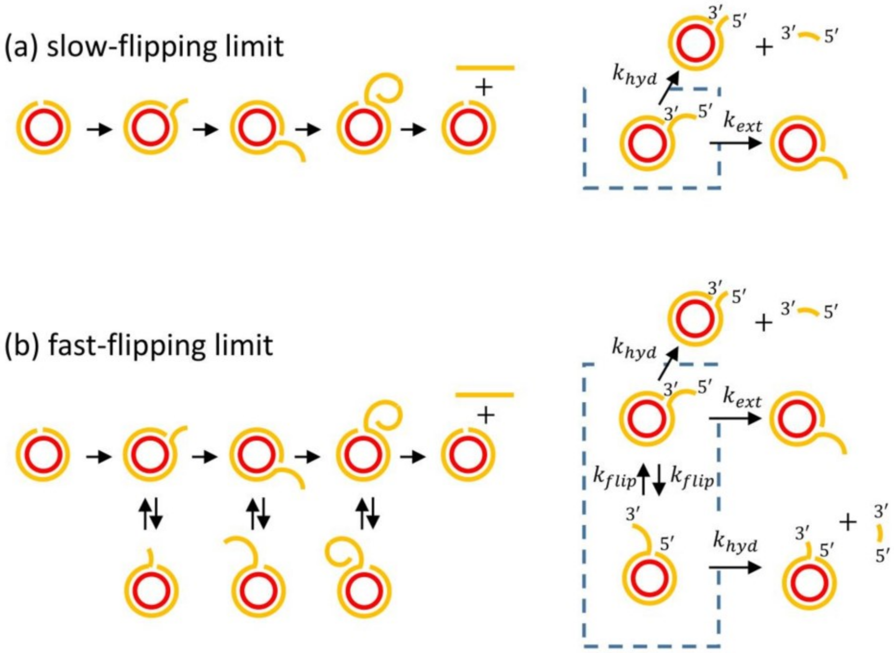


Figure 1: Two limits of rolling-circle synthesis: (a) when tail-flipping is much slower than primer extension, and (b) when tail-flipping is much faster than primer extension. In both cases, the dashed blue box on the right represents one hypothetical state of a simulation. In the slow-flipping limit, the next reaction to occur is either hydrolysis of the 5’ tail or primer extension of the 3’ annealed end. In the fast-flipping limit, the template flips between a 3’ end annealed and 5’ end annealed state. In this case, hydrolysis of the 3’ tail is also possible.

Primer extension in this model is assumed to occur at a constant fidelity, q_{ext}. Table 1 provides some experimental reference values. Depending on the reaction conditions, the fidelity of RNA primer extension varies between roughly $0.83 - 0.99$. Under non-enzymatic conditions, the fidelity of primer extension is highest when a downstream oligomer contributes an additional stacking interaction with the incoming nucleotide (Prywes et al. 2016; Tam et al. 2017). During rolling-circle synthesis, the displaced tail would contribute a similar downstream stacking interaction and would likely have a similar primer extension fidelity. We will start by assuming a fidelity of $q_{ext} = 0.99$ for our model, but will later consider fidelities ranging from 0.9 to 0.9999 to illustrate the influence this has on the dynamics of the model.

Polymer	Catalysis	Template Nucleotide → Product Nucleotide					Reference
		C→G	G→C	A→U(T)	U(T)→A	AVG	
RNA	Non-enzymatic	0.9917	0.9435	0.8397	0.5515	0.8316	Leu et al. 2011
DNA	Non-enzymatic	0.9939	0.9487	0.9058	0.8471	0.9288	Leu et al. 2011
RNA	Non-enzymatic	0.9997	0.9994	0.9843	0.9479	0.9828	Prywes et al. 2016
RNA	Ribozyme	0.992	0.997	0.996	0.948	0.983	Attwater et al. 2013
RNA	Ribozyme	0.990	0.991	0.974	0.913	0.967	Horning and Joyce 2016

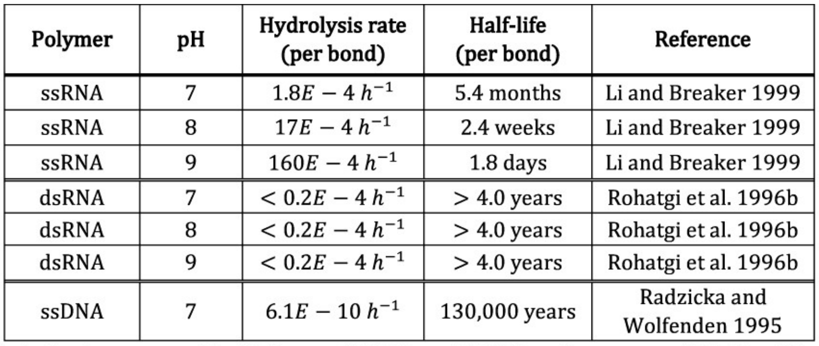
Table 1: Non-enzymatic and ribozyme-catalyzed fidelities of primer extension. The fidelities listed for Leu et al. 2011 were determined with non-equimolar concentrations; the concentration of U (or T) was 4x greater than the others nucleotide. The fidelities listed for Prywes et al. 2016 are for primer extension with a downstream oligomer. The fidelities listed for Horning and Joyce 2016 are for the wild type polymerase ribozyme, the 24-3 polymerase ribozyme has worse fidelity.

During rolling-circle synthesis, primer extension results in RNA synthesis and an increasing tail length. For simplicity, this primer extension is assumed to occur at a constant rate of k_{ext} in the 5' to 3' direction on a circular template of length L. Table 2 provides experimental reference values under non-enzymatic, ribozyme, and protein enzyme catalyzed conditions. While the focus of this paper is on non-enzymatic synthesis, the rates of ribozyme and protein enzymes provide a valuable reference to the range of possible extension rates. Under non-enzymatic conditions, the extension rates can be as fast as $27\ h^{-1}$ using activated nucleotides at millimolar concentrations and a downstream oligomer (Walton et al. 2019). Whereas using triphosphate activation results in ligation rates of roughly $9 * 10^{-5}\ h^{-1}$ (Rohatgi et al. 1996a). Since this latter value was determined for a bound oligomer, the net rate of primer extension by single nucleotides would be even slower. Ribozyme polymerases are able to significantly enhance the rate of triphosphate-based ligation to achieve primer extension rates of roughly $72\ h^{-1}$ (Horning and Joyce 2016). In comparison, protein polymerases achieve primer extension rates of $72{,}000\ h^{-1}$ at a fraction of the nucleotide concentration (Olsen et al. 2013). In terms of prebiotic chemistry, faster primer extension is less plausible than slower primer extensions as it requires harsher activation chemistry and higher nucleotide concentration.

Polymer	Catalysis	Nucleotide	[Nucl]	pH	Extension rate	Reference
RNA	Non-enzymatic	5′-triphosphate	–	7.8	$9.0E - 5h^{-1}$	Rohatgi et al. 1996a
RNA	Non-enzymatic	5′-phosphorimidazolide 3′-amino group	$10\ mM$	7	$11\ h^{-1}$	Leu et al. 2011
RNA	Non-enzymatic	5′-phosphorimidazolide 3′-downstream stacking	$2\ mM$	8	$27\ h^{-1}$	Walton et al. 2019
RNA	Ribozyme	5′-triphosphate	$4\ mM$	8.3	$3.9\ h^{-1}$	Attwater et al. 2013
RNA	Ribozyme	5′-triphosphate	$4\ mM$	8.3	$72\ h^{-1}$	Horning and Joyce 2016
DNA	Enzyme	5′-triphosphate	$10\ \mu M$	7.8	$7.2E + 4\ h^{-1}$	Olsen et al. 2013

Table 2: Non-enzymatic, ribozyme, and enzyme-catalyzed rates of primer extension. Rates for Leu et al. 2011 are average primer extension rates for G and C extension. The rate listed for Rohatgi et al. 1996a is the V_{max} for ligation based on a bound oligomer at saturating concentrations. Primer extension rate for Horning & Joyce 2016 is for the 24-3 ribozyme polymerase. Note: concentrations of Mg^{2+} and K^{+}, and temperature, vary between experimental setups.

Competing with the process of tail growth from primer extension is random hydrolysis and site-specific cleavage. Random hydrolysis is assumed to occur at a constant rate of k_{hyd} per bond in all ssRNA. Table 3 provides experimental reference values for ssRNA, dsRNA, and ssDNA at varying pH. At neutral pH, each bond in ssRNA breaks at a rate of $k_{hyd} \approx 1.8 * 10^{-4} h^{-1}$, corresponding to a per bond half-life of roughly 5 months. When the pH is increased to pH 9, the hydrolysis rate increases to $k_{hyd} \approx 160 * 10^{-4}\ h^{-1}$, corresponding to a per bond half-life of roughly 2 days. In contrast to ssRNA, dsRNA appears to be incredibly resistant to hydrolysis. While the rate of ssRNA hydrolysis increases two orders of magnitude between pH 7 and pH 9, the rate of dsRNA hydrolysis remains below detection levels (Rohatgi et al. 1996b). At pH 9, dsRNA appears to be at least 1000x more stable than ssRNA. A value which appears to be corroborated by the variability in hydrolysis rates of bonds found within folded RNA structures (Soukup and Breaker 1999). Due to the predicted stability of dsRNA bonds over ssRNA bonds, the hydrolysis of the dsRNA template is ignored in this model. For reference, the hydrolysis rate of ssDNA is also provided in Table 2. While DNA is unlikely under prebiotic conditions, we should not omit the possibility that the first non-enzymatic replication occurred with a polymer more stable than RNA.



Polymer	pH	Hydrolysis rate (per bond)	Half-life (per bond)	Reference
ssRNA	7	$1.8E - 4\,h^{-1}$	5.4 months	Li and Breaker 1999
ssRNA	8	$17E - 4\,h^{-1}$	2.4 weeks	Li and Breaker 1999
ssRNA	9	$160E - 4\,h^{-1}$	1.8 days	Li and Breaker 1999
dsRNA	7	$< 0.2E - 4\,h^{-1}$	$>$ 4.0 years	Rohatgi et al. 1996b
dsRNA	8	$< 0.2E - 4\,h^{-1}$	$>$ 4.0 years	Rohatgi et al. 1996b
dsRNA	9	$< 0.2E - 4\,h^{-1}$	$>$ 4.0 years	Rohatgi et al. 1996b
ssDNA	7	$6.1E - 10\,h^{-1}$	130,000 years	Radzicka and Wolfenden 1995

Table 3: Hydrolysis rates and half-lives of RNA and DNA polymers at pH 7-9. All ssRNA and dsRNA rates and half-lives are based on a [Mg^{2+}] of 100 mM, [K^{+}] of 50 mM, at 37°C. The rate of dsRNA hydrolysis reported by Rohatgi et al. 1996b was below detection limits regardless of pH. Therefore, hydrolysis rates for dsRNA are maximum, and half-lives are minimum possible values based on instrument sensitivity. Single stranded DNA hydrolysis rate was approximated by Radzicka and Wolfenden 1995 at [Mg^{2+}] of 0mM, [K^{+}] of 100 mM, and 23°C, using an analogous chemical compound.

In addition to random cleavage of ssRNA by hydrolysis, the presence of a self-cleaving ribozyme can specifically enhance the rate of hydrolysis for a single bond within ssRNA. The specificity of this cleavage is important in rolling-circle replication since it allows the newly formed linear fragment to circularize and undergo further rolling-circle synthesis. In viroids, this cleavage is catalyzed by hammerhead ribozymes which are roughly 50 nucleotides in length and cleave at a rate of roughly $k_{cleave} \approx 60 - 6000\ h^{-1}$ (Hertel et al. 1994; O'Rourke and Scott 2018). For simplicity, we will assume that there is a contiguous set of 50 nucleotides in the template which encode for a hammerhead ribozyme. When this ribozyme is synthesized without error, the ribozyme cleaves itself in half. However, when errors are incorporated, we will assume that the ribozyme has lost its ability to self-cleave. While a more accurate model of self-cleaving could be incorporated, the specifics of cleavage are not of importance for this paper. The important bit is that when the ribozyme is present, it can cleave to generate new product sequences to undergo further rolling-circle synthesis.

At this point, we have introduced three important rates: k_{cleave}, k_{ext}, and k_{hyd}. Since these rates tend to vary by many orders of magnitude, we can introduce a simplification. For RNA replication to be productive, the rate of primer extension must be faster than the rate of hydrolysis, $k_{ext} > k_{hyd}$. If this was not the case, then synthesis of RNA would not occur past a few nucleotides and replication would not be possible. Based on the rate of hammerhead ribozyme cleavage, it also seems likely that cleavage is faster than extension, $k_{cleave} > k_{ext}$. In this case, we can assume that whenever the self-cleaving ribozyme is present in a ssRNA tail, then it instantly cleaves. If primer extension is slow, as it is predicted to be based on non-enzymatic experiments and ribozymes, then this is a close approximation since any self-cleaving ribozyme will likely cleave before a single primer extension event occurs. Therefore, we will let $k_{cleave} = \infty$, and are now left with the important ratio $R = k_{ext}/k_{hyd}$. This simply means that the rate of hydrolysis sets the timescale of replication, whereas R controls the dynamics of replication. As an example, if the rate of non-enzymatic primer extension was $k_{ext} = 72\ h^{-1}$ (fastest polymerase ribozyme), and hydrolysis was $k_{hyd} = 17 * 10^{-4}\ h^{-1}$ (ssRNA hydrolysis rate when accounting for pH), then $R \approx 40{,}000$. Alternatively, if the rate on non-enzymatic primer extension was $k_{ext} = 9 * 10^{-5}\ h^{-1}$ (non-enzymatic triphosphate ligation), but hydrolysis was $6.1 * 10^{-10}\ h^{-1}$ (ssDNA hydrolysis rate), then $R \approx 150{,}000$. To give a comprehensive overview of these possibilities, we will consider a wide range of R values between 1 and 10^{12}.

Lastly, and perhaps most importantly, we need to consider the relative rate of tail-flipping. When flipping is slow, the model simplifies to the top scheme in Figure 1, which will be referred to hereafter as the slow-flipping limit. In the slow-flipping limit, primer extension occurs at a constant rate of k_{ext} since the 3' end of the product strand is always annealed to the template. Modern rolling-circle replication is expected to follow this limit since primer extension is catalyzed by a protein polymerase which prevents the tail flipping. At the other extreme, when flipping is fast, the template flips between the 3' end annealed state and the 5' end annealed state many times before primer extension occurs. This will be referred to as the fast-flipping limit of the model. In the fast-flipping limit, the fraction of time the template spends in these two flipped states is determined by their difference in free energy. If we let ΔG_3^0 be the standard free energy of the 3' annealed state, and ΔG_5^0 be the energy free energy of the 5' annealed state, then the probability of being in the 3' annealed state, P_3, can be approximated as:

$$P_3 = \frac{\exp(\,-(\Delta G_3^0 - \Delta G_5^0)/k_B T\,)}{1+\exp(-\,(\Delta G_3^0 - \Delta G_5^0)/k_B T\,)},$$

where k_B is the Boltzmann constant, and T is the temperature. In the fast-flipping limit, the net extension rate is therefore $k_{ext} * P_3$ as opposed to the slow-flipping limit which is always k_{ext}. A further consequence of this equilibrium of flipped states if that each bond in the 5' tail hydrolyzes at a net rate of $k_{hyd} * P_3$, and each bond in the 3' tail hydrolyzes at a net rate of $k_{hyd} * (1 - P_3)$, which contrasts to the non-flipping model in which only the 5' tail hydrolyzes at a net rate of k_{hyd} per bond.

In this paper, we will approximate the free energy parameters ΔG_3^0 and ΔG_5^0 based on the nearest-neighbor stacking energy rules between the template and the annealed product sequence (Turner and Mathews 2009). For simplicity, we will ignore the sequence dependence of stacking energy and instead assume that all adjacent Watson-Crick base pairs contribute a constant stacking energy of $\Delta\bar{G}^0_{stack}$ and all errors contribute a stacking energy of $0\ kcal/mol$. In which case, $\Delta G_3^0 = n_3 \Delta\bar{G}^0_{stack}$, and $\Delta G_5^0 = n_5 \Delta\bar{G}^0_{stack}$, where n_3 and n_5 are the number of stacking interactions present in each state respectively. Under these assumptions, the probability of being in the 3' end annealed state is:

$$P_3 = \frac{\exp(-\,(n_3 - n_5) * \Delta\bar{G}^0_{stack}/k_B T\,)}{1+\exp(\,-\,(n_3 - n_5) * \Delta\bar{G}^0_{stack}/k_B T\,)}.$$

When no errors are present in the product strand, $n_3 = n_5$, and $P_3 = 0.5$, implying that the template spends half of its time with the 3' end annealed to the template. When an error is present in the 3' end annealed state, then $n_3 = n_5 - 2$, i.e. two stacking interactions are lost. In which case, $P_3 \approx 0.001$ (assuming $\Delta\bar{G}^0_{stack} = -2\ kcal/mol$, $T = 37°\text{C}$), and the template spends most of its time in the 5' end annealed state. In the fast-flipping limit, the fraction of time spent in the 3' end annealed state changes drastically when errors are incorporated. In this paper, we will assume a constant temperature of 37°C and an average stacking energy of $\Delta\bar{G}^0_{stack} = -2\ kcal/mol$.

The primer purpose of this paper is to show that the rate of tail flipping has a dramatic impact on the fidelity of product synthesis. As such, it is critical to know the relative rate of tail flipping in relation to the other reactions. Based on experimental data, Srinivas et al. 2013 developed a theoretical model for the process of flipping which matches the experimental observations. According to this theory, the process of flipping has two main phases. In the first phase, a branch point forms when the annealed end lifts of the template and the tail replaces the lost base-pair. This occurs at a rate k_{first}. After branch point formation, the branch point migrates through a random walk process at a rate of k_{bm} per step. If the random walk reaches the left boundary, then tail flipping has occurred. If the random walk reaches the right boundary then tail flipping has not occurred, and we must wait for another branch formation event to occur. Based on this process, the net rate of tail flipping is predicted to be:

$$k_{flip} = \frac{1}{l}\left(\frac{1}{k_{first}} + \frac{l-1}{k_{bm}}\right)^{-1}.$$

Where l is the length of the tail. While these rate constants for RNA are unknown, for DNA they are estimated to be: $k_{first} \cong 1.1 * 10^6\ \ h^{-1}$ and $k_{bm} \cong 6.5 * 10^7\ h^{-1}$ (Srinivas et al. 2013). As an example, let us consider a tail length of 10 nucleotides. With this tail length, the rate of flipping is predicted to be $k_{flip} \approx 10^5\ h^{-1}$. For a tail length of 100 nucleotides, the rate of flipping slows down to $k_{flip} \approx 4 * 10^3\ h^{-1}$, which is still $\sim$50x faster than primer extension by polymerase ribozymes. It is not until tail lengths approach 1000 nucleotides that the rate of flipping becomes slower than polymerase ribozymes. Based on these parameters, the fast-flipping limit appears to be relevant under prebiotic conditions whereas the slow-flipping limit is where modern viroids and viruses replicate.

2.2 Modeling the Dynamics of Tail Length, Hydrolysis, and Cleavage

In this section, we will begin to uncover the key differences between the slow-flipping and fast-flipping limits of rolling-circle synthesis. Using the previously defined model, we implemented a simple Gillespie algorithm (see Materials and Methods) which tracks the synthesis of RNA on a single template undergoing rolling-circle RNA synthesis.

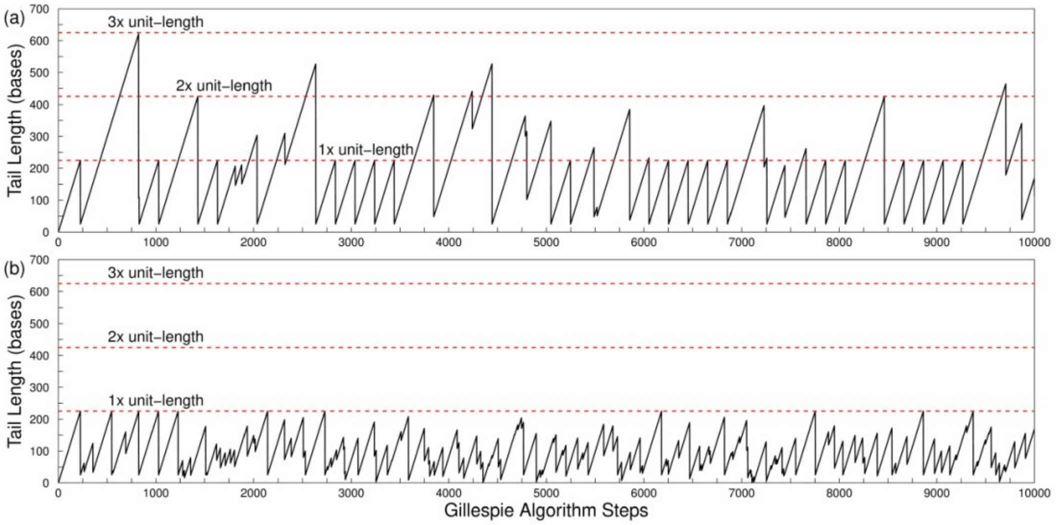


Figure 2: Tail length as a function of Gillespie algorithm steps for the (a) slow-flipping and (b) fast-flipping limits for a template containing 200 base-pairs. Dashed horizontal lines are shown at lengths of 225, 425, and 625 bases corresponding to cleavage events resulting in unit-length, 2x unit-length, and 3x unit-length products, respectively. In the slow-flipping limit, hydrolysis events are rare and genome duplication or triplication events are observed. In the fast-flipping limit, hydrolysis events are frequent, and no genome duplication events are observed. Parameters used: $\Delta \bar{G}^{0}_{stack} = -2\ kcal/mol$, $T = 37$°C, $q_{ext} = 0.99$, and $R = 10^5$.

We will start our analysis by considering a specific set of parameters and illustrate how the length of the ssRNA tail changes as a function of time. In figure 2, tail length is shown as a function of Gillespie algorithm steps in the specific case of $R = \left(\frac{k_{ext}}{k_{hyd}}\right) = 10^5$, $q_{ext} = 0.99$, and a template of length 200. In the slow-flipping limit (top panel), a regular pattern of tail growth and cleavage is observed. Under this set of parameters, hydrolysis of the tail is rare, and most tail cleavage events correspond to ribozyme activity. Ribozyme cleavage is observed to occur at a tail length of 225 nucleotides, immediately after the synthesis of a self-cleaving ribozyme. This corresponds to the unit-length template of 200 bases, and half of a self-cleaving ribozyme at 25 bases. After cleavage, a unit-length product of 200 nucleotide is formed, and 25 nucleotides are left in the tail, corresponding to half of the self-cleaving ribozyme. We also observe ribozyme cleavage at lengths of $225 + 200n$, where n is any integer greater than 0. The extended lengths of these products are the result of errors incorporated into self-cleaving ribozymes which prevented their cleavage. When considering a more complete model of replication, these products could result in whole genome duplication and be an important mechanism of increasing genome size and complexity.

In the fast-flipping limit, the time evolution of the tail length is strikingly different than the slow-flipping limit, despite the same set of parameters being used (Figure 2, bottom panel). The first observable difference is the highly "jagged" nature of the tail length resulting from random hydrolysis events. To be clear, this is not due to an increase in the rate of hydrolysis, but conversely, is due to the decrease in the net rate of primer extension. In the fast-flipping limit, the net rate of primer extension is $P_3 * k_{ext}$, where P_3 is the probability that the template is in the 3' end annealed state. Prior to the incorporation of any errors, $P_3 = 0.5$, and thus the extension is only marginally slower than the slow-flipping limit. However, after incorporation of an error, P_3 drops precipitously. By flipping into the 5' end annealed state, the incorporated error is forced into the 3' tail and two stacking interactions are recovered. The result of which is a $P_3 \approx 0.001$, and a 500-fold reduction in the net rate of primer extension. If additional errors are added, P_3 further decreases, and primer extension grinds to a halt. The result of which is an increased likelihood of hydrolysis as opposed to further primer extension.

The second observable difference is the absence of tail lengths greater than 225 nucleotides. The reason for this difference is simply due to the removal of all errors before they reached the 5' tail, and thus all self-cleaving ribozymes maintained their cleavage activity. To illustrate this, let us consider what is occurring at the Gillespie algorithm step following the error (see Figure 3). Assuming $P_3 \approx 0.001$, as calculated previously, we can determine the probability of primer extension, hydrolysis of the 5' tail, and hydrolysis of the 3' tail being the next reaction. If we let the length of the tail, L_T, be 100 nucleotides after the error is incorporated, then the probability that the next reaction is primer extension is:

$$Pr(primer\ extension) = \frac{R * P_3}{R * P_3 + L_T} = 0.5,$$

whereas the probability of 5' tail hydrolysis is:

$$Pr(5'\ tail\ hydrolysis) = \frac{L_T * P_3}{R * P_3 + L_T} = 0.005,$$

and the probability of 3' tail hydrolysis is:

$$Pr(3'\ tail\ hydrolysis) = \frac{L_T * (1 - P_3)}{R * P_3 + L_T} = 0.495.$$

With near equal probability, the next reaction will be either primer extension by one nucleotide or hydrolysis of the 3' tail. Since the error is located at the end of the 3' tail, any hydrolysis event occurring in the 3' tail will remove the error, resulting in error correction. Therefore, in a single step, the probability of error correction is $Pr(3'\ tail\ hydrolysis) \approx 0.5$. Even if the next reaction is primer extension, these probabilities remain nearly unchanged for subsequent steps. The probability of this error surviving for N extensions is therefore $\sim 0.5^N$, which becomes vanishingly small for large N. In contrast to the fast-flipping limit, the slow-flipping limit cannot undergo error correction as the 5' end annealed state is unreachable. In the slow-flipping limit, $Pr(3'\ tail\ hydrolysis) = 0$, hydrolysis cannot remove the error, and thus no error-correction occurs.

Based on the equations from this example, we would predict that the extent of error-correction is dependent on both R and the length of the template. When R is small, error correction should be more likely. Whereas in the limit of $R \to \infty$, no error correction is possible as only primer extension can occur. Similarly, as the template length increases, the length of the tail must also increase. When the tail is short, the net rate of hydrolysis is slower, and thus error-correction is less likely. Whereas when the tail is long, the net rate of hydrolysis is faster, and error-correction is more likely.

