UNIVERSITY OF CALIFORNIA SAN DIEGO

The Origin of Life: Could prebiotically plausible Peptides have aided the Emergence of Ribozymes?

A Thesis submitted in partial satisfaction of the requirements
for the degree Master of Science

in

Chemistry

by

Shayan Poordian

Committee in charge:

Professor Ulrich Muller, Chair
Professor Simpson Joseph
Professor Navtej Toor

2024

The Thesis of Shayan Poordian is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2024

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to express my gratitude to Professor Uli Muller, my advisor and the chair of my committee. Over the past year, their unwavering support and invaluable guidance have been instrumental in the progress of my project. Their insights and teachings have significantly enhanced my abilities as a researcher. Additionally, I am thankful to the members of Professor Muller's lab—Xu Han, Debolina Sarkar, and Tommy Lee—for generously sharing their knowledge and technical expertise with me.

Chapter 1, in full, is a reprint of the material as it appears in Weak Effects of Prebiotically Plausible Peptides on Ribozyme Function. Arriola, Josh T., Poordian, Shayan, Martínez Valdivia, Estefanía, Le, Tommy, Leman, Luke J., Schellinger, Joan G., & Müller, Ulrich F. The Thesis author was the co-investigator and co-author of this paper.

ABSTRACT OF THE THESIS

The Origin of Life: Could prebiotically plausible Peptides have aided the Emergence of Ribozymes?

by

Shayan Poordian

Master of Science in Chemistry

University of California San Diego, 2024

Professor Ulrich Muller, Chair

Prebiotically plausible peptides likely existed during the evolution of catalytic ribozymes. The purpose of my thesis was to test whether prebiotically plausible peptides could benefit a model ribozyme. After testing 10 ribozymes individually in the presence of 10 different peptides, a mild beneficial effect from peptides on ribozyme activity was observed.

INTRODUCTION

Origin of Life:

The question of how life originated on Earth has intrigued many thinkers. To address this question scientifically, we must consider both the history of Earth and the evidence of life on it. Earth is widely believed to have been formed around 4.54 billion years ago (1). Before the 20th century, methods like calculating the time it took for Earth to cool from an initial molten state or sediment accumulation were used to estimate its age, resulting in a wide range of estimates (1).

Later, scientists such as Henry N. Russell used the decay of uranium to lead to estimate Earth's age (2). The invention of the mass spectrometer enabled more accurate estimates. Rutherford used the ratio of 235U to 238U, assuming an initial 1:1 ratio, to estimate Earth's age at 3.4 billion years (3). F. G. Houtermans and C. C. Patterson later used isotope ratios in meteorites as proxies for initial isotope ratios on early Earth, arriving at an estimate of 4.5 billion years (4). The current estimate of 4.54 billion years was determined by F. Tera, on the basis that the lead isotope ratios of different galenas converge at a single point that determines Earth's age (5).

Sometime after the formation of earth, life must have emerged. Microfossils and stromatolites are two generally accepted forms of evidence for early life. Stromatolites are defined as “accretionary sedimentary structures, commonly thinly layered, megascopic and calcareous, interpreted to have been produced by the activities of mat-building communities of mucilage-secreting micro-organisms” (6). However, there is still debate about what constitutes the earliest evidence of life on Earth, because there are differing views on which microfossils and stromatolites are truly biogenic. The earliest commonly accepted evidence of life is the 3.43 billion years old stromatolites of Strelley Pool Chert (SPC) in Australia. Allwood, Walter, Kamber, Marshall, and Burch provided strong evidence for the biogenicity of SPC stromatolites

(7). They identified seven distinct stromatolites with unique morphologies and distributions incompatible with a purely abiogenic origin. The conical shape of the stromatolites, and textural differences between cone and interspace laminae, suggests biogenic processes. Finally, a 250-fold enrichment of rare earth elements in carbonate laminae relative to chert laminae aligns with known biosedimentary signatures.

Therefore, life must have emerged somewhere between 4.5 and 3.5 billion years ago. It is widely accepted that "the origin of life must have resulted from a long process or a series of processes, not a sudden event, for the complexity of a cell could not have appeared instantaneously" (8). One of the leading hypotheses about this intermediate stage in the emergence of life is the RNA world hypothesis.

RNA World Hypothesis**:**

According to the RNA world hypothesis, during the early stages of life's emergence, RNA molecules served both as the genome and as catalysts. The notion that RNA could have played a central role in the origin of life first emerged in the 1960s (9, 10, 11, 12). In contemporary organisms, enzymatic activities are primarily carried out by proteins, while DNA is responsible for preserving genetic information, with mRNA acting as the intermediