Self-replicating RNA molecules

A view on the RNA world hypothesis

Jerre G. Edens
Essay, February 2017

Supervisor: Prof. Dr. Sijbren Otto
Stratingh Institute of Chemistry

Contents

Abstract

The RNA world hypothesis suggests that RNA molecules were generated from prebiotic components enabling the generation of early life forms. The catalytic and informational storage capacity of RNA may have played a role in the origin of a self-sustained Darwinian evolutionary system. Bottom-up research has resulted in many findings suggesting that RNA molecules could find its origin in small building blocks present in the prebiotic pool. On the other hand top-down research has led to a promising RNA enzyme (ribozyme) system capable of exponential self-replication enabling Darwinian evolution. This ribozyme system may help closing the knowledge gap in the generation of the biological spectrum from the prebiotic pool.

Introduction

The RNA world hypothesis suggests that RNA molecules were generated from prebiotic components enabling the generation of early life forms. This widely accepted hypothesis is supported by the catalytic features and informational storage capacity of RNA. Also RNA has many critical functions in modern life forms suggesting that modern life may be an evolutionary remnant of a RNA world.

Catalytically active RNA molecules have been found in nature and have interesting mechanisms. Chapter I describes how some of these RNA catalysts operate. Also experiments to improve their catalytic activity or to find new artificial RNA catalysts are described.

Estimating of the content of the prebiotic pool and considering the extreme conditions, there is strong evidence that precursors for ribonucleotides were present in the prebiotic world (Sutherland, 2016). Chapter II will elaborate on the development of RNA from these precursors. Conceivable pathways towards nucleobases, the backbone and eventually protocells from mere prebiotic components are demonstrated. Because RNA has catalytic properties, chapter III elaborates on how this feature of RNA can be enhanced to archive specific binding of ligands, self-polymerization and eventually self-replication.

The RNA world hypothesis still entails many problems. The rare catalytic activity, the instability and the complexity of RNA keeps this hypothesis very questionable (Bernhardt, 2012). The results approaching the self-replication of RNA will be discussed in Chapter IV to give an answer to the main questions: What are the main problems of RNA synthesis from the prebiotic world and how far are we with complete Darwinian RNA self-replication?

Chapter I: Ribozymes

Together with DNA and proteins, RNA molecules are generally known for their essential role in life. In the general process of genetic information expression, DNA is transcribed into RNA which codes for proteins. In this model the RNA molecules are capable of storing information. In 1980, Cech was first to discover RNA molecules which not only function as genetic material but also as a biological catalyst (Zaug & Cech, 1980). He found evidence for an excision of a part of ribosomal RNA in the thermophilic bacteria *Tetrahymena*, which remained stable *in vitro* and showed catalytic activity. In 1982 was reported that this catalytic RNA molecule was actually self-splicing. This revolutionary discovery of RNA molecules capable of enzymatic activity, led to the development of the term "Ribozymes" (Kruger, et al., 1982). Cech and Altman were therefore rewarded with the 1989 Nobel Prize in chemistry. This finding supported the hypothesis of a RNA world which will be discussed later in this chapter.

Biological functions of ribozymes

After the discovery of the first ribozyme, more and more ribozymes with multiple functions have been reported. The most common catalytic activity of ribozymes entails a phosphate-group transfer and a peptide-bond formation of RNA and DNA (Fedor & Williamson, 2005). A ubiquitous ribozyme which cleaves tRNA precursor molecules is ribonuclease P (RNase P) found in *E. Coli* and *B. Subtillus* (Guerrier-Takada, Gardiner, Marsh, Pace, & Altman, 1983). This ribozyme cleaves RNA in a protein-like way. The bacterial ribozyme consist of a RNA strand and a polypeptide chain. *In vivo* these two parts are essential for activity but *in vitro* the RNA shows catalytic features by itself. The complex structure of this ribozyme could not give answers to which motifs were essential for RNA catalysis.

Later a smaller ribozyme called the hammerhead ribozyme was discovered in multiple satellite virus RNAs. The hammerhead motif consists of a conserved domain which uses the general reaction mechanisms for self-cleaving. This mechanism entails an S_N2-type reaction to enable phosphodiester-cleavage (Figure 1).

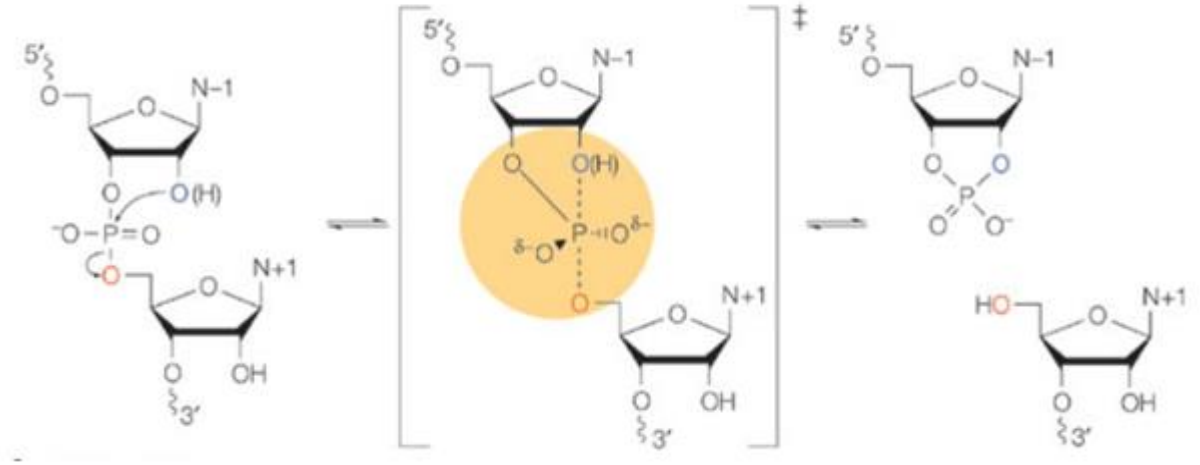

Figure 1: Self-cleaving mechanism of ribozymes. From (Fedor & Williamson, 2005)

The reversible reaction 1a shows a phosphodiester-cleavage in which the 2’ hydroxyl is attacking the phosphorous making the 5’ oxygen the leaving group.

The hammerhead ribozyme is only able to cleave itself in the presence of a divalent metal ion (Scott, Murray, Arnold, Stoddard, & Klug, 1996) or a high concentration of monovalent cations (Murray, Seyhan, Walter, Burke, & Scott, 1998). In natural state this ribozyme does not act as a catalyst since it is consumed by the reaction, however this ribozyme can be engineered into two RNA strands. Supplying the strand that gets cleaved in excess allows the ribozyme to catalyze with multiple turnover.

Another well-known ribozyme is the hepatitis delta virus (HPV) ribozyme found in the satellite virus of hepatitis B virus (HBV) (Ferré-D'Amaré, Zhou, & Doudna, 1998). This ribozyme also utilizes the S_N2-type reaction stabilized by divalent cations, however for this reaction also stabilization by cytidine is required. This led to the proposal of the general acid-base catalysis mechanism for this ribozyme shown in Figure 2 (Ke, Zhou, Ding, J.H.D., & Doudna, 2003).

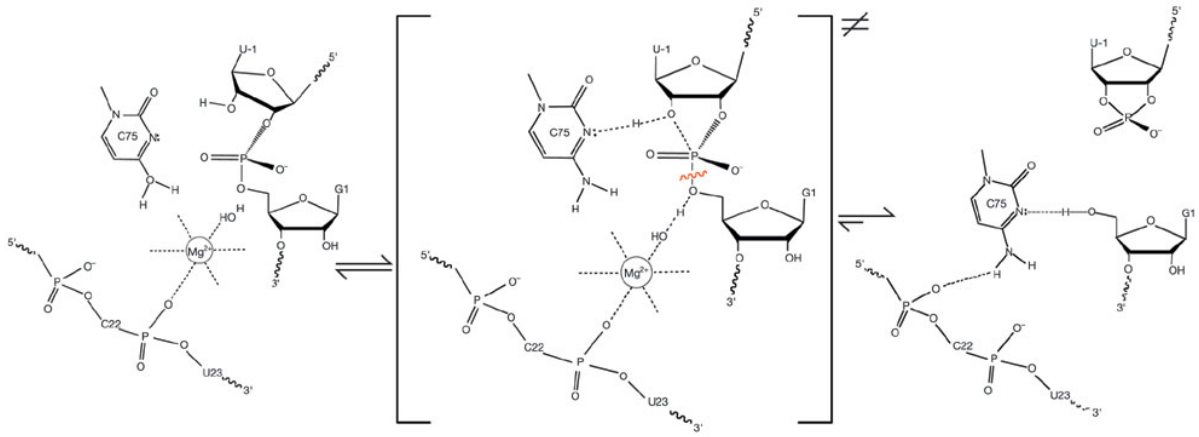

Figure 2. Proposed general acid-base catalysis mechanism. From (Ke, Zhou, Ding, J.H.D., & Doudna, 2003).

The structure of the precursor contains a cytidine nucleotide which acts as base in this mechanism; it deprotonates the 2’-OH of the backbone of the RNA. Simultaneously the hydrated metal ion acting as a general acid, protonates the 5’-O leaving group resulting in the cleavage of the backbone. Conformational changes occur where the catalytic metal ion is discharged and the cytidine is down-shifted to bind the 5’-OH of the substrate resulting in the self-cleaved product.

Other ribozymes have the ability of self-splicing. An example of this group is the Group-I self-splicing introns. These ribozymes are present in a wide range of prokaryote and eukaryote organisms (Nielsen & Johansen, 2009). They catalyze their own excision from mRNA, tRNA and rRNA. The mechanism for these ribozyme also entails a S_N2-type phosphate transfer (Figure 3). Self-splicing requires the attack of the 3’-hydroxyl of exogenous guanosine or the 2’-hydroxyl of adenosine resulting in the spliced product.

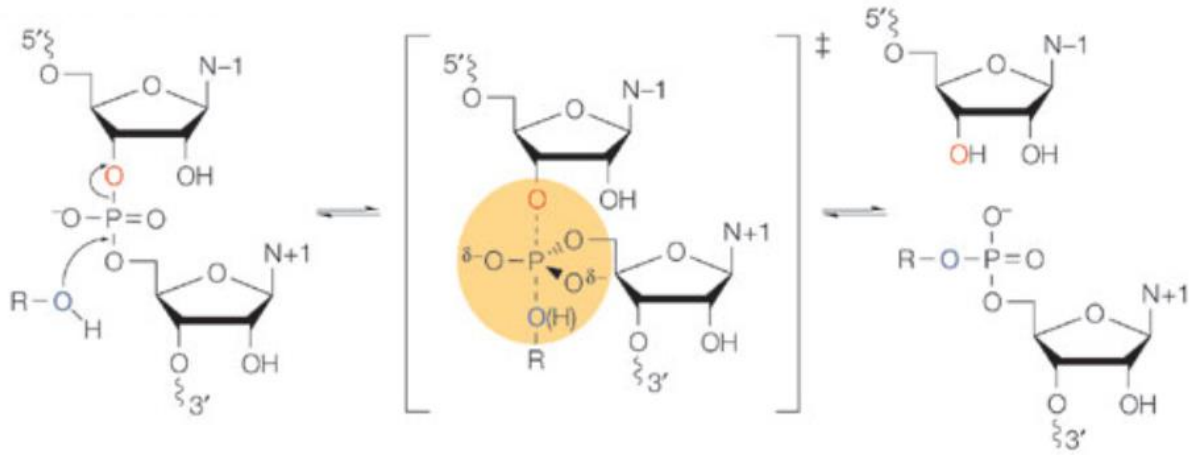

Figure 3. Mechanism for self-splicing of ribozymes. From (Fedor & Williamson, 2005).

Multiple ribozymes have been found in modern life forms, but they can also be engineered with *in vitro* mutagenesis to enhance or change their function. These artificial ribozymes have been an interesting field of study for the past decades.

In vitro directed evolution of RNA molecules

A widely used method to obtain artificial ribozymes is *in vitro* directed evolution. This approach makes use of Darwinian evolution of catalytic RNA *in vitro*. Besides obtaining many interesting ribozymes, also Darwinian evolution itself has been investigated with this powerful method (Joyce, 2004). This technique starts with a population of random RNA sequences. Bond forming or breaking reactions between the RNA sequences and chemically tagged substrate enables activity determination; a selection procedure can separate the catalyst based on whether or not they are tagged. For instance, oligonucleotides attached to the substrates can be used as tags. Bond forming or bond breaking reactions will respectively increase or decrease the weight of the substrate confirming whether or not and which RNA sequence shows actual catalytic activity. Also biotin-tagged catalyzed products can be captured by streptavidin enabling the isolation of active RNA sequences. Reverse transcription is used to transform the captured RNA sequences into cDNA which subsequently is ready for amplification.

PCR is used to amplify the sequence up to a population of 10^{12}-10^{14} individuals. Another isothermal method for amplification is using the enzymes reversed transcriptase, RNase H, and a DNA-dependent RNA polymerase (Guatelli, Whitfield, & Kwoh, 1990). Reversed transcription results in cDNA which together with the RNA is amplified by the DNA dependent RNA polymerase by using forward and reversed transcription. The final amplification method is to clone the active sequences into bacteria and after the growth to a large population, harvest a 10^{6}-10^{8} fold amplification of the DNA sequences.