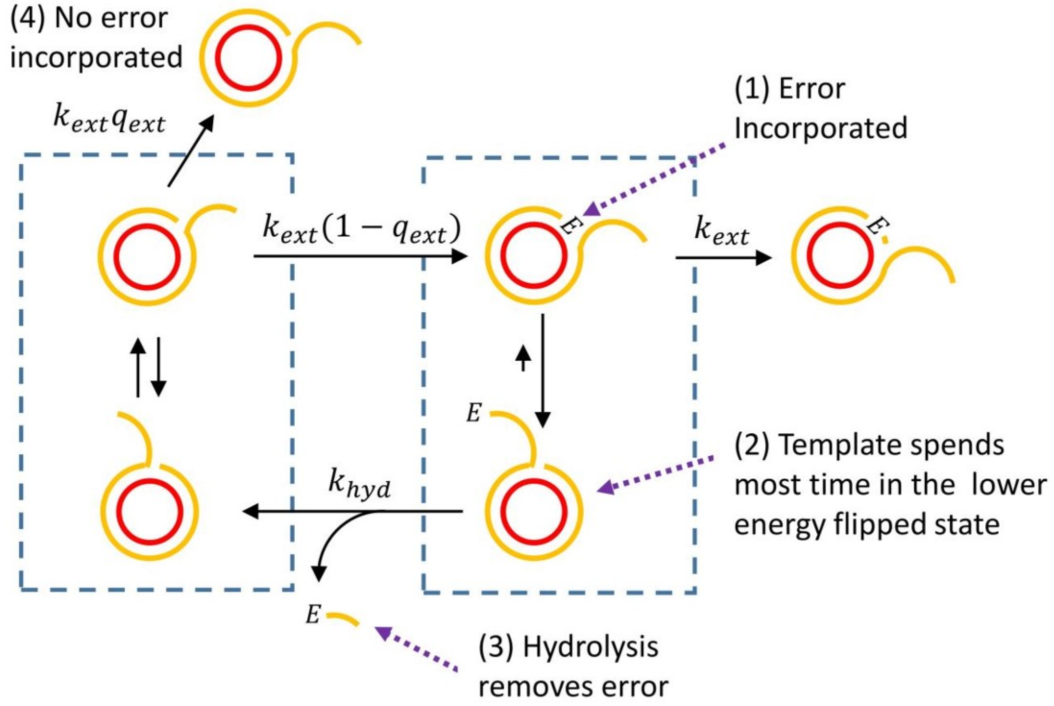

(4) No error incorporated
$k_{ext}q_{ext}$
$k_{ext}(1-q_{ext})$
(1) Error Incorporated
E
k_{ext}
E
E
(2) Template spends most time in the lower energy flipped state
k_{hyd}
E
(3) Hydrolysis removes error

Figure 3: Simplified diagram illustrating the error-correction cycle. After an error is incorporated (1), the template preferentially spends most of its time in the lower energy flipped state (2). This difference in energy is due to the thermodynamic penalty of having the error E annealed to the template. From the lower energy state, hydrolysis of the 3' tail becomes more likely. In which case, the error is removed (3) and the template is once again in the unbiased non-error state. Further extension can then result in correct incorporation (4), or incorrect incorporation (1), in which case the error-correction cycle is repeated.

2.3 Non-enzymatic Error Correction

In this section, we will expand our focus and determine the extent of error-correction as a function of R, q_{ext}, and template length. Since we are primarily interested in showing that this error-correction avoids Eigen's paradox, we will report our results in terms of the per-base product fidelity, q_{prod}. When no error-correction is present, $q_{prod} = q_{ext}$, i.e. the fidelity of product sequences is the same as the fidelity of extension. Whereas error correction results in $q_{prod} >$

q_{ext}, i.e. errors are removed before incorporation into a product sequence. To get an accurate measurement of the per-base product fidelity, q_{prod}, we restrict our analysis to unit-length product sequences, particularly those bases which are not part of the self-cleaving ribozyme. These bases are free to mutate and thus do not bias our measurement of per-base product fidelity.

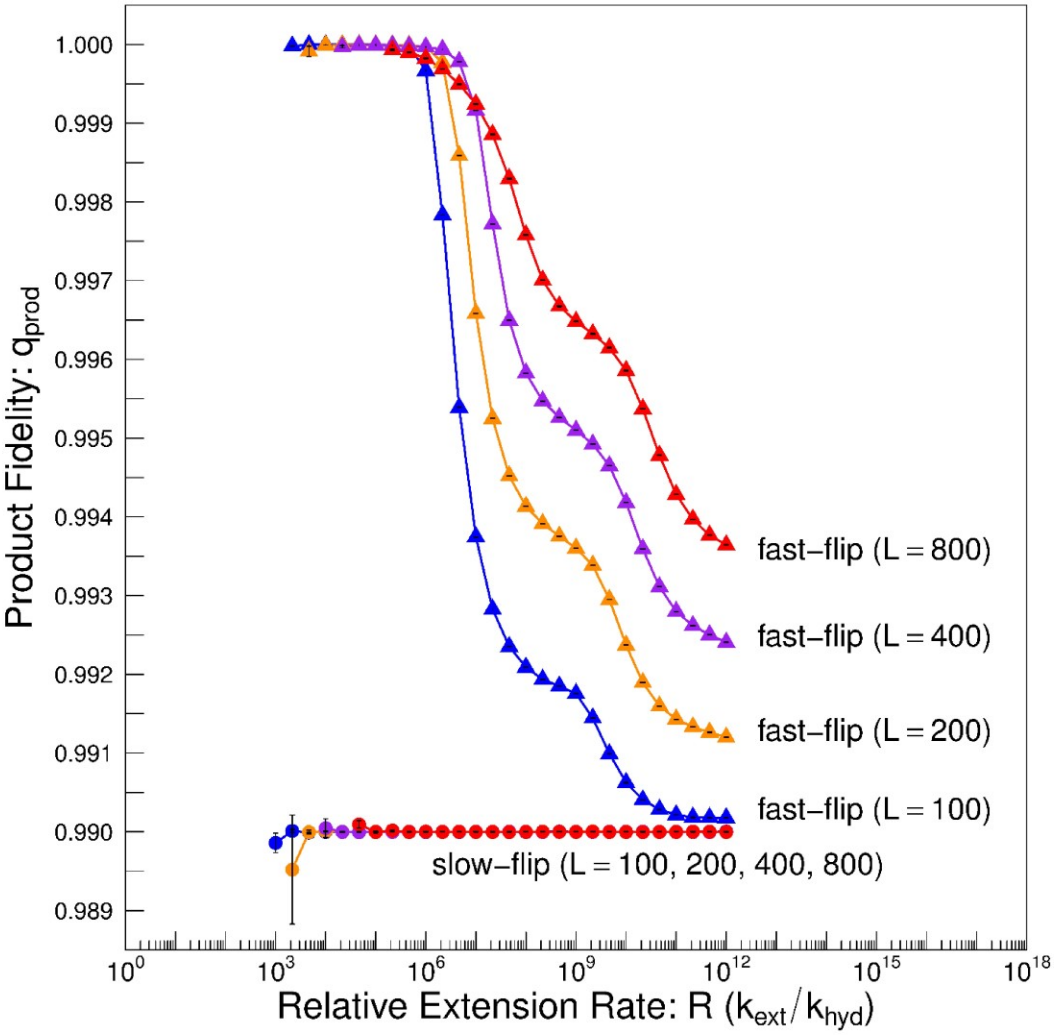


Figure 4: The per-base product fidelity, q_{prod}, is shown as a function of R, the ratio of extension rate to hydrolysis rate. In the slow-flipping limit, the per-base product fidelity is the same as the extension fidelity, $q_{prod} = q_{ext}$. In the fast-flipping limit, the per-base product fidelity in greater, $q_{prod} \gg q_{ext}$. When R is small, e.g. primer extension is slow, $q_{prod} \approx 1$. In this regime, error-correction dominants. When R is large, e.g. primer extension is fast, q_{prod} approaches q_{ext}. The extent of error-correction is also observed to be dependent on template length, with longer

templates achieving higher q_{prod} than shorter templates. Error bars represent the standard error in the mean. Parameters used: $\Delta\bar{G}^0_{stack} = -2\ kcal/mol$, $T = 37$°C, and $q_{ext} = 0.99$.

In Figure 4, the per-base fidelity of product sequences, q_{prod}, is plotted as a function of increasing R, assuming a constant extension fidelity of $q_{ext} = 0.99$. In the slow-flipping limit, the per-base product fidelity is the same as the per-base extension fidelity, and is independent of template length, i.e. $q_{prod} = q_{ext}$. This is to be expected as there is no error-correction occurring in the slow-flipping limit. Note that for small values of R, complete synthesis of unit-length product sequences becomes unlikely due to the relatively fast rate of hydrolysis. As such, the error bars for our measurements are higher in this regime, and for longer templates, an accurate value cannot be measured. The net rate of product synthesis will be discussed later.

In the fast-flipping limit, the per-base fidelity of product sequences is dependent on both R and template length (Figure 4), as was predicted from the example in the previous section. When R is small ($R \leq 10^5$), e.g. primer extension is slow, nearly all incorporated errors are removed by error correction. For reference, experimental rates of ribozyme-catalyzed extension result in $R \approx 10^4 - 10^5$, and thus fall within this $q_{prod} \approx 1$ regime. Therefore, if prebiotic synthesis were occurring at a similar rate, we would expect highly accurate synthesis. When plotted on a log axis, we actually observe a maximum fidelity of $q_{prod} \approx 0.99999 - 0.999999$, and a decrease in q_{prod} for very small values of R (Figure S1). This decrease is due to the complex way in which an incorporated error modifies P_3, and provides a slight benefit to the synthesis of product sequences containing an error. However, even when R is very small, $q_{prod} > 0.9999$, and thus any error catastrophe would be easily avoided. As R increases, e.g. primer extension becomes faster, q_{prod} decreases but always remains higher than q_{ext}, i.e. $q_{prod} > q_{ext}$. In the limit of $R \rightarrow \infty$, the simulations verify that $q_{prod} = q_{ext}$ (data not shown). This limit of R approaching infinity, however, has no basis in the chemistry of RNA. For all prebiotically and biologically relevant values of R, error-correction is predicted, and thus $q_{prod} > q_{ext}$.

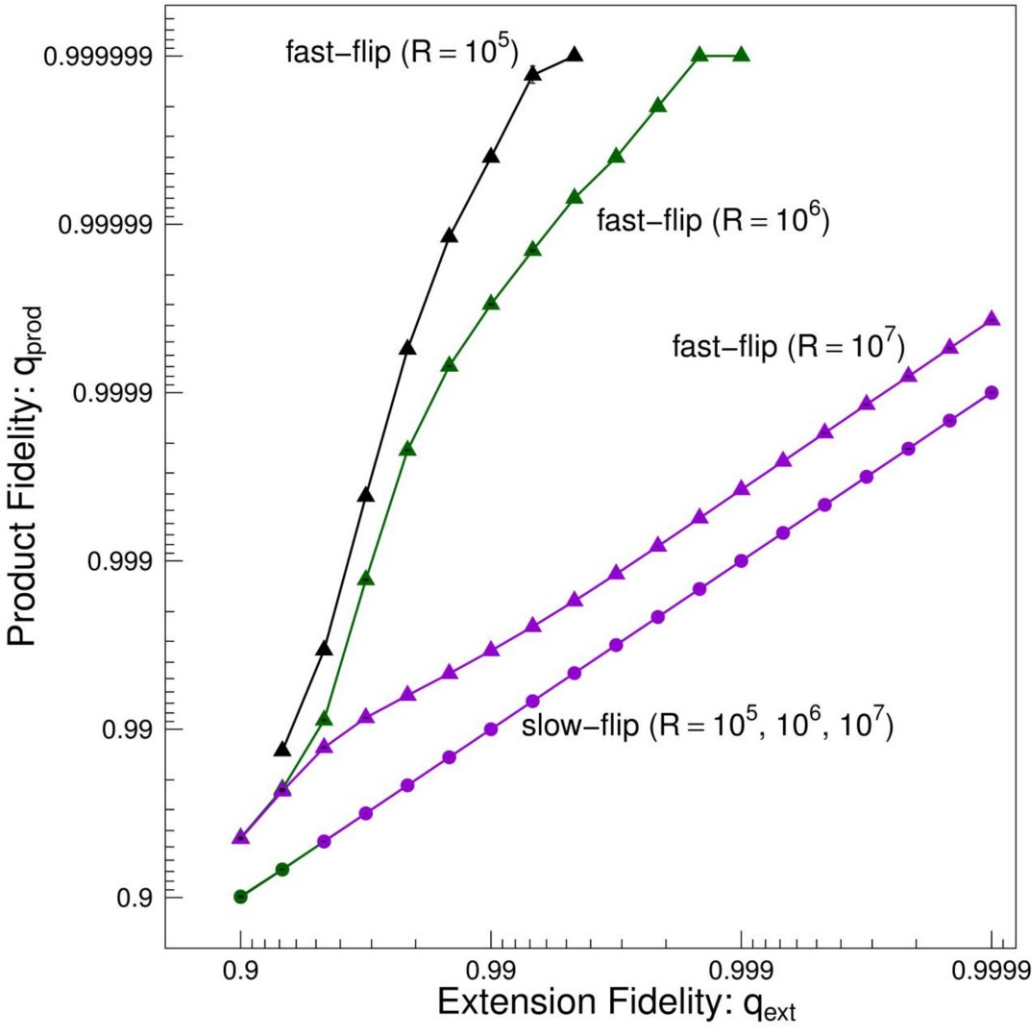


Figure 5: The per-base product fidelity, q_{prod}, is shown as a function of the per-base extension fidelity, q_{ext}, for a fixed template length of 200 nucleotides. In the slow-flipping limit, the per-base product fidelity is the same as the extension fidelity, $q_{prod} = q_{ext}$, regardless of the value of R. In the fast-flipping limit, the per-base product fidelity in greater, $q_{prod} \gg q_{ext}$. For reference, fidelity in experiments is measures at $q_{ext} \approx 0.99$, and $R \approx 10^4 - 10^5$ for the faster known polymerase ribozyme. Error bars represent the standard error in the mean. Parameters used: $\Delta\bar{G}^0_{stack} = -2\frac{kcal}{mol}$, $T = 37$°C, and $L = 200$.

We also considered alternative values of q_{ext} ranging from 0.9 to 0.9999, for a fixed template length containing 200 bases (Figure 5). In the slow-flipping limit, no error correction is observed and $q_{prod} = q_{ext}$ regardless of q_{ext} and R. In the fast-flipping limit, increasing q_{ext} results in proportionally higher q_{prod}. Note that for extremely high q_{ext}, the curves appear to abruptly end. This is due to the simulations providing a q_{prod} estimate of 1 which cannot be plotted on these log axes. Decreasing q_{ext} results in lower q_{prod} since error-correction struggles to keep up with the enhanced rate of error incorporation. Nevertheless, even with poor extension fidelity, $q_{prod} > q_{ext}$. When considering the experimental estimate of $q_{ext} \approx 0.99$ and $R \approx 10^4 - 10^5$, q_{prod} is expected to be sufficiently high to avoid an error catastrophe.

The error-correction described here for the fast-flipping limit relies on hydrolysis to remove the error. Since this hydrolysis can be relatively slow, the net rate of product synthesis differs in the two limits of slow and fast-flipping. In the slow-flipping limit, the rate of product synthesis, k_{prod}, is expected to be proportional to the rate of primer extension: $k_{prod} \propto k_{ext}/L$. Or equivalently, the relative rate of product synthesis, R_{prod}, is proportional to R: $R_{prod} \propto R/L$. In figure 5, the relative rate of product synthesis, R_{prod}, is plotted as a function of increasing R. When R is large, hydrolysis is comparatively rare, and this relationship is observed. Note that even in the limit of $R \rightarrow \infty$, $R_{prod} < R/L$ since poor extension fidelity limits the formation of a unit-length product. When R is small, the rate of product synthesis becomes limited by hydrolysis. If R is too small, then the rate of product synthesis becomes too slow to measure accurately, and thus the curves appear to end abruptly.

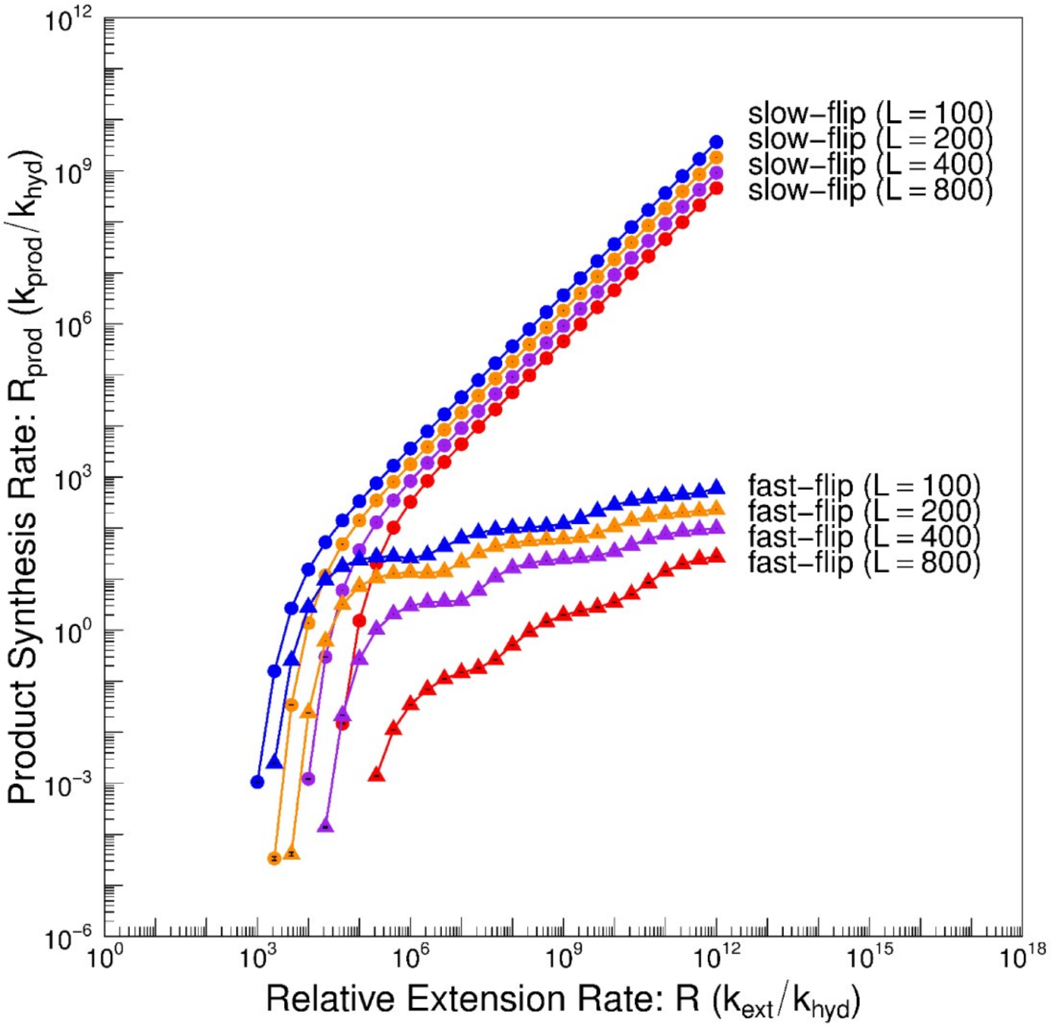


Figure 6: The relative rates of product synthesis, R_{prod}, are shown as a function of R for the slow and fast-flipping limits. In both limits, the rate of product synthesis approaches zero when R is small due to the relatively fast rate of hydrolysis. In the slow-flipping limit, the rate of product synthesis scales linearly with extension rate for large R. In the fast-flipping limit, the rate of product synthesis increases slowly with R since error correction and incorporated errors slow primer extension. Error bars represent the standard error in the mean. Parameters used: $\Delta\bar{G}^{0}_{stack} = -2\ kcal/mol$, $T = 37$°C, and $q_{ext} = 0.99$.

In contrast to the slow-flipping limit, the rate of product synthesis in the fast-flipping limit is more complex. When R is small, the rate of product synthesis in the fast-flipping limit is similar to the slow-flipping limit. In this regime, hydrolysis is fast, and error-correction is not rate limiting. In this regime, the observed difference in product synthesis rates stems from the 2x slowdown in the extension rate by the factor P_3. When accounting for this difference, the slow and fast-flipping limits converge to the same values. As R increases, error-correction becomes rate limiting as hydrolysis events are comparatively slow. When R is sufficiently high ($R > 10^6$), error-correction is unable to efficiently remove incorporated error. In this regime, the average rate of primer extension, $k_{ext} * P_3$, increases very slowly since P_3 tends to decrease with increasing k_{ext}. Interestingly, in the fast-flipping limit, increasing the fidelity of primer extension provides a relatively large rate enhancement to product synthesis (Figure 7). For instance, at $q_{ext} = 0.99$, increase R by a factor of a 10, increases R_{prod} roughly 2-fold. Whereas, decreasing the error rate by a factor of 10, equivalent to increasing the extension fidelity from $q_{ext} = 0.99$ to $q_{ext} = 0.999$, results in a ~20-fold increase in the rate of product synthesis. Therefore, if tail-flipping is fast, a ribozyme which enhances the fidelity, as opposed to the rate of primer extension, may provide a much larger replicative advantage for a template.

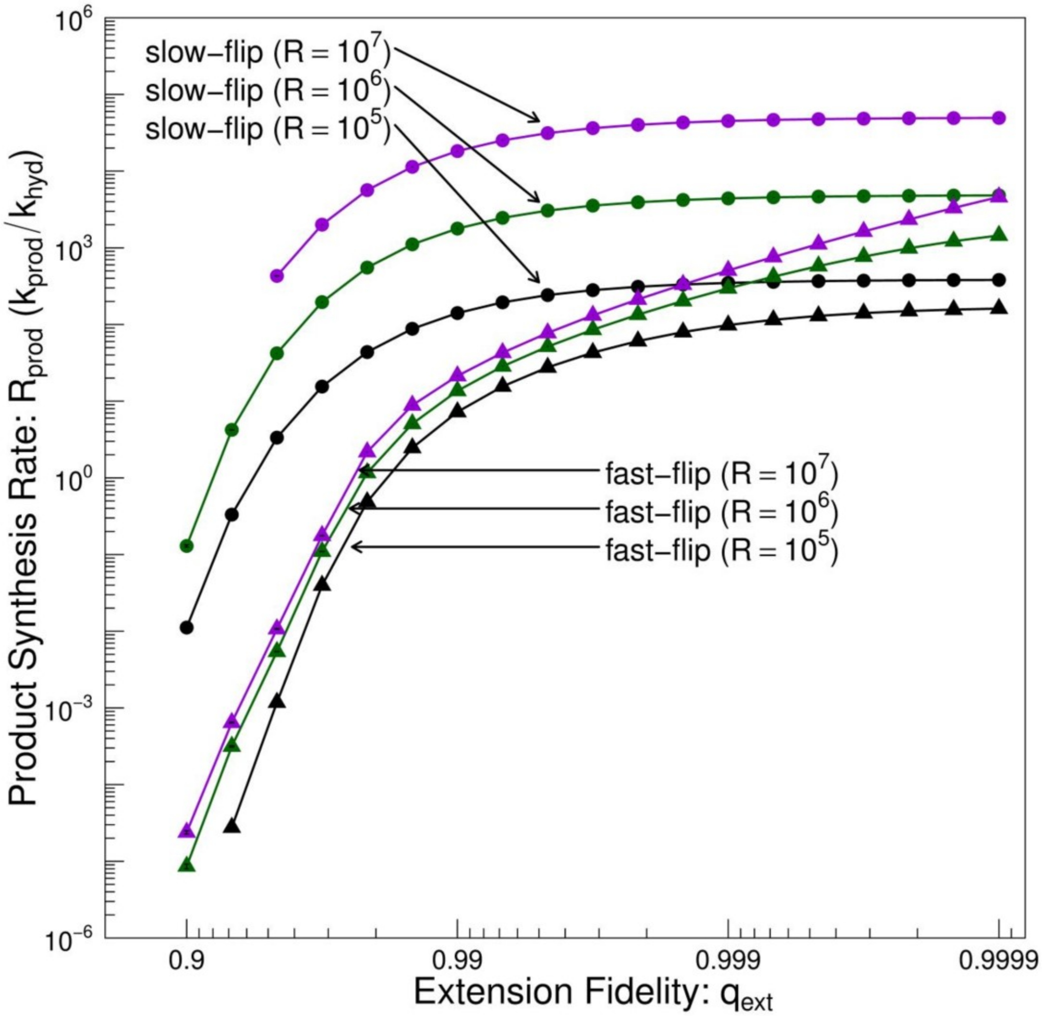


Figure 7: The relative rate of product synthesis, R_{prod}, is plotted as a function of q_{ext}, the fidelity of primer extension. In both limits, the rate of product synthesis increases with fidelity due to the higher probability of synthesizing an intact self-cleaving ribozyme. In the fast-flipping limit, the rate of product synthesis is limited by the process of error-correction. In this limit, increasing the fidelity of extension, q_{ext}, has a much greater impact on R_{prod} than increasing R. Error bars represent the standard error in the mean. Parameters used: $\Delta\bar{G}^0_{stack} = -2\ kcal/mol$, $T = 37$°C, and $L = 200$.

3. Discussion

The model used in the paper is simplistic and designed to show the importance of tail-flipping during rolling-circle synthesis. Even with this simplicity, there are vast variations in the parameter set which were not thoroughly explored due to computation constraints. In the figures shown here, we assumed a constant $\Delta\bar{G}^0_{stack}$ of -2 kcal/mol, which is consistent with stacking energies in the nearest neighbor model (Turner and Mathews 2009). Under prebiotic conditions, the value of $\Delta\bar{G}^0_{stack}$ would depend on the sequence identity of the template, with a larger GC content resulting in a higher $\Delta\bar{G}^0_{stack}$. We also tested smaller and larger values of $\Delta\bar{G}^0_{stack}$ with predictable results. Since $\Delta\bar{G}^0_{stack}$ only changes the value of P_3 after incorporation of an error, it has no impact on the slow-flipping limit. In the fast-flipping limit, a large $\Delta\bar{G}^0_{stack}$ results in more error correction and thus higher product fidelity. Conversely, smaller values of $\Delta\bar{G}^0_{stack}$ result in less error-correction and worse fidelity. Even for smaller values of $\Delta\bar{G}^0_{stack}$, the resulting product fidelity appears to be sufficiently high to avert an error catastrophe when R is small. Similar to $\Delta\bar{G}^0_{stack}$, changing the temperature would also change the value of P_3 and the extent of error-correction. With high temperatures providing less error-correction than cold temperatures. However, changing the temperature would also change the rates of primer extension, fidelity, hydrolysis, and tail-flipping. As such, the simple model used here is unable to capture the importance of temperature.

We also tested a more complex version of the model which included the experimentally observed stalling and poor fidelity after the incorporation of an error (Rajamani et al. 2010; Leu et al. 2013). These additions are not essential for error-correction in the fast-flipping limit; however, they are beneficial. After an error is incorporated, the value of R is artificially smaller and thus error-correction is more likely. In essence, product synthesis proceeds faster due to the higher R, while achieving the enhanced fidelity of a smaller R. Similarly, the poor fidelity after an error enhances the likelihood two adjacent errors being incorporated. Since this introduces an additional stacking penalty in the 3' end annealed state, P_3 is even smaller, and error correction is likely to remove both errors with one hydrolysis event. Therefore, the poor fidelity after incorporation of an error can paradoxically increase the synthesis fidelity.

There are likely to be prebiotic conditions which prevent tail-flipping, and thus prevent error correction. However, these conditions should not detract from the conditions which allow fast tail-flipping. If the prebiotic conditions make tail-flipping slow, then emerging life would be evolutionarily stuck due to the poor fidelity of replication and Eigen's paradox. Whereas life emerging under fast-flipping conditions would be free to evolve and find new ribozymes without succumbing to an error catastrophe. As such, it is critical to understand the conditions which allow for fast tail-flipping. Currently, our estimate for the rate of tail flipping rate is based on the measurements of toehold-mediated strand displacement and its underlying theory (Zhang and Winfree 2009; Srinivas et al. 2013; Šulc et al. 2015). These measurements, however, are based on relatively short tail lengths without errors. Presumably, the observed rate of tail-flipping would be faster immediately after incorporation of an error due to the lower free energy penalty of branch point formation. In which case, the rate of tail-flipping may be faster than assumed here, and thus the fast-flipping limit may be more likely. However, the experiments on toehold-mediated strand displacement also omit folded structures. Presumably, the rate of tail-flipping would decrease with length if the sequence can adopt a folded structure. In an extreme case, the folded structure could be so stable to prevent flipping entirely. However, these folded structures would also be unlikely to act as a template for further replication. At this stage of evolution, there may be a selective pressure against folded structures as they hinder replication, unless of course, the folded structure catalyzes a reaction which enhances replication.