An essential aspect of Darwinian evolution is the introduction of mutations in sequences so a new species can emerge. Starting with a random population of RNA sequences will generate genetic diversity but introducing errors during the amplification process combined with selection will entail the full characteristics of Darwinian evolution. This complete cycle of directed evolution was optimized and has led to the development of useful artificial ribozymes. Chapter III elaborates on these results and their applications.

Chapter II: Prebiotic soup and the origin of RNA molecules:

The availability of the compounds in the prebiotic world has led to the discussion about which of the subsystems in biology has evolved first (Sutherland, 2016). The RNA world hypothesis claims that self-replicating RNA systems constituted an important phase in the origin of life as we know it. Another hypothesis claims that proteins catalysis had to be first because of the proteins ability to replicate and polymerize the RNA strands. Because life as we know it consist of individual isolated cells, other theories therefore insist that compartmentation in cells instead of biological molecules was essential at the origin of life. In this chapter the synthetic routes from the prebiotic soup to RNA molecules and to protocells are described.

RNA components from prebiotic molecules

Pentose development for backbone

The backbone of RNA consist of a ribose and a phosphate group. Where a relatively apparent molecule like phosphate probably was present in the prebiotic soup, a more complex and less stable molecule like ribose was probably not or at least not in high concentration. Multiple attempts were performed to synthesize this sugar from well-known C_1-molecules which were available in the prebiotic environment (Decker, Schweer, & Pohlamnn, 1982) (Kim, et al., 2011). These attempts starting from glycolaldehyde resulted in low yields, but pentose molecules were formed including ribose. Crossed aldolisation of glycolaldehyde phosphate with formaldehyde however, did yield pentose-2,4-diphosphate. This compound cyclizes into the stable hemiacetal form (Figure 4.) (Ohloff, 1990).

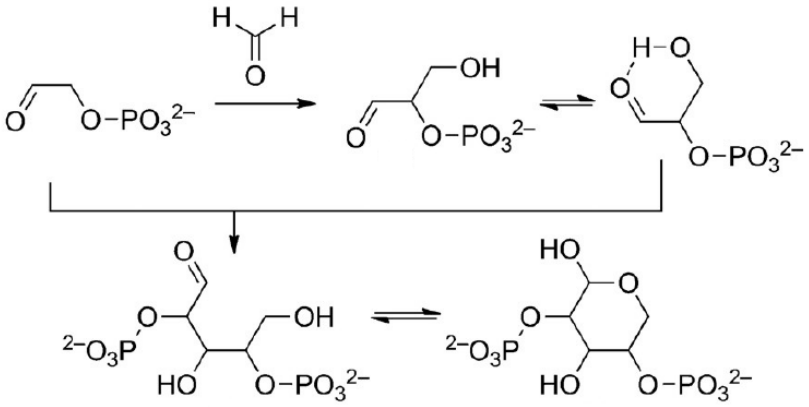

Figure 4: Synthesis of pentose-2,4-diphosphate from glycolaldehyde phosptate and formaldehyde and subsequent cyclization. From (Ohloff, 1990).

This reaction requires C_2- and C_3-sugar molecules. These components are more complex than the known carbon derivatives present in prebiotic environment. The hypothesis about this reaction to be the origin of ribose remains questionable. Recent research has shown the generation of multiple sugars may have occurred in the icy cosmic environment (Meinert, et al., 2016). This astrophysical scenario entails the synthesis of ribose and numerous other sugars using the condition which comets may have encountered including low temperatures (T = 78 K) and low pressure (p = 10^{-7} mbar). This would be a reasonable explanation for the presence of ribose during prebiotic times.

RNA bases from HCN

The development of the four bases: adenine, guanine, uracil and cytosine also remains underdetermined. These bases all consist of cyclic carbon structures, which probably were not present in the prebiotic soup. However, a synthesis of adenine from simply hydrogen cyanide and ammonia was developed by Oró (Oró, 1960) and later optimized by Ferris and Orgel (Figure 5) (Ferris & Orgel, 1966). This synthesis requires only 5 HCN molecules making the presence of adenine during the prebiotic stage very plausible. However, this synthesis requires a concentration of HCN of at least 0.01 M and the reaction is usually performed at pH 9.2 (Shapiro, 1995). This would limit the adenine synthesis to only a few places during the prebiotic period.

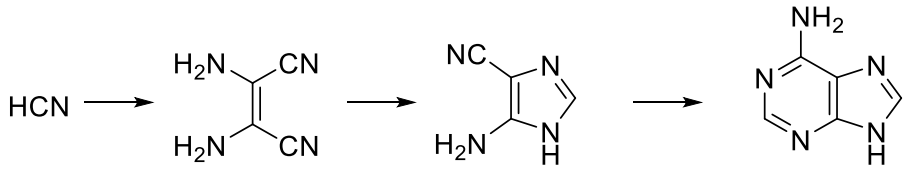

Figure 5: Synthesis of adenine using only 5 HCN molecules. From (Ferris & Orgel, 1966).

A prebiotic route for the complete synthesis of a nucleobase has been proposed (Becker, et al., 2016). Ribose was synthesized from glycolaldehyde and formaldehyde in a more refined fashion than described before. Prebiotically available formamidopyrimidines and molecules derived from NH_4CN were regioselectively coupled to ribose resulting in nucleobase precursor furanoside. This route towards complete nucleobases from prebiotically available compounds is called the FaPy pathway. Starting materials NH_3, NH_4CN and formic acid derivatives have been found on comets, suggesting that this synthesis might be a very reasonable explanation for the origin of RNA precursors.

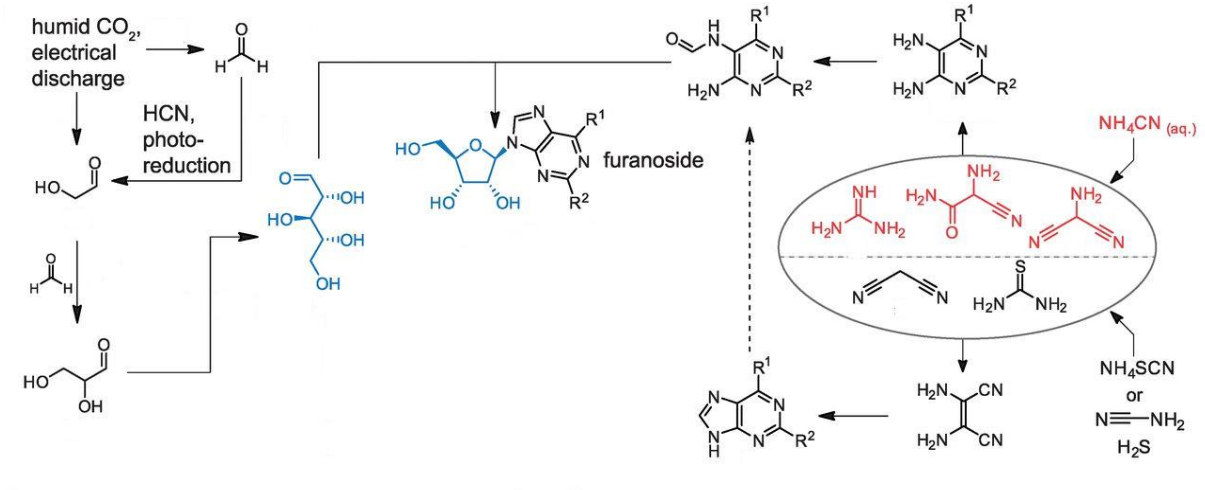

Figure 6: FaPy pathway describing the synthesis of nucleobase precursor furanoside from mere prebiotic compounds. Adapted from (Becker, et al., 2016).

This sophisticated pathway describes the plausible origin of nucleotides, however there are still many questions left. How would these nucleobases polymerize? How would the relatively weak hydrogen bonding between nucleobases remain intact and how would the individual RNA molecules compartmentalize and be able to undergo Darwinian evolution?

Protocells

Compartmentalization is essential in early stages of life's development. Membranes form a semipermeable barrier which prevents small molecules from entering the cells. Also, from an evolutionary perspective, membranes separate different genomes, allowing them to have a barrier against inactive parasites (Chen & Walde, 2010). Functional primitive synthetic cells embodied by a membrane have not been developed yet (Chen, 2006). To understand the issues of this process, the amphiphilic components of membrane vesicle have to be investigated. The membrane of modern cells consists of a phospholipid bilayer. These phospholipids entail a hydrophilic headgroup and two acyl chains. These membranes show slow kinetics and impermeability, making them inadequate for early life forms (Paula, Volkov, Van Hoek, Haines, & Deamer, 1996). There are other known amphiphiles having multiple advantages over each other (Chen & Walde, 2010). Single-chain amphiphiles have fast kinetics, high water solubility and high permeability. Primitive cellular life forms (protocells) probably had membranes comprising these primitive amphiphilic compounds.

The most plausible compounds in the prebiotic membranes were fatty acids. These amphiphilic compounds were found in meteorites (Lawless & Yuen, 1979) and may have originated from prebiotic compounds (Mansy, 2009). Vesicles from fatty acids show various similarities with phospholipid membranes including: vesicle size, thermostability and their ability to contain intravesicular protein and RNA (Mansy, 2009). However fatty acids have a much higher degree of flip-flop dynamics than phospholipid membranes allowing them to be more permeable. This allows the fatty acids to uptake critical ingredients at a higher rate which may have been an advantage at early life stages (Monnard & Deamer, 2001).

Synthesis of fatty acids in hydrothermal conditions from formic acid and oxalic acid resulted in multiple fatty acids ranging from C_2 to C_{35}-fatty acids (McCollom, Ritter, & Simoneit, 1999). Due to their amphiphilic properties, fatty acids have the tendency to self-assemble into bilayer vesicles in water. Dissolving fatty acids (sodium octanoate and oleic acid) in water resulted in spontaneous self-assembly into micelles and vesicles and later self-reproduction of these vesicles (Blöchliger, Blocher, Walde, & Luisi, 1998). The probability of protocells to have fatty acid membranes is therefore high. If these fatty acid vesicles were the origin of protocells the next question is how RNA would have been able to enter these vesicles and start showing characteristics of early life forms.

RNA in protocells

While the synthesis of RNA from prebiotic components is still questionable, the concept of how this RNA was encapsulated by vesicles to allow Darwinian evolution is also investigated. Multiple hypothesis about this "coupling of phenotype and genotype" are suggested. RNA polymerization may have occurred in a clay called montmorillonite (Hanczyc, Fujikawa, & Szostak, 2003). This clay catalyzes RNA polymerization but also accelerate the conversion of fatty acids into micelles and vesicles. This encapsulated RNA did not show leakage for over 24 hours. An early RNA catalyzing prebiotic membrane synthesis may have been encapsulated. This would allow the growth and division of early protocells (Murtas, 2013). If in later stages also RNA replicators have been encapsulated, the division of protocells containing RNA would be realistic. From this point on, Darwinian evolution could take place.

The origin of RNA from mere prebiotic compounds and the development of then concept division and replication still remains an open question in the RNA world hypothesis. Another approach of investigation early stages of life is to conduct research from a top-down perspective. The development of RNA replication systems would give a great insight in how early life has evolved.

Chapter III: Biological approach to develop a self-replicating RNA system

Although we still do not know much about the exact origin of RNA, top-down research can give clues about what RNA molecules are required for catalysis and eventually self-replication. Catalytic properties of RNA enzymes found in nature have been described in Chapter I. Techniques to develop artificial ribozymes allow investigation in the catalytic properties for ribozymes essential for early life forms. So far, no self-replicating RNA system has been developed capable of complete self-replication with the efficiently, fidelity and rate to be a potential ribozyme responsible for the origin and replication of early protocells.

Catalytic activity of RNA

Self-cleaving or self-splicing ribozymes described in the introduction may give clues for the mechanism used by the early self-replicating ribozymes. For better understanding self-replicating RNA systems, the structure of catalytic RNA has to be investigated. Artificial ribozymes obtained from *in vitro* directed evolution experiments have enhanced catalytic activity. *In vitro* selection is applied to obtain ligands called aptamers. These aptamers specifically bind ions, small molecules, organelles or even entire cells (Wilson & Szostak, 1999). The structure of aptamers binding small molecules can give insight in how RNA was able to recognize molecules essential for early life forms.

An early *in vitro* developed aptamer containing 15 nucleobases recognizes theophylline (Zimmermann, Shields, Jenison, Wick, & Pardi, 1998). This aptamer is a short hairpin with two small internal loops. Mere five nucleobase pairs play a role as a pocket binding theophylline with high affinity (Figure 7).

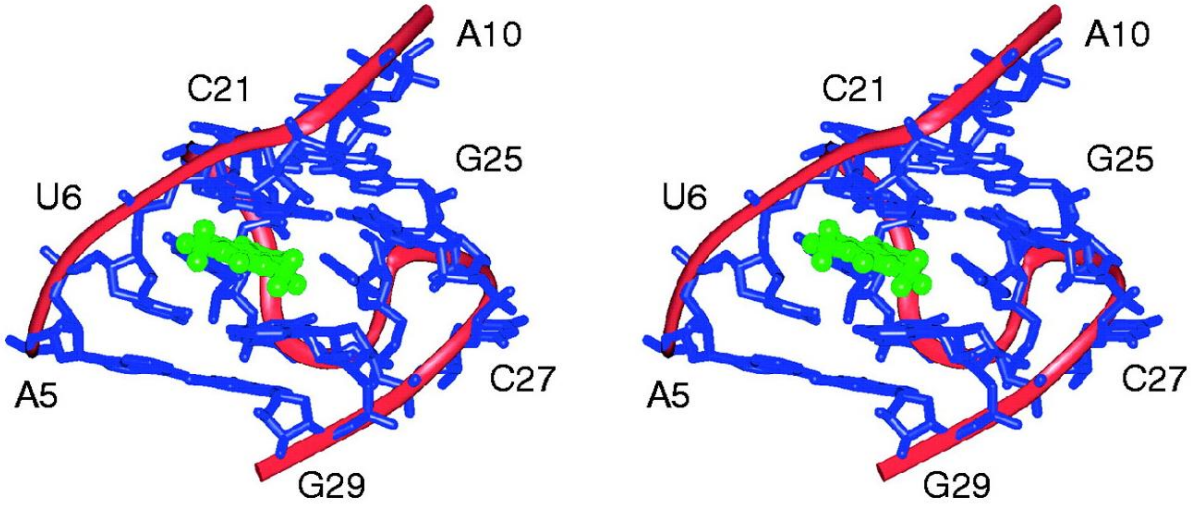

Figure 7: Binding pocket of an aptamer binding the substrate theophylline. From (Wilson & Szostak, 1999).