Polymerase ribozymes are often considered to be an important early ribozyme, in the hope that they can mimic the speed and fidelity of modern protein polymerases. Modern protein polymerases, however, are highly complex enzymes, often catalyzing both primer extension as well as nuclease activity to remove incorporated errors. If the fidelity of replication under non-enzymatic conditions is high, as it appears to be in the fast-flipping limit, then the slow emergence of a complex polymerase through refinement of structure may be possible. Based on the results of the fast-flipping limit, we find that there would be relatively little benefit of evolving an exceedingly fast polymerase as this prevents error-correction and provides little enhancement to the net rate of product synthesis. Whereas the evolution of a polymerase which enhances the fidelity of primer extension, which in turn, enhances the rate of replication, would be of much greater benefit. Through this evolutionary path, the fidelity of replication continues to increase, while the need for error-correction decreases. When the extension fidelity becomes sufficiently high, the error-correction from flipping is no longer needed, and the slow-flipping limit is survivable.

Interestingly, the presence of error-correction appears to allow for the emergence of completely different type of ribozyme. Instead of enhancing the rate or accuracy of bond formation, the new ribozyme could enhance the rate of bond breaking. Since the newly incorporated error is always on the 3' end, any ribozyme which specifically cleaves the 3' end will enhance the rate of error-correction. This would decouple the error-correction's reliance on hydrolysis and would enhance the net rate of product synthesis. We should not be quick to dismiss the existence of such a ribozyme as it parallels the modern RNase P, which is one of the two universal conserved ribozymes in modern biology. If such a ribozyme emerged, then its continued evolution could have allowed it to remove the error while the error is still annealed to the template. In which case, tail-flipping would no longer be required to maintain a high replication accuracy. This evolutionary path would be akin to the emergence of the nuclease activity of a polymerase prior to the emergence of its primer extension activity.

Regardless of the evolutionary path which emerging life took, the error-correction present in the fast-flipping limit would have circumvented Eigen's paradox. Initially, the emerging life would have been replicating with high fidelity under non-enzymatic conditions due to error-correction. With high fidelity replication, the emergence of a complex ribozyme could have been a gradual process where ribozyme activity is slowly improved and refined over time. Importantly, any ribozyme which prevented the error-correction, without first increasing the fidelity of replication, would be lost due to the inevitable error catastrophe. In this scenario, the emerging life is trapped in a state of high-fidelity replication. Mutant lineages lacking error-correction are lost due to the ensuing error catastrophe. Whereas mutant lineages which enhance the fidelity of replication, and indirectly the rate of replication, tend to become fixed in the population. The result of which is a population of RNA organisms which are evolutionarily driven towards high fidelity replication.

4. Materials and Methods

The simplicity of this model allows it to be easily implemented according to the Gillespie algorithm (Gillespie 1977). At each algorithm step, the length of the tail sequence was determined, and in the fast-flipping limit, the number of stacking interactions for each flipped state. Based on the reaction rate parameters, and the current state of the template, each reaction was given a probability of occurring in proportion to its rate and one was chosen at random. When the next reaction was determined to be primer extension, a subsequent random number determined whether the added nucleotide was correct or incorrect. For instantaneous self-cleaving of synthesized ribozymes, the reaction rate was set to be infinite such that if self-cleavage could occur, it would always be the next reaction to occur.

Simulations were run for a total of 1 billion Gillespie algorithm steps and the product strands generated were analyzed for fidelity and the simulation time at which synthesis completed was recorded. Average fidelity was then calculated by summing the errors found in all unit-length product strands, omitting the region containing the self-cleaving ribozyme. This provides a simulation estimate of product fidelity. Similarly, the average rate of product synthesis was calculated by summing the inverse of times between product synthesis events. To avoid biasing our measurements, the first product strand produced was ignored as its fidelity is unknown. If the product was the result of annealing, it would have fidelity q_{prod}, whereas if it was synthesized *de novo* it would have fidelity of q_{ext}. The results shown assumed that the initial product annealed to the template was synthesized with fidelity of unity. However, the same results are observed when starting with a product of fidelity q_{ext}. This provides reasonable confidence that the reported results are independent of the starting conditions. Each set of simulation parameters were replicated 10 times and the reported value is the average of simulation estimates, with error bars corresponding to the standard error in the mean.

References

Attwater J, Raguram A, Morgunov AS, Gianni E, Holliger P. 2018. Ribozyme-catalysed RNA synthesis using triplet building blocks. Elife. 7. doi:10.7554/eLife.35255.

Attwater J, Wochner A, Holliger P. 2013. In-ice evolution of RNA polymerase ribozyme activity. Nat Chem. 5(12):1011–1018. doi:10.1038/nchem.1781.

Eigen M, McCaskill J, Schuster P. 1988. Molecular quasi-species. J Phys Chem. 92(24):6881–6891. doi:10.1021/j100335a010.

Gillespie DT. 1977. Exact stochastic simulation of coupled chemical reactions. J Phys Chem. 81(25):2340–2361. doi:10.1021/j100540a008.

Hertel KJ, Uhlenbeck OC, Herschlag D. 1994. A Kinetic and Thermodynamic Framework for the Hammerhead Ribozyme Reaction. Biochemistry. 33(11):3374–3385. doi:10.1021/bi00177a031.

Higgs PG. 2017. Chemical Evolution and the Evolutionary Definition of Life. J Mol Evol. 84(5–6):225–235. doi:10.1007/s00239-017-9799-3.

Higgs PG, Lehman N. 2015. The RNA World: Molecular cooperation at the origins of life. Nat Rev Genet. 16(1):7–17. doi:10.1038/nrg3841.

Horning DP, Joyce GF. 2016. Amplification of RNA by an RNA polymerase ribozyme. Proc Natl Acad Sci U S A. 113(35):9786–9791. doi:10.1073/pnas.1610103113.

Jeffares DC, Poole AM, Penny D. 1998. Relics from the RNA world. J Mol Evol. 46(1):18–36. doi:10.1007/PL00006280.

Joyce GF. 2002. The antiquity of RNA-based evolution. Nature. 418(6894):214–221. doi:10.1038/418214a.

Kun Á, Santos M, Szathmáry E. 2005. Real ribozymes suggest a relaxed error threshold. Nat Genet. doi:10.1038/ng1621.

Kun Á, Szilágyi A, Könnyu B, Boza G, Zachar I, Szathmáry E. 2015. The dynamics of the RNA world: Insights and challenges. Ann N Y Acad Sci. doi:10.1111/nyas.12700.

Leu K, Kervio E, Obermayer B, Turk-Macleod RM, Yuan C, Luevano JM, Chen E, Gerland U, Richert C, Chen IA. 2013. Cascade of reduced speed and accuracy after errors in enzyme-free copying of nucleic acid sequences. J Am Chem Soc. 135(1):354–366. doi:10.1021/ja3095558.

Leu K, Obermayer B, Rajamani S, Gerland U, Chen IA. 2011. The prebiotic evolutionary advantage of transferring genetic information from RNA to DNA. Nucleic Acids Res. doi:10.1093/nar/gkr525.

Li Y, Breaker RR. 1999. Kinetics of RNA degradation by specific base catalysis of transesterification involving the 2γ-hydroxyl group. J Am Chem Soc. 121(23):5364–5372. doi:10.1021/ja990592p.

O'Rourke SM, Scott WG. 2018. Structural Simplicity and Mechanistic Complexity in the Hammerhead Ribozyme. Prog Mol Biol Transl Sci. 159:177–202. doi:10.1016/bs.pmbts.2018.07.006.

Olsen TJ, Choi Y, Sims PC, Gul OT, Corso BL, Dong C, Brown WA, Collins PG, Weiss GA. 2013. Electronic measurements of single-molecule processing by DNA polymerase i (Klenow fragment). J Am Chem Soc. 135(21):7855–7860. doi:10.1021/ja311603r.

Prywes N, Blain JC, Del Frate F, Szostak JW. 2016. Nonenzymatic copying of RNA templates containing all four letters is catalyzed by activated oligonucleotides. Elife. 5(JUN2016). doi:10.7554/eLife.17756.

Radzicka A, Wolfenden R. 1995. A proficient enzyme. Science (80-). 267(5194):90–93. doi:10.1126/science.7809611.

Rajamani S, Ichida JK, Antal T, Treco DA, Leu K, Nowak MA, Szostak JW, Chen IA. 2010. Effect of stalling after mismatches on the error catastrophe in nonenzymatic nucleic acid replication. J Am Chem Soc. 132(16):5880–5885. doi:10.1021/ja100780p.

Rohatgi R, Bartel DP, Szostak JW. 1996a. Kinetic and mechanistic analysis of nonenzymatic, template-directed oligoribonucleotide ligation. J Am Chem Soc. 118(14):3332–3339. doi:10.1021/ja953712b.

Rohatgi R, Bartel DP, Szostak JW. 1996b. Nonenzymatic, template-directed ligation of oligoribonucleotides is highly regioselective for the formation of 3′-5′ phosphodiester bonds. J Am Chem Soc. 118(14):3340–3344. doi:10.1021/ja9537134.

Smith JM. 1983. Models of evolution. Proc R Soc London - Biol Sci. 219(1216):315–325. doi:10.1098/rspb.1983.0076.

Soukup GA, Breaker RR. 1999. Relationship between internucleotide linkage geometry and the stability of RNA. RNA. doi:10.1017/S1355838299990891.

Srinivas N, Ouldridge TE, Šulc P, Schaeffer JM, Yurke B, Louis AA, Doye JPK, Winfree E. 2013. On the biophysics and kinetics of toehold-mediated DNA strand displacement. Nucleic Acids Res. 41(22):10641–10658. doi:10.1093/nar/gkt801.

Šulc P, Ouldridge TE, Romano F, Doye JPK, Louis AA. 2015. Modelling toehold-mediated RNA strand displacement. Biophys J. 108(5):1238–1247. doi:10.1016/j.bpj.2015.01.023.

Takeuchi N, Poorthuis PH, Hogeweg P. 2005. Phenotypic error threshold; additivity and epistasis in RNA evolution. BMC Evol Biol. 5. doi:10.1186/1471-2148-5-9.

Tam CP, Fahrenbach AC, Björkbom A, Prywes N, Izgu EC, Szostak JW. 2017. Downstream oligonucleotides strongly enhance the affinity of GMP to RNA primer-template complexes. J Am Chem Soc. 139(2):571–574. doi:10.1021/jacs.6b09760.

Turner DH, Mathews DH. 2009. NNDB: The nearest neighbor parameter database for predicting stability of nucleic acid secondary structure. Nucleic Acids Res. 38(SUPPL.1). doi:10.1093/nar/gkp892.

Walton T, Pazienza L, Szostak JW. 2019. Template-Directed Catalysis of a Multistep Reaction Pathway for Nonenzymatic RNA Primer Extension. Biochemistry. 58(6):755–762. doi:10.1021/acs.biochem.8b01156.

Zhang DY, Winfree E. 2009. Control of DNA strand displacement kinetics using toehold exchange. J Am Chem Soc. 131(47):17303–17314. doi:10.1021/ja906987s.

Supplementary Figures:

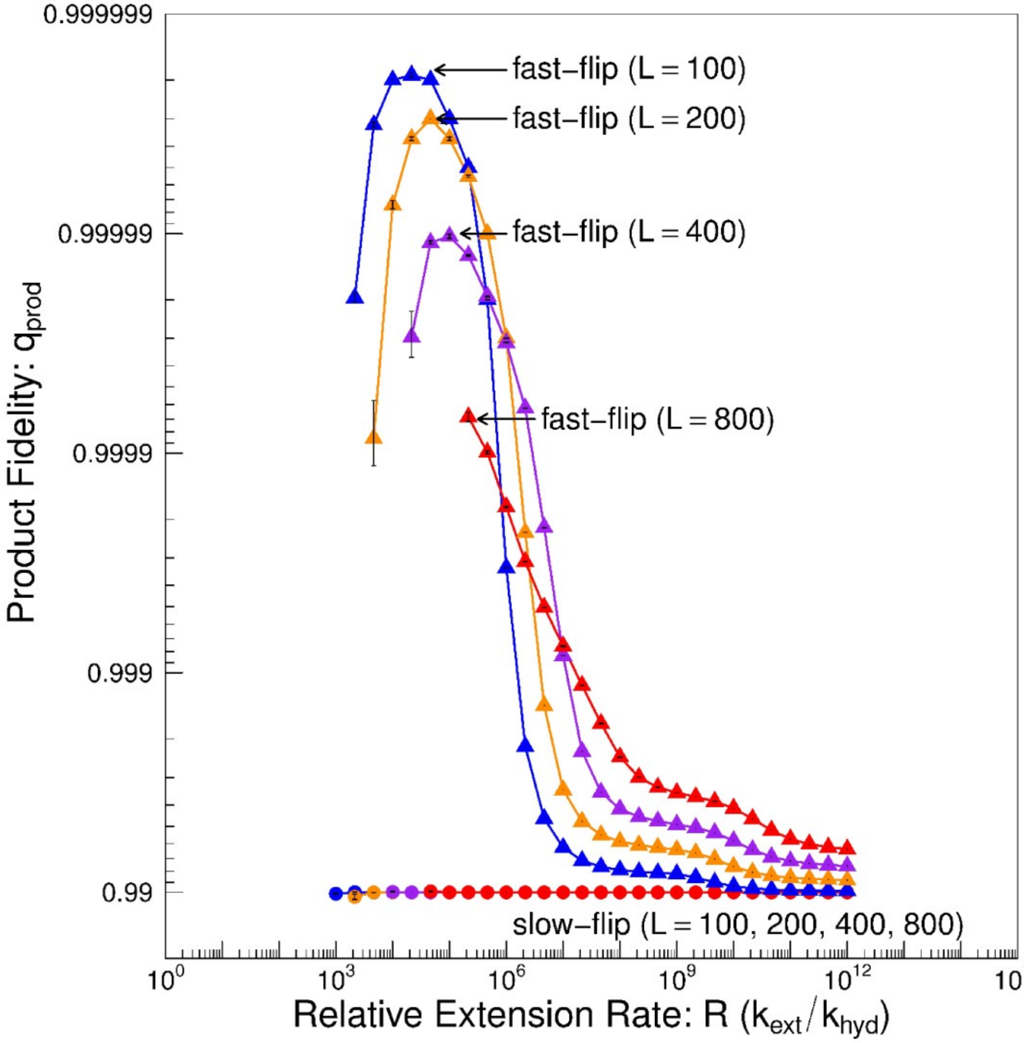


Figure S1: The per-base product fidelity, q_{prod}, is plotted on a log axis as a function of R, the ratio of extension rate to hydrolysis rate. In the slow-flipping limit, the per-base product fidelity is the same as the extension fidelity, $q_{prod} = q_{ext}$. In the fast-flipping limit, the per-base product fidelity in greater, $q_{prod} \gg q_{ext}$. When R is small, a maximum q_{prod} is observed which is dependent on the template length. Error bars represent the standard error in the mean. Parameters used: $\Delta\bar{G}^0_{stack} = -2\ kcal/mol$, $T = 37$°C, and $q_{ext} = 0.99$.