Where aptamers specifically recognize compounds, ribozymes also catalyze their target. This requires more sophisticated RNA molecules. To understand the range of catalytic possibilities of RNA, abundant investigation is conducted.

Many known ribozymes use phosphodiester transfers to catalyze itself. Catalysis entails the acceleration of a reaction. Therefore it may sound rather counterintuitive when the catalyst consumes itself, since the reaction would stop further catalysis. Full regeneration in its original form

would be essential to meet all the criteria of an enzyme. To accomplish full regeneration, the RNA molecule is separated into two distinct parts: the catalyst part and the substrate part. This enzymatic mechanism is achieved by adding an oligonucleotide to the RNA molecule (Zaug & Cech, 1986). This resulted in a model where the active RNA site RRRRRR binds and cleaves the oligonucleotide C_5, forming a covalent bound intermediate (Figure 8). Subsequently another C_5-oligonucleotide binds resulting in the hydrolysis of the intermediate regenerating the original RNA molecule. This cleavage-ligation process is catalyzed with K_m = 42 µM, k_{cat} = 2 min^{-1}, and k_{cat}/K_m = 1 x 10^3 sec^{-1}. These values resemble the rates of known proteins catalyzing chain cleavage or polymerization initiation.

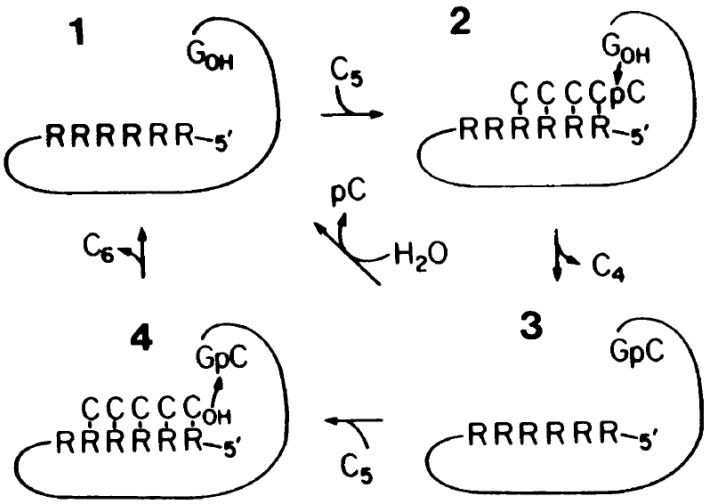

Figure 8: Model for enzymatic RNA catalysis. From (Zaug & Cech, 1986).

Other more complex models for ribozyme activity were discovered showing different mechanisms. Rather than using directed evolution, engineering of RNA based on rational design and known RNA structures has led to more sophisticated ribozymes with broad functionality and application.

Template-directed polymerization model

The ability of catalyzing ligation of ribozymes could support the theory of a RNA polymerization model. Although still not fully understood, the modern protein polymerization process could give clues about how RNA polymerases would function. Modern DNA or RNA polymerases recognize double-stranded nucleic acids in a sequence independent matter (McGinness & Joyce, 2003). This would be an important feature for early ribozyme polymerases.

In vitro evolution is employed to develop a ribozyme capable of polymerization (Johnston, Unrau, Lawrence, Glasner, & Bartel, 2001). Starting from a pool of 10^{15} RNA sequences 10 rounds of *in vitro* selection resulted in a RNA molecule able to extent a primer template (Figure 9). This was tested by using radiolabeled NTPs for polymerization and subsequent analyses with gel electrophoreses. The ribozyme was tested with multiple primer-template template pairs and in each case it was able to extent the primer with matching base pairs (Figure 9A).

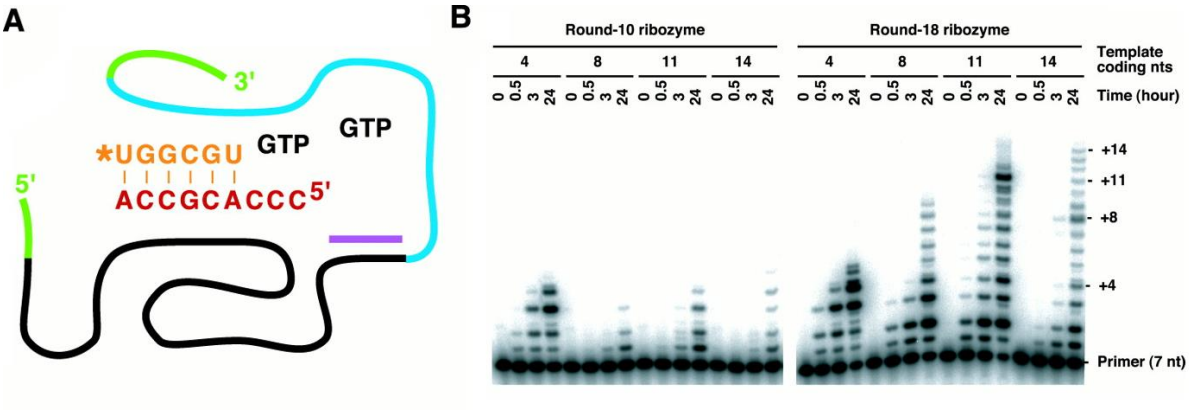

Figure 9: A: Sequence of the active site of the RNA polymerase. B: Gel electrophoresis of round-10 and round-18 ribozyme showing that round-18 ribozyme is able to polymerize templates up to 15 nucleotides. Adapted from (Johnston, Unrau, Lawrence, Glasner, & Bartel, 2001).

For optimization of this ribozyme another 8 rounds of vitro selection were performed resulting in ribozyme R18 which capable of extending the primer up to 14 nucleotides (Figure 9B). The overall fidelity of the polymerization process by this artificial ribozyme was 0.976. This approaches viral RNA polymerases replicating with 0.996 fidelity. Although this ribozyme has polymerase properties, it is still sequence-dependent and needs further optimization to approach a ribozyme able to replicate full RNA molecules.