Chapter 7: Discussion

The aim of this thesis was to utilize computational modeling to further our understanding about the emergence of RNA replication in an RNA world. The results of chapter 2 have shifted the research direction of our research group, and hopefully others as well, from surface environments towards protocell like environments. Follow up research has further bolstered this shift by directly comparing surface and protocell environment (Shah et al. 2019). In chapter 3, we proposed that RNA could have emerged from a prebiotic soup due to its superior ability to act as a template for replication. Recent experimental research now supports this, and the authors of the study make a very similar claim (Kim et al. 2020). While chapter 4 was only recently published, we are hopefully that it will renew interest in Earth-based sources of nucleotides for an RNA world, as opposed to meteoritic sources of organics. We are also hopeful that chapters 5 and 6 will persuade researchers to consider the rolling-circle mechanism of replication, both under non-enzymatic and ribozyme-catalyzed conditions.

7.1 Remaining Problems in RNA replication

Despite being a focus of research for the better half of a century, many problems with RNA replication persist. In Szostak's 2012 review, eight outstanding problems with non-enzymatic RNA replication are discussed, and are copied below for easy reference (Szostak 2012).

1. Regiospecificity: chemical template copying generates complementary strands with a heterogeneous backbone containing randomly interspersed 2'-5' and 3'-5' linkages.
2. High Tm of long RNA duplexes: because of the high melting temperature of RNA duplexes, the accurate copying of an RNA template will generate a dead-end duplex product.

3. Fidelity of template copying chemistry: the accuracy of chemical replication is insufficient to allow for the propagation of functional genetic information.
4. Rate of template copying chemistry: chemical template copying occurs on the same timescale as template and substrate degradation.
5. Reactivation chemistry: the efficiency of template copying is limited by substrate hydrolysis, but current means of re-activating hydrolyzed substrates lead to damaging side reactions that would destroy both templates and substrates.
6. Divalent metal ions: the high concentrations of divalent cations required for RNA template-copying reactions catalyze RNA degradation and are incompatible with known vesicle replication systems.
7. Primer-independent RNA replication: replication based on primer-extension is incompatible with a protocell model system, because protocells cannot take up exogenous primers.
8. Strand reannealing: the reannealing of separated strands prevents template copying, but the rate of strand reannealing is orders of magnitude faster than current copying chemistry.

In chapter 5, we investigated and provided a theoretical foundation for problem #2, referred to here as the product inhibition problem. To do so, we considered the mechanisms of RNA replication found in the viral world, as well as the abiotic mechanisms proposed in the literature. Starting with viral mechanisms of replication seemed intuitive, given that we know viruses replicate exponentially and have overcome the product inhibition problem. Based on currently understood reaction kinetics, we found that rolling-circle replication may work non-enzymatically, resulting in exponential growth and solving the product-inhibition problem. Furthermore, since the rolling-circle template is dsRNA, the fast strand reannealing problem (#8) is avoided. And interestingly, the “problem” of fast strand reannealing allows rolling-circle replication to become primer-independent at protocell-like concentration, thus solving problem #7. In chapter 6, further investigations predict that non-enzymatic rolling-circle synthesis undergoes error-correction, solving the problem of poor fidelity of RNA synthesis (#3), and likely the regiospecificity problem (#1). Therefore, by simply changing the mechanism of replication, five of the eight problems of non-enzymatic replication appear to be solved.

Due to the nature of the remaining problems (#4, #5, #6), it seems unlikely that they will be solved by rolling-circle replication. These problems appear to stem from the relatively fast rate of RNA hydrolysis. Instead of trying to enhance the rate of replication, e.g. using highly activated nucleotides, it may be beneficial to consider ways in which RNA is stabilized. In modern life, RNA can be chemically modified through methylation of the 2'-hydroxyl group. This is known to be a universally conserved trait of life and it likely emerged pre-LUCA (Jeffares et al. 1998; Poole et al. 2000; Rana and Ankri 2016). For the emergence of life, this methylation may be critical as it greatly enhances the stability of RNA. With methylated-RNA, the problem of fast template hydrolysis may therefore be solved (#4), in addition to the problem of template susceptibility to divalent cations (#6). Furthermore, slower template hydrolysis may also allow the simpler NTP's to be used, thereby solving the activation chemistry problem (#5).

In this scenario, life would have emerged in a methylated-RNA world (Poole et al. 2000), likely undergoing rolling-circle replication. Once life achieved sufficient complexity such that DNA was discovered, there would be little evolutionary benefit to maintaining highly methylated RNA genomes. RNA methylation then decreased in prominence but remained in critical areas such as within the ribosome, where it remains in modern life (Jeffares et al. 1998). The strength to scenarios such as this one is that it adds minimal complexity to the origin of life. No new chemistry, nor mechanisms, have been invented. The emergence of methylation needs to be explained, this scenario simply assumes that methylation emerged earlier than previously thought, likely originating non-enzymatically.

7.2 The Search for Life

In addition to understanding the origin of life on Earth, the study of RNA replication may have profound implications for the search for life elsewhere in the universe. When attempting to determine the number of intelligent civilizations in the universe, or our own galaxy, the Drake equation predicts an abundance of life (Frank and Sullivan 2016; Engler and von Wehrden 2019). However, the apparent silence of these neighbors has led to the idea of a "Great Filter" (Hanson 1998; Haqq-Misra et al. 2020). The idea that there are incredibly difficult hurdles in the origin and evolution of complex life which reconcile this discrepancy. This filter, in essence, decreases the likelihood that a planet would allow for the emergence of complex life. The contents of this thesis contribute to a different, but equally important, type of filter for the origin of life, one which decreases the diversity of life as complexity increases.

If we consider these thesis chapters in their order for the emergence of life, as opposed to the order in which they are presented, we find that each chapter decreases the diversity of the resulting life. For instance, in Chapter 3, we considered a pool of nucleotide precursors which were polymerizing into random polymers, and then through replication, were selected for fast replication. The result of which was a decrease in complexity of the solution and the emergence of uniform RNA polymers. If true, this suggests that RNA was selected from the prebiotic soup due to its ability for short polymers to replicate fast. Would a similar prebiotic soup on a distant world result in similar selection for an RNA-like polymer?

Continuing our story, Chapters 5 and 6, considered the mechanisms of non-enzymatic replication. Once again, despite the numerous possibilities, only one mechanism appears to work non-enzymatically, that being rolling-circle replication. Since self-cleavage is a necessity for rolling-circle replication, any polymers which are unable to self-cleave would be unable to replicate, and thus a second filter is observed. In Chapter 4, we considered smaller nucleotide alphabets, in particular those which lack cytosine. Once again, we observe a filter. Without cytosine, the emergence of life appears highly improbable due to the limited complexity of folded structures. Lastly, in Chapter 2, we considered the replication of a hypothetical polymerase ribozyme and found that high processivity is vital for survival. As each new filter is applied, the diversity of polymers which could give rise to complex life decreases. The implication being that if we do find life elsewhere, it will likely look similar to us.

If the existence of diversity filters is real, then the search for life elsewhere becomes the search for viruses, viroids, and life as we know it. In some ways, this is good news since Earth life has been well documented and continues to be an active area of research. It is also convenient that viroids and viruses, the simplest replicators on Earth, appear to resemble the first life on Earth. Such entities may therefore have been present throughout most of Earth's history and may be abundant on distant worlds regardless of its stage of evolution. Furthermore, if it is true that life has passed through many filters, then the genetic material of these viruses and viroids should be remarkably similar to our RNA and DNA, and thus should be easily detectable.

However, such diversity filters also imply something far grimmer, that these other worldly viruses may be compatible with Earth life. In which case, exploration of distant worlds may bring the possibility of mutually assured destruction, and thus contribute to the "Great Filter". Does the curiosity of complex life, which provides the tools to search for extraterrestrial life, also lead to its own destruction?

References

Abramov O, Kring DA, Mojzsis SJ. 2013. The impact environment of the Hadean Earth. Geochemistry. 73(3):227–248. doi:10.1016/j.chemer.2013.08.004.

Achilles T, von Kiedrowski G. 1993. A Self-Replicating System from Three Starting Materials. Angew Chemie Int Ed English. 32(8):1198–1201. doi:10.1002/anie.199311981.

Aldersley MF, Joshi PC, Huang Y. 2017. The Comparison of Hydrochloric Acid and Phosphoric Acid Treatments in the Preparation of Montmorillonite Catalysts for RNA Synthesis. Orig Life Evol Biosph. 47(3):297–304. doi:10.1007/s11084-017-9533-6.

Attwater J, Raguram A, Morgunov AS, Gianni E, Holliger P. 2018. Ribozyme-catalysed RNA synthesis using triplet building blocks. Elife. 7. doi:10.7554/eLife.35255.

Attwater J, Wochner A, Holliger P. 2013. In-ice evolution of RNA polymerase ribozyme activity. Nat Chem. 5(12):1011–1018. doi:10.1038/nchem.1781.

Ban N, Nissen P, Hansen J, Moore PB, Steitz TA. 2000. The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. Science (80-). 289(5481):905–920. doi:10.1126/science.289.5481.905.

Barboni M, Boehnke P, Keller B, Kohl IE, Schoene B, Young ED, McKeegan KD. 2017. Early formation of the Moon 4.51 billion years ago. Sci Adv. 3(1):e1602365. doi:10.1126/sciadv.1602365.

Bartel DP, Szostak JW. 1993. Isolation of new ribozymes from a large pool of random sequences. Science (80-). 261(5127):1411–1418. doi:10.1126/science.7690155.

Becker S, Feldmann J, Wiedemann S, Okamura H, Schneider C, Iwan K, Crisp A, Rossa M, Amatov T, Carell T. 2019. Unified prebiotically plausible synthesis of pyrimidine and purine RNA ribonucleotides. Science (80-). 366(6461):76–82. doi:10.1126/science.aax2747.

Bell EA, Boehnke P, Harrison TM, Mao WL. 2015. Potentially biogenic carbon preserved in a 4.1 billion-year-old zircon. Proc Natl Acad Sci. 112(47):14518–14521. doi:10.1073/pnas.1517557112.

Benner SA, Bell EA, Biondi E, Brasser R, Carell T, Kim H, Mojzsis SJ, Omran A, Pasek MA, Trail D. 2020. When Did Life Likely Emerge on Earth in an RNA-First Process? ChemSystemsChem. 2(2). doi:10.1002/syst.201900035.

Benner SA, Ellington AD, Tauer A. 1989. Modern metabolism as a palimpsest of the RNA world. Proc Natl Acad Sci U S A. 86(18):7054–7058. doi:10.1073/pnas.86.18.7054.

Benner SA, Kim HJ, Biondi E. 2019. Prebiotic chemistry that could not not have happened. Life. 9(4):84. doi:10.3390/life9040084.

Berliner AJ, Mochizuki T, Stedman KM. 2018. Astrovirology: viruses at large in the universe. Astrobiology. 18(2):207–223. doi:10.1089/ast.2017.1649.

Bernhardt HS. 2012. The RNA world hypothesis: the worst theory of the early evolution of life (except for all the others)a. Biol Direct. 7. doi:10.1186/1745-6150-7-23.

Blackburn EH, Greider CW, Szostak JW. 2006. Telomeres and telomerase: The path from maize, Tetrahymena and yeast to human cancer and aging. Nat Med. 12(10):1133–1138. doi:10.1038/nm1006-1133.

Boehnke P, Harrison TM. 2016. Illusory Late Heavy Bombardments. Proc Natl Acad Sci. 113(39):10802–10806. doi:10.1073/pnas.1611535113.

Bolli M, Micura R, Eschenmoser A. 1997. Pyranosyl-RNA: chiroselective self-assembly of base sequences by ligative oligomerization of tetra nucleotide-2′,3′-cyclophosphates (with a commentary concerning the origin of biomolecular homochirality). Chem Biol. 4(4):309–320. doi:10.1016/S1074-5521(97)90074-0.

Busigny V, Planavsky NJ, Jézéquel D, Crowe S, Louvat P, Moureau J, Viollier E, Lyons TW. 2014. Iron isotopes in an Archean ocean analogue. Geochim Cosmochim Acta. 133:443–462. doi:10.1016/j.gca.2014.03.004.

Butlerow A. 1861. Formation synthétique d'une substance sucrée. CR Acad Sci. 53:145–147.

Callahan MP, Smith KE, Cleaves HJ, Ruzicka J, Stern JC, Glavin DP, House CH, Dworkin JP. 2011. Carbonaceous meteorites contain a wide range of extraterrestrial nucleobases. Proc Natl Acad Sci U S A. 108(34):13995–13998. doi:10.1073/pnas.1106493108.

Cameron CE, Raney KD, Götte M. 2009. Viral genome replication. Viral Genome Replication.:1–636. doi:10.1007/b135974.

Cavosie AJ, Valley JW, Wilde SA, E.I.M.F. 2005. Magmatic δ18O in 4400–3900 Ma detrital zircons: A record of the alteration and recycling of crust in the Early Archean. Earth Planet Sci Lett. 235(3–4):663–681. doi:10.1016/j.epsl.2005.04.028.

Cech TR. 1987. The chemistry of self-splicing RNA and RNA enzymes. Science (80-). 236(4808):1532–1539. doi:10.1126/science.2438771.

Cech TR. 2000. The ribosome is a ribozyme. Science (80-). 289(5481):878–879. doi:10.1126/science.289.5481.878.

Cheng LKL, Unrau PJ. 2010. Closing the circle: replicating RNA with RNA. Cold Spring Harb Perspect Biol. 2(10). doi:10.1101/cshperspect.a002204.

Cleaves HJ, Bada JL. 2012. The Prebiotic Chemistry of Alternative Nucleic Acids. :3–33. doi:10.1007/978-94-007-2941-4_1.

Costanzo G, Pino S, Timperio AM, Šponer JE, Šponer J, Nováková O, Šedo O, Zdráhal Z, Di Mauro E. 2016. Non-enzymatic oligomerization of 3′,5′ cyclic AMP. PLoS One. 11(11). doi:10.1371/journal.pone.0165723.

Costanzo G, Saladino R, Botta G, Giorgi A, Scipioni A, Pino S, Di Mauro E. 2012. Generation of RNA Molecules by a Base-Catalysed Click-Like Reaction. ChemBioChem. 13(7):999–1008. doi:10.1002/cbic.201200068.

Crick F. 1970. Central dogma of molecular biology. Nature. 227(5258):561–563. doi:10.1038/227561a0.

Dickson KS, Burns CM, Richardson JP. 2000. Determination of the free-energy change for repair of a DNA phosphodiester bond. J Biol Chem. 275(21):15828–15831. doi:10.1074/jbc.M910044199.

Diener TO. 1989. Circular RNAs: Relics of precellular evolution? Proc Natl Acad Sci U S A. 86(23):9370–9374. doi:10.1073/pnas.86.23.9370.

Diener TO. 2016. Viroids: “living fossils” of primordial RNAs? Biol Direct. 11(1). doi:10.1186/s13062-016-0116-7.

Doolittle WF. 2000. The nature of the universal ancestor and the evolution of the proteome. Curr Opin Struct Biol. 10(3):355–358. doi:10.1016/S0959-440X(00)00096-8.

Doudna JA, Cech TR. 2002. The chemical repertoire of natural ribozymes. Nature. 418(6894):222–228. doi:10.1038/418222a.

Draper WE, Hayden EJ, Lehman N. 2008. Mechanisms of covalent self-assembly of the Azoarcus ribozyme from four fragment oligonucleotides. Nucleic Acids Res. 36(2):520–531. doi:10.1093/nar/gkm1055.

Edeleva E, Salditt A, Stamp J, Schwintek P, Boekhoven J, Braun D. 2019. Continuous nonenzymatic cross-replication of DNA strands with in situ activated DNA oligonucleotides. Chem Sci. 10(22):5807–5814. doi:10.1039/c9sc00770a.

Ekland EH, Bartel DP. 1995. The secondary structure and sequence optimization of an RNA ligase ribozyme. Nucleic Acids Res. 23(16):3231–3238. doi:10.1093/nar/23.16.3231.

Ekland EH, Szostak JW, Bartel DP. 1995. Structurally complex and highly active RNA ligases derived from random RNA sequences. Science (80-). 269(5222):364–370. doi:10.1126/science.7618102.

Engler J-O, von Wehrden H. 2019. 'Where is everybody?' An empirical appraisal of occurrence, prevalence and sustainability of technological species in the Universe. Int J Astrobiol. 18(6):495–501. doi:10.1017/S1473550418000496.

Ferris JP. 2002. Montmorillonite catalysis of 30-50 mer oligonucleotides: Laboratory demonstration of potential steps in the origin of the RNA world. Orig Life Evol Biosph. 32(4):311–332. doi:10.1023/A:1020543312109.

Ferus M, Pietrucci F, Saitta AM, Knížek A, Kubelík P, Ivanek O, Shestivska V, Civiš S. 2017. Formation of nucleobases in a Miller-Urey reducing atmosphere. Proc Natl Acad Sci U S A. 114(17):4306–4311. doi:10.1073/pnas.1700010114.

Flores R, Gago-Zachert S, Serra P, Sanjuán R, Elena SF. 2014. Viroids: Survivors from the RNA World? Annu Rev Microbiol. 68(1):395–414. doi:10.1146/annurev-micro-091313-103416.

Flores R, Grubb D, Elleuch A, Nohales MÁ, Delgado S, Gago S. 2011. Rolling-circle replication of viroids, viroid-like satellite RNAs and hepatitis delta virus: Variations on a theme. RNA Biol. 8(2):200–206. doi:10.4161/rna.8.2.14238.

Frank A, Sullivan WT. 2016. A New Empirical Constraint on the Prevalence of Technological Species in the Universe. Astrobiology. 16(5):359–362. doi:10.1089/ast.2015.1418.

Frick DN, Richardson CC. 2001. DNA Primases. Annu Rev Biochem. 70(1):39–80. doi:10.1146/annurev.biochem.70.1.39.

Fuller WD, Sanchez RA, Orgel LE. 1972. Studies in prebiotic synthesis. VII - Solid-State Synthesis of Purine Nucleosides. J Mol Evol. 1(3):249–257. doi:10.1007/BF01660244.

Furukawa Y, Chikaraishi Y, Ohkouchi N, Ogawa NO, Glavin DP, Dworkin JP, Abe C, Nakamura T. 2019. Extraterrestrial ribose and other sugars in primitive meteorites. Proc Natl Acad Sci U S A. 116(49):24440–24445. doi:10.1073/pnas.1907169116.

Gavette J V., Stoop M, Hud N V., Krishnamurthy R. 2016. RNA–DNA Chimeras in the Context of an RNA World Transition to an RNA/DNA World. Angew Chemie - Int Ed. 55(42):13204–13209. doi:10.1002/anie.201607919.

Gilbert W. 1986. Origin of life: The RNA world. Nature. 319(6055):618. doi:10.1038/319618a0.

Giovanna C, Pino S, Ciciriello F, Di Mauro E. 2009. Generation of long RNA chains in water. J Biol Chem. 284(48):33206–33216. doi:10.1074/jbc.M109.041905.

Di Giulio M. 2011. The Last Universal Common Ancestor (LUCA) and the Ancestors of Archaea and Bacteria were Progenotes. J Mol Evol. 72(1):119–126. doi:10.1007/s00239-010-9407-2.

Greider CW, Blackburn EH. 1985. Identification of a specific telomere terminal transferase activity in tetrahymena extracts. Cell. 43(2 PART 1):405–413. doi:10.1016/0092-8674(85)90170-9.

Grew ES, Bada JL, Hazen RM. 2011. Borate Minerals and Origin of the RNA World. Orig Life Evol Biosph. 41(4):307–316. doi:10.1007/s11084-010-9233-y.

Hammann C, Luptak A, Perreault J, De La Peña M. 2012. The ubiquitous hammerhead ribozyme. Rna. 18(5):871–885. doi:10.1261/rna.031401.111.

Hanson R. 1998. The Great Filter—are we almost past it? Robin Hanson's website.

Haqq-Misra J, Kopparapu RK, Schwieterman E. 2020. Observational Constraints on the Great Filter. Astrobiology. 20(5):572–579. doi:10.1089/ast.2019.2154.

Harrison TM. 2005. Heterogeneous Hadean Hafnium: Evidence of Continental Crust at 4.4 to 4.5 Ga. Science (80-). 310(5756):1947–1950. doi:10.1126/science.1117926.

Hayden EJ, Von Kiedrowski G, Lehman N. 2008. Systems chemistry on ribozyme self-construction: Evidence for anabolic autocatalysis in a recombination network. Angew Chemie - Int Ed. 47(44):8424–8428. doi:10.1002/anie.200802177.

He C, Gállego I, Laughlin B, Grover MA, Hud N V. 2017. A viscous solvent enables information transfer from gene-length nucleic acids in a model prebiotic replication cycle. Nat Chem. 9(4):318–324. doi:10.1038/nchem.2628.

Higgs PG. 2016. The effect of limited diffusion and wet–dry cycling on reversible polymerization reactions: Implications for prebiotic synthesis of nucleic acids. Life. 6(2). doi:10.3390/life6020024.

Higgs PG. 2019. Three Ways to Make an RNA Sequence. Handb Astrobiol.:395–407. doi:10.1201/b22230-28.

Higgs PG, Lehman N. 2015. The RNA World: Molecular cooperation at the origins of life. Nat Rev Genet. 16(1):7–17. doi:10.1038/nrg3841.

Horning DP, Joyce GF. 2016. Amplification of RNA by an RNA polymerase ribozyme. Proc Natl Acad Sci U S A. 113(35):9786–9791. doi:10.1073/pnas.1610103113.

Huang W, Ferris JP. 2006. One-step, regioselective synthesis of up to 50-mers of RNA oligomers by montmorillonite catalysis. J Am Chem Soc. 128(27):8914–8919. doi:10.1021/ja061782k.

Hug LA, Baker BJ, Anantharaman K, Brown CT, Probst AJ, Castelle CJ, Butterfield CN, Hernsdorf AW, Amano Y, Ise K, et al. 2016. A new view of the tree of life. Nat Microbiol. 1(5). doi:10.1038/nmicrobiol.2016.48.

Hulo C, De Castro E, Masson P, Bougueleret L, Bairoch A, Xenarios I, Le Mercier P. 2011. ViralZone: A knowledge resource to understand virus diversity. Nucleic Acids Res. 39(SUPPL. 1). doi:10.1093/nar/gkq901.

Jacobson SA, Morbidelli A, Raymond SN, O'Brien DP, Walsh KJ, Rubie DC. 2014. Highly siderophile elements in Earth's mantle as a clock for the Moon-forming impact. Nature. 508(7494):84–87. doi:10.1038/nature13172.

Jayathilaka TS, Lehman N. 2018. Spontaneous Covalent Self-Assembly of the Azoarcus Ribozyme from Five Fragments. ChemBioChem. 19(3):217–220. doi:10.1002/cbic.201700591.

Jeffares DC, Poole AM, Penny D. 1998. Relics from the RNA world. J Mol Evol. 46(1):18–36. doi:10.1007/PL00006280.

Jheeta S, Joshi PC. 2014. Prebiotic RNA synthesis by montmorillonite catalysis. Life. 4(3):318–330. doi:10.3390/life4030318.

Johnston WK, Unrau PJ, Lawrence MS, Glasner ME, Bartel DP. 2001. RNA-catalyzed RNA polymerization: Accurate and general RNA-templated primer extension. Science (80-). 292(5520):1319–1325. doi:10.1126/science.1060786.

Joyce GF. 2002. The antiquity of RNA-based evolution. Nature. 418(6894):214–221. doi:10.1038/418214a.

Joyce GF, Visser GM, van Boeckel CAA, van Boom JH, Orgel LE, van Westrenen J. 1984. Chiral selection in poly(C)-directed synthesis of oligo(G). Nature. 310(5978):602–604. doi:10.1038/310602a0.

Keller MA, Kampjut D, Harrison SA, Ralser M. 2017. Sulfate radicals enable a non-enzymatic Krebs cycle precursor. Nat Ecol Evol. 1(4). doi:10.1038/s41559-017-0083.

Keller MA, Turchyn A V., Ralser M. 2014. Non-enzymatic glycolysis and pentose phosphate pathway-like reactions in a plausible Archean ocean. Mol Syst Biol. 10(4). doi:10.1002/msb.20145228.

Keller MA, Zylstra A, Castro C, Turchyn A V., Griffin JL, Ralser M. 2016. Conditional iron and pH-dependent activity of a non-enzymatic glycolysis and pentose phosphate pathway. Sci Adv. 2(1). doi:10.1126/sciadv.1501235.

von Kiedrowski G. 1986. A Self-Replicating Hexadeoxynucleotide. Angew Chemie Int Ed English. 25(10):932–935. doi:10.1002/anie.198609322.

Kim DE, Joyce GF. 2004. Cross-catalytic replication of an RNA ligase ribozyme. Chem Biol. 11(11):1505–1512. doi:10.1016/j.chembiol.2004.08.021.

Kim SC, Zhou L, Zhang W, O'Flaherty DK, Rondo-Brovetto V, Szostak JW. 2020. A Model for the Emergence of RNA from a Prebiotically Plausible Mixture of Ribonucleotides, Arabinonucleotides, and 2′-Deoxynucleotides. J Am Chem Soc. 142(5):2317–2326. doi:10.1021/jacs.9b11239.

Kim YE, Higgs PG. 2016. Co-operation between Polymerases and Nucleotide Synthetases in the RNA World. Wilke CO, editor. PLOS Comput Biol. 12(11):e1005161. doi:10.1371/journal.pcbi.1005161.

Koonin E V. 2003. Comparative genomics, minimal gene-sets and the last universal common ancestor. Nat Rev Microbiol. 1(2):127–136. doi:10.1038/nrmicro751.

Kozlov IA, Orgel LE. 2000. Nonenzymatic template-directed synthesis of RNA from monomers. Mol Biol. 34(6):781–789. doi:10.1023/A:1026663422976.

Krishnamurthy R. 2015. On the Emergence of RNA. Isr J Chem. 55(8):837–850. doi:10.1002/ijch.201400180.

Lai LB, Vioque A, Kirsebom LA, Gopalan V. 2010. Unexpected diversity of RNase P, an ancient tRNA processing enzyme: Challenges and prospects. FEBS Lett. 584(2):287–296. doi:10.1016/j.febslet.2009.11.048.

Laurino P, Tawfik DS. 2017. Spontaneous Emergence of S-Adenosylmethionine and the Evolution of Methylation. Angew Chemie - Int Ed. 56(1):343–345. doi:10.1002/anie.201609615.