Further *in vitro* selection and engineering of this ligase ribozymes resulted in the ribozymes capable of synthesizing RNAs up to 95 nucleotides in length (Wochner, Attwater, Coulson, & Holliger, 2011). Full secondary structures of these ribozymes could clarify more about their activity (Figure 10). Unlike the R18 ribozyme, this new ribozyme possesses the yellow-marked ssC19 region. This region has therefore shown to be essential for increased activity. Further selection experiments also improved sequence generality, decreasing the sequence-dependency. This allows the ribozyme to replicate more templates and therefore this ribozyme may show characteristics comparable with replicases in early life forms. However to generate a fully functional RNA replicase ribozyme, the fidelity and efficiency of the primer extension have to be improved.

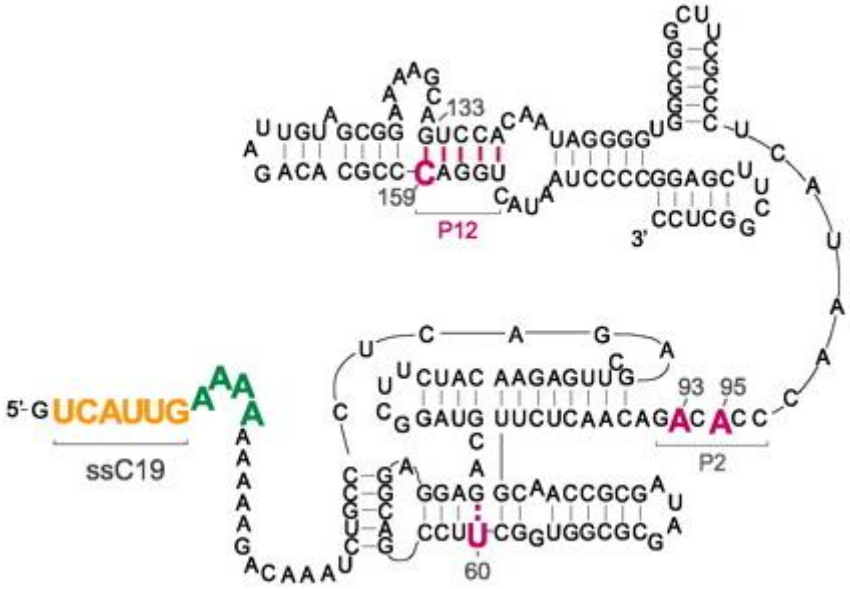

Figure 10: Secondary structure of the engineered R18 ribozyme. From (Wochner, Attwater, Coulson, & Holliger, 2011).

Highly efficient self-replication

Isothermal amplification

A key core of replication is the amplification of the RNA molecule. Modern methods for amplification of DNA and RNA include PCR or LCR, which require multiple temperature cycling (Wiedmann, et al., 1994) (Bartlett & Stirling, 2003). However, the prebiotic environment did probably not entail this regulated temperature cycling. Therefore an isothermal amplification of DNA and RNA method using constant temperatures was developed as described in the introduction (Guatelli, Whitfield, & Kwoh, 1990). This *in vitro* method to amplify RNA in a retroviral fashion was not only useful for the detection and nucleotide sequence analysis of ribozymes but enabled a novel approach for developing self-replicating ribozymes using cross replication.

Cross replication

By enabling the amplification to proceed without the use of any enzymes, a system using cross replication was developed (Kim & Joyce, 2004). This gave rise to a novel replication system using self-sustained exponential amplification. Using *in vitro* evolution a previous developed RNA ligase (R3C ligase) was catalytically improved to optimize the two components in this cross replication (Lincoln & Joyce, 2009). Enzyme E catalyzes the assembly of the two substrates A' and B' forming E' by forming a 3',5'-phosphodiester linkage. E' subsequently catalyzes the assembly of A and B. The kinetic properties of this system were later investigated (Ferretti & Joyce, 2013) (Figure 11). By using different starting concentrations for this reaction, the rate limiting step has shown to be the availability of the free substrates. The exponential self-sustained replication occurs at a rate of 0.03 min^{-1}.

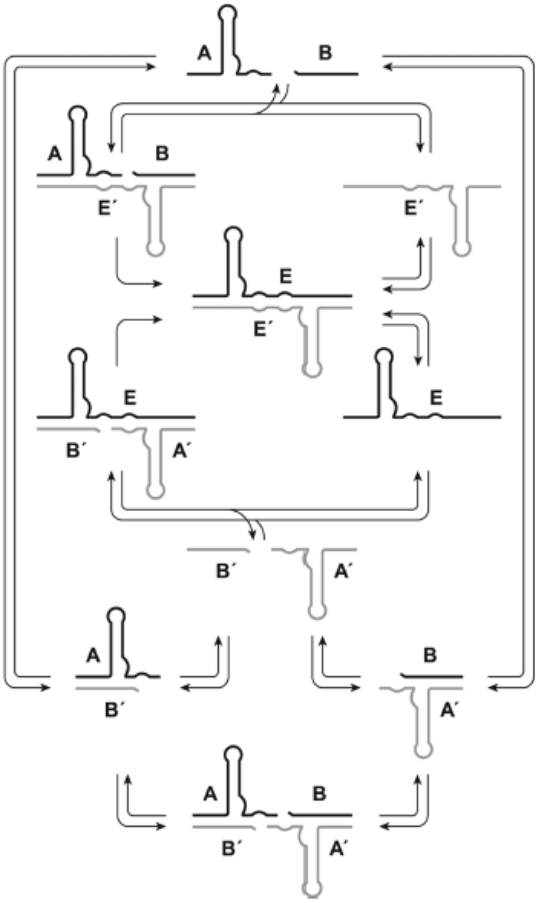

Figure 11: Cross replicating system where assembly of A' and B' is catalyzed by E and assembly of A and B is catalyzed by E'. From (Ferretti & Joyce, 2013).

In vitro directed evolution of this ribozyme improved the rate of this to 0.14 min^{-1} which corresponds to a doubling time of 5 minutes, making this system to be a great model for Darwinian evolution during the RNA world (Robertson & Joyce, 2014).

Another system uses the recognition of a ligand for initiation of a self-replicating process (Lam & Joyce, 2009). When the ligand binds the aptazyme, the catalytic domain is assembled enabling the subsequent ligand-independent cross replication of the RNA enzyme at an exponential rate (Figure 12). The ligand concentration has shown to be the rate limiting step for this reaction (Lam & Joyce, 2011). Therefore this autocatalytic aptazyme can be used for quantitative detection for target ligands (Olea & Joyce, 2015). Also, the ligand could been used for regulation of the cross replication during the RNA world.

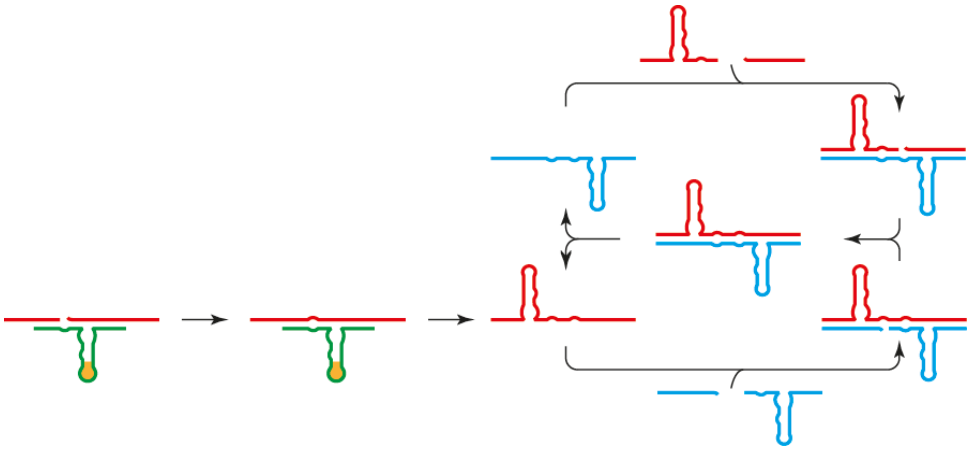

Figure 12: Ligand-dependent catalysis and subsequent ligand-independent cross replication. From (Lam & Joyce, 2011).

The possible applications for these ribozymes have yet to be discovered. However they may give us further insight about how early life forms may have arisen and sustained themselves during the RNA world.

Chapter IV: Conclusively elaborating on the likelihood of the RNA world hypothesis

What are the main problems of RNA synthesis from the prebiotic world?

There is strong evidence for RNA world before life had emerged on earth, but still many questions remain unanswered. One of the question is whether or not all the ingredients for RNA were present in the prebiotic pool. Chapter II shows that nucleobases of RNA, the ribose phosphate backbone of RNA and fatty acids may have been present in the prebiotic environment. There are also indications that comets have brought some of the necessary ingredients facilitating the origin of early life forms.

Another issue which needs further research entails the stereospecificity of RNA and DNA in modern life forms. All known life on earth today contains L-RNA and there is no particular explanation why there is no D-RNA present. It is generally believed that modern life has evolved from molecules with the same stereospecificity. However, recently a cross chiral RNA polymerase was developed (Sczepanski & Joyce, 2014). Both D- and L-RNA catalyze the assembly of the full length copies of their own chirality. This finding may give clues about why only one enantiomer has remained in modern life forms.

The two nucleophilic 2' and 3' hydroxyl groups of a ribonucleotide can both attack the activated 5'-phosphate of another nucleotide. Natural ribonucleotide structures have 3'-5' linkages, but RNA polymerization yields a mixture of 3'-5' and 2'-5' linkages (Szostak, 2012). It may be possible that early life stages possessed backbone heterogeneity, however, no specific reason has been found why mere 3'-5' linkages sustained in modern life.

There is strong evidence indicating that fatty acids may have been present in the prebiotic pool. The presence of vesicles is therefore imaginable. First protocells entailing a membrane of fatty acids may have played a role in de development of protocells. There are clues about the simultaneous development of RNA and fatty acids on the clay montmorillonite (Hanczyc, Fujikawa, & Szostak, 2003). However, RNA replication requires nucleotides for replication based on primer-extension. This is incompatible with the encapsulated RNA model since these protocells are not able to take up exogenous RNA (Szostak, 2012). How protocells would have obtained all features of life therefore still remains unknown.

Investigation of the components present in the prebiotic soup has given many clues about the origin of life, yet still many questions still have to be answered before the RNA world hypothesis can be confirmed.

How far are we with complete Darwinian RNA self-replication?

In vitro directed evolution has led to many novel functional ribozymes. Artificial aptazymes, RNA polymerases and RNA replicases have shown the catalytic possibilities of RNA. The cross-replicating ribozyme model described in Chapter III may give clues about the ability to replicate information in the RNA world. The rate and fidelity of this cross-replication approach the rate and fidelity of modern replicases. This RNA replicase is able to undergo exponential amplification and could therefore enable self-sustained Darwinian evolution. Although this ribozyme is not able to invent new RNA functions, it may lead the research towards new directions closer to the ultimate goal for a Darwinian engineer: Find a self-sustaining system that can evolve on its own.

References

Bartlett, J. M., & Stirling, D. (2003). A Short History of the Polymerase Chain Reaction. *Methods in Molecular Biology, 226(2nd)*, 3–6.

Becker, S., Thoma, I., Deutsch, A., Gehrke, T., Mayer, P., Zipse, H., & Carell, T. (2016). A high-yielding, strictly regioselective prebiotic purine nucleoside formation pathway. *Science, 352*(6287), 833-836.

Bernhardt, H. (2012). The RNA world hypothesis: the worst theory of the early evolution of life (except for all the others). *Biology Direct, 7*, 23.

Blöchliger, E., Blocher, M., Walde, P., & Luisi, P. (1998). Matrix Effect in the Size Distribution of Fatty Acid Vesicles. *J. Phys. Chem. B, 102*, 10383-10390.

Chen, I. (2006). The Emergence of Cells During the Origin of Life. *Science, 314*(5805), 1558-1559.

Chen, I., & Walde, P. (2010). From Self-Assembled Vesicles to Protocells. *Cold Spring Harb Perspect Biol, 2*(7), 1-13.

Decker, P., Schweer, H., & Pohlamnn, R. (1982). Bioids : X. Identification of formose sugars, presumable prebiotic metabolites, using capillary gas chromatography/gas chromatography—masspectrometry of n-butoxime trifluoroacetates on OV-225. *Journal of Chromatography A, 244*, 281-291.

Fedor, M., & Williamson, J. (2005). The catalytic diversity of RNAs. *Nature Reviews Molecular Cell Biology, 6*, 399-412.

Ferré-D'Amaré, A., Zhou, K., & Doudna, J. (1998). Crystal structure of a hepatitis delta virus ribozyme. *Nature, 395*, 567-574.

Ferretti, A., & Joyce, G. (2013). Kinetic Properties of an RNA Enzyme That Undergoes Self-Sustained. *biochemistry, 52*, 1227-1235.

Ferris, J., & Orgel, L. (1966). An Unusual Photochemical Rearrangement in the Synthesis of Adenine from Hydrogen Cyanide. *J. Am. Chem. Soc., 88*(5), 1047-1047.

Guatelli, J., Whitfield, K., & Kwoh, D. (1990). Isothermal, in vitro amplification of nucleic acids by a multienzyme reaction modeled after retroviral replication. *Proc Natl Acad Sci USA, 87*(5), 1874–1878.

Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., & Altman, S. (1983). The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell, 35*(3), 849-857.