Lawrence MS, Bartel DP. 2003. Processivity of ribozyme-catalyzed RNA polymerization. Biochemistry. 42(29):8748–8755. doi:10.1021/bi034228l.

Lawrence MS, Bartel DP. 2005. New ligase-derived RNA polymerase ribozymes. Rna. 11(8):1173–1180. doi:10.1261/rna.2110905.

Lebrun T, Massol H, Chassefière E, Davaille A, Marcq E, Sarda P, Leblanc F, Brandeis G. 2013. Thermal evolution of an early magma ocean in interaction with the atmosphere. J Geophys Res Planets. 118(6):1155–1176. doi:10.1002/jgre.20068.

Levy M, Miller SL. 1998. The stability of the RNA bases: Implications for the origin of life. Proc Natl Acad Sci U S A. 95(14):7933–7938. doi:10.1073/pnas.95.14.7933.

Lincoln TA, Joyce GF. 2009. Self-sustained replication of an RNA enzyme. Science (80-). 323(5918):1229–1232. doi:10.1126/science.1167856.

Lorenz R, Bernhart SH, Höner zu Siederdissen C, Tafer H, Flamm C, Stadler PF, Hofacker IL. 2011. ViennaRNA Package 2.0. Algorithms Mol Biol. 6(1). doi:10.1186/1748-7188-6-26.

Lundin D, Berggren G, Logan DT, Sjöberg BM. 2015. The origin and evolution of ribonucleotide reduction. Life. 5(1):604–636. doi:10.3390/life5010604.

Maurel MC, Leclerc F, Vergne J, Zaccai G. 2019. RNA back and Forth: Looking through Ribozyme and Viroid Motifs. Viruses. 11(3). doi:10.3390/v11030283.

Messner CB, Driscoll PC, Piedrafita G, De Volder MFL, Ralser M. 2017. Nonenzymatic gluconeogenesis-like formation of fructose 1,6-bisphosphate in ice. Proc Natl Acad Sci U S A. 114(28):7403–7407. doi:10.1073/pnas.1702274114.

Miller SL. 1953. A production of amino acids under possible primitive earth conditions. Science (80-). 117(3046):528–529. doi:10.1126/science.117.3046.528.

Monteux J, Andrault D, Samuel H. 2016. On the cooling of a deep terrestrial magma ocean. Earth Planet Sci Lett. 448:140–149. doi:10.1016/j.epsl.2016.05.010.

Moore PB, Steitz TA. 2002. The involvement of RNA in ribosome function. Nature. 418(6894):229–235. doi:10.1038/418229a.

Muchowska KB, Varma SJ, Chevallot-Beroux E, Lethuillier-Karl L, Li G, Moran J. 2017. Metals promote sequences of the reverse Krebs cycle. Nat Ecol Evol. 1(11):1716–1721. doi:10.1038/s41559-017-0311-7.

Nelson JW, Breaker RR. 2017. The lost language of the RNA World. Sci Signal. 10(483). doi:10.1126/scisignal.aam8812.

Nemchin AA, Whitehouse MJ, Menneken M, Geisler T, Pidgeon RT, Wilde SA. 2008. A light carbon reservoir recorded in zircon-hosted diamond from the Jack Hills. Nature. 454(7200):92–95. doi:10.1038/nature07102.

Nissen P, Hansen J, Ban N, Moore PB, Steitz TA. 2000. The structural basis of ribosome activity in peptide bond synthesis. Science (80-). 289(5481):920–930. doi:10.1126/science.289.5481.920.

Nutman AP, Bennett VC, Friend CRL, Van Kranendonk MJ, Chivas AR. 2016. Rapid emergence of life shown by discovery of 3,700-million-year-old microbial structures. Nature. 537(7621):535–538. doi:10.1038/nature19355.

O'Donnell M, Langston L, Stillman B. 2013. Principles and concepts of DNA replication in bacteria, archaea, and eukarya. Cold Spring Harb Perspect Biol. 5(7). doi:10.1101/cshperspect.a010108.

Ohtomo Y, Kakegawa T, Ishida A, Nagase T, Rosing MT. 2014. Evidence for biogenic graphite in early Archaean Isua metasedimentary rocks. Nat Geosci. 7(1):25–28. doi:10.1038/ngeo2025.

Orgel LE. 1995. Unnatural Selection in Chemical Systems. Acc Chem Res. 28(3):109–118. doi:10.1021/ar00051a004.

Orgel LE. 2004. Prebiotic chemistry and the origin of the RNA world. Crit Rev Biochem Mol Biol. 39(2):99–123. doi:10.1080/10409230490460765.

Oró J. 1960. Synthesis of adenine from ammonium cyanide. Biochem Biophys Res Commun. 2(6):407–412. doi:10.1016/0006-291X(60)90138-8.

Ouzounis CA, Kunin V, Darzentas N, Goldovsky L. 2006. A minimal estimate for the gene content of the last universal common ancestor - Exobiology from a terrestrial perspective. Res Microbiol. 157(1):57–68. doi:10.1016/j.resmic.2005.06.015.

Paul N, Joyce GF. 2002. A self-replicating ligase ribozyme. Proc Natl Acad Sci U S A. 99(20):12733–12740. doi:10.1073/pnas.202471099.

Pearce BKD, Pudritz RE. 2015. Seeding the Pregenetic Earth: Meteoritic Abundances of Nucleobases and Potential Reaction Pathways. Astrophys J. 807(1). doi:10.1088/0004-637X/807/1/85.

Pearce BKD, Tupper AS, Pudritz RE, Higgs PG. 2018. Constraining the Time Interval for the Origin of Life on Earth. Astrobiology. 18(3):343–364. doi:10.1089/ast.2017.1674.

Perreault J, Weinberg Z, Roth A, Popescu O, Chartrand P, Ferbeyre G, Breaker RR. 2011. Identification of Hammerhead Ribozymes in All Domains of Life Reveals Novel Structural Variations. PLoS Comput Biol. 7(5). doi:10.1371/journal.pcbi.1002031.

Poole AM, Logan DT. 2005. Modern mRNA Proofreading and Repair: Clues that the Last Universal Common Ancestor Possessed an RNA Genome? Mol Biol Evol. 22(6):1444–1455. doi:10.1093/molbev/msi132.

Poole AM, Logan DT, Sjöberg BM. 2002. The evolution of the ribonucleotide reductases: Much ado about oxygen. J Mol Evol. 55(2):180–196. doi:10.1007/s00239-002-2315-3.

Poole AM, Penny D, Sjöberg BM. 2000. Methyl-RNA: An evolutionary bridge between RNA and DNA? Chem Biol. doi:10.1016/S1074-5521(00)00042-9.

Powner MW, Gerland B, Sutherland JD. 2009. Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. Nature. 459(7244):239–242. doi:10.1038/nature08013.

Prywes N, Blain JC, Del Frate F, Szostak JW. 2016. Nonenzymatic copying of RNA templates containing all four letters is catalyzed by activated oligonucleotides. Elife. 5(JUN2016). doi:10.7554/eLife.17756.

Raffaelli N. 2011. Nicotinamide Coenzyme Synthesis: A Case of Ribonucleotide Emergence or a Byproduct of the RNA World? Orig Life Primal Self-Organization.:185–208. doi:10.1007/978-3-642-21625-1_9.

Rajamani S, Vlassov A V., Benner S, Coombs A, Olasagasti F, Deamer DW. 2008. Lipid-assisted synthesis of RNA-like polymers from mononucleotides. Orig Life Evol Biosph. 38(1):57–74. doi:10.1007/s11084-007-9113-2.

Ralser M. 2014. The RNA world and the origin of metabolic enzymes. Biochem Soc Trans. 42(4):985–988. doi:10.1042/BST20140132.

Ralser M. 2018. An appeal to magic? The discovery of a non-enzymatic metabolism and its role in the origins of life. Biochem J. 475(16):2577–2592. doi:10.1042/BCJ20160866.

Rana AK, Ankri S. 2016. Reviving the RNA world: An insight into the appearance of RNA methyltransferases. Front Genet. 7(JUN). doi:10.3389/fgene.2016.00099.

Raymann K, Brochier-Armanet C, Gribaldo S. 2015. The two-domain tree of life is linked to a new root for the Archaea. Proc Natl Acad Sci U S A. 112(21):6670–6675. doi:10.1073/pnas.1420858112.

Ricardo A, Carrigan MA, Olcott AN, Benner SA. 2004. Borate Minerals Stabilize Ribose. Science (80-). 303(5655):196. doi:10.1126/science.1092464.

Robertson MP, Joyce GF. 2012. The origins of the RNA World. Cold Spring Harb Perspect Biol. 4(5):1. doi:10.1101/cshperspect.a003608.

Rosing MT. 1999. 13C-depleted carbon microparticles in ≥3700-Ma sea-floor sedimentary rocks from west Greenland. Science (80-). 283(5402):674–676. doi:10.1126/science.283.5402.674.

Rouxel OJ, Bekker A, Edwards KJ. 2005. Iron isotope constraints on the Archean and Paleoproterozoic ocean redox state. Science (80-). 307(5712):1088–1091. doi:10.1126/science.1105692.

Rudolph SA, Johnson EM, Greengard P. 1971. The enthalpy of hydrolysis of various 3’,5’-and 2’,3’-cyclic nucleotides. J Biol Chem. 246(5):1271–1273.

Shah V, de Bouter J, Pauli Q, Tupper AS, Higgs PG. 2019. Survival of RNA replicators is much easier in protocells than in surface-based, spatial systems. Life. doi:10.3390/life9030065.

Shay JA, Huynh C, Higgs PG. 2015. The origin and spread of a cooperative replicase in a prebiotic chemical system. J Theor Biol. doi:10.1016/j.jtbi.2014.09.019.

Sievers D, Von Kiedrowski G. 1994. Self-replication of complementary nucleotide-based oligomers. Nature. 369(6477):221–224. doi:10.1038/369221a0.

Da Silva L, Maurel MC, Deamer DW. 2015. Salt-Promoted Synthesis of RNA-like Molecules in Simulated Hydrothermal Conditions. J Mol Evol. 80(2):86–97. doi:10.1007/s00239-014-9661-9.

Spaeth A, Hargrave M. 2020. A polyaddition model for the prebiotic polymerization of RNA and RNA-like polymers. Life. 10(2):12. doi:10.3390/life10020012.

Spudis PD, Wilhelms DE, Robinson MS. 2011. The Sculptured Hills of the Taurus Highlands: Implications for the relative age of Serenitatis, basin chronologies and the cratering history of the Moon. J Geophys Res. 116(12):E00H03. doi:10.1029/2011JE003903.

Szilágyi A, Zachar I, Scheuring I, Kun Á, Könnyű B, Czárán T. 2017. Ecology and Evolution in the RNA World Dynamics and Stability of Prebiotic Replicator Systems. Life. 7(4):48. doi:10.3390/life7040048.

Szostak JW. 2012. The eightfold path to non-enzymatic RNA replication. J Syst Chem. 3(1). doi:10.1186/1759-2208-3-2.

Takeuchi N, Hogeweg P. 2009. Multilevel Selection in Models of Prebiotic Evolution II: A Direct Comparison of Compartmentalization and Spatial Self-Organization. Stormo GD, editor. PLoS Comput Biol. 5(10):e1000542. doi:10.1371/journal.pcbi.1000542.

Talini G, Gallori E, Maurel MC. 2009. Natural and unnatural ribozymes: Back to the primordial RNA world. Res Microbiol. 160(7):457–465. doi:10.1016/j.resmic.2009.05.005.

Valadkhan S. 2007. The spliceosome: A ribozyme at heart? Biol Chem. 388(7):693–697. doi:10.1515/BC.2007.080.

Wacey D, Kilburn MR, Saunders M, Cliff J, Brasier MD. 2011. Microfossils of sulphur-metabolizing cells in 3.4-billion-year-old rocks of Western Australia. Nat Geosci. 4(10):698–702. doi:10.1038/ngeo1238.

Watson J. 1972. Origin of Concatemeric T7DNA. Nature. 239(94):197–201. doi:10.1038/10.1038/newbio239197a0.

Weiss MC, Sousa FL, Mrnjavac N, Neukirchen S, Roettger M, Nelson-Sathi S, Martin WF. 2016. The physiology and habitat of the last universal common ancestor. Nat Microbiol. 1(9). doi:10.1038/nmicrobiol.2016.116.

White HB. 1976. Coenzymes as fossils of an earlier metabolic state. J Mol Evol. 7(2):101–104. doi:10.1007/BF01732468.

Williams TA, Foster PG, Cox CJ, Embley TM. 2013. An archaeal origin of eukaryotes supports only two primary domains of life. Nature. 504(7479):231–236. doi:10.1038/nature12779.

Wochner A, Attwater J, Coulson A, Holliger P. 2011. Ribozyme-catalyzed transcription of an active ribozyme. Science (80-). 332(6026):209–212. doi:10.1126/science.1200752.

Woese CR. 1987. Bacterial evolution. Microbiol Rev. 51(2):221.

Woese CR, Kandler O, Wheelis ML. 1990. Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria, and Eucarya. Proc Natl Acad Sci U S A. 87(12):4576–4579. doi:10.1073/pnas.87.12.4576.

Zaher HS, Unrau PJ. 2007. Selection of an improved RNA polymerase ribozyme with superior extension and fidelity. Rna. 13(7):1017–1026. doi:10.1261/rna.548807.

Zahnle K, Arndt N, Cockell C, Halliday A, Nisbet E, Selsis F, Sleep NH. 2007. Emergence of a Habitable Planet. Space Sci Rev. 129(1–3):35–78. doi:10.1007/s11214-007-9225-z.

Zhang X V., Martin ST. 2006. Driving parts of Krebs cycle in reverse through mineral photochemistry. J Am Chem Soc. 128(50):16032–16033. doi:10.1021/ja066103k.

Zielinski WS, Orgel LE. 1987. Autocatalytic synthesis of a tetranucleotide analogue. Nature. 327(6120):346–347. doi:10.1038/327346a0.