Hanczyc, H., Fujikawa, S., & Szostak, J. (2003). Experimental Models of Primitive Cellular Compartments: Encapsulation, Growth, and Division. *Science, 302*(5645), 618-622.

Johnston, W., Unrau, P., Lawrence, M., Glasner, M., & Bartel, D. (2001). RNA-Catalyzed RNA Polymerization: Accurate and General RNA-Templated Primer Extension. *Science, 292*(5520), 1319-1325.

Joyce, G. (2004). Directed evolution of nucleic acid enzymes. *Annual Reviews Biochemistry, 73*, 791–836.

Ke, A., Zhou, k., Ding, F., J.H.D., C., & Doudna, J. (2003). A conformational switch controls hepatitis delta virus ribozyme catalysis. *Nature, 429*, 201-205.

Kim, D., & Joyce, G. (2004). Cross-catalytic replication of an RNA ligase ribozyme. *Cell chemical biology, 11*, 1505-1512.

Kim, H., Ricardo, A., Illangkoon, H., M.J., K., Carrigan, M., Fry, F., & Benner, S. (2011). Synthesis of Carbohydrates in Mineral-Guided Prebiotic Cycles. *J. Am. Chem. Soc., 133*(24), 9457–9468.

Kruger, K., Grabowski, P., Zaug, A., Sands, J., Gottschling, D., & Cech, T. (1982). Self-splicing RNA: Autoexcision and autocyclization of the ribosomal RNA intervening sequence of tetrahymena. *Cell, 31*(1), 147-157.

Lam, B. J., & Joyce, G. F. (2009). Autocatalytic aptazymes enable ligand-dependent exponential amplification of RNA. *Nature Biotechnology, 27*, 288-292.

Lam, B., & Joyce, G. (2011). An Isothermal System that Couples Ligand-Dependent Catalysis to Ligand-Independent Exponential Amplification. *Journal of the American Chemical Society, 133*, 3191-3197.

Lawless, J., & Yuen, G. (1979). Quantification of monocarboxylic acids in the Murchison carbonaceous meteorite. *Nature, 282*, 396–398.

Lincoln, T., & Joyce, G. (2009). Self-Sustained Replication of an RNA Enzyme. *Science, 323*, 1229-1232.

Mansy, S. (2009). Model Protocells from Single-Chain Lipids. *Int J Mol Sci, 10*, 835–843.

McCollom, T., Ritter, G., & Simoneit, B. (1999). Lipid Synthesis Under Hydrothermal Conditions by Fischer-Tropsch-Type Reactions. *Origins of life and evolution of the biosphere, 29*(2), 153–166.

McGinness, K., & Joyce, G. (2003). In Search of an RNA Replicase Ribozyme. *Chemistry & biology, 10*(1), 5-14.

Meinert, C., Myrgorodska, I., de Marcellus, P., Buhse, T., Nahon, L., Hoffman, S., . . . Meierhenrich, U. (2016). Ribose and related sugars from ultraviolet irradiation of interstellar ice analogs. *Science, 352*(6282), 208-212.

Monnard, P., & Deamer, D. (2001). Nutrient uptake by protocells: a liposome model system. *Orig. Life. Evol. Biosph., 31*, 147-155.

Murray, J., Seyhan, A., Walter, N., Burke, J., & Scott, W. (1998). The hammerhead, hairpin and VS ribozymes are catalytically proficient in monovalent cations alone. *Chem Biol, 5*(10), 587–595.

Murtas, G. (2013). Early self-reproduction, the emergence of division mechanisms in protocells. *Mol Biosyst., 9*(2), 195-204.

Nielsen, H., & Johansen, S. D. (2009). Group I introns: Moving in new directions. *RNA Biology, 6*(4), 375-383.

Ohloff, G. (1990). Chemistry of α-Aminonitriles. Aldomerisation of glycolaldehyde phosphate to rac-hexose 2,4,6,-triphosphates and (in presence of formaldehyde) rac-pentose 2,4-diphosphates: rac-Allose 2,4,6-triphosphates and rac-ribose 2,4-diphosphate are the main reaction. *Helvetica Chimica Acta, 73*(5), 1410-1468.

Olea, C. j., & Joyce, G. (2015). Chapter Two – Ligand-Dependent Exponential Amplification of Self-Replicating RNA Enzymes. *Methods in Enzymology, 550*, 23-39.

Oró, J. (1960). Synthesis of adenine from ammonium cyanide. *Biochemical and Biophysical Research Communications, 2*(6), 407-412.

Paula, S., Volkov, A., Van Hoek, A., Haines, T. H., & Deamer, D. (1996). Permeation of protons, potassium ions, and small polar molecules through phospholipid bilayers as a function of membrane thickness. *Biophys J, 70*, 339–348.

Robertson, M., & Joyce, G. (2014). Highly Efficient Self-Replicating RNA Enzymes. *Chemistry & Biology, 21*, 238-245.

Scott, W., Murray, J., Arnold, J., Stoddard, B., & Klug, A. (1996). Capturing the Structure of a Catalytic RNA Intermediate: The Hammerhead Ribozyme. *Science, 274*(5295), 2065-2069.

Sczepanski, J., & Joyce, G. (2014). A Cross-chiral RNA Polymerase Ribozyme. *Nature, 515*(7527), 440-442.

Shapiro, R. (1995). The prebiotic role of adenine: a critical analysis. *Orig Life Evol Biosph., 25*(1-3), 83-98.

Sutherland, J. (2016). The Origin of Life—Out of the Blue. *Angew. Chem. Int. Ed., 55*, 104 –121.

Szostak, J. (2012). Highly efficient self-replicating RNA enzymes. *Journal of Systems Chemistry, 3*(2), 1-14.

Wiedmann, M., Wilson, W., Czajka, J., Luo, J., Barany, F., & Batt, C. (1994). Ligase chain reaction (LCR) - Overview and applications. *PCR Methods and Applications, 3*, 51-64.

Wilson, D., & Szostak, J. (1999). In Vitro Selection of Functional Nucleic Acids. *Annual Review of Biochemistry, 68*, 611-647.

Wochner, A., Attwater, J., Coulson, A., & Holliger, P. (2011). Ribozyme-Catalyzed Transcription of an Active Ribozyme. *Science, 332*(6026), 209-211.

Zaug, A., & Cech, T. (1980). In vitro splicing of the ribosomal RNA precursor in nuclei of tetrahymena. *Cell, 19*(2), 331-338.

Zaug, A., & Cech, T. (1986). The intervening sequence RNA of Tetrahymena is an enzyme. *Science, 231*, 470-475.

Zimmermann, G., Shields, T., Jenison, R., Wick, C., & Pardi, A. (1998). A Semiconserved Residue Inhibits Complex Formation by Stabilizing Interactions in the Free State of a Theophylline-Binding RNA. *Biochemistry, 37*(25), 1998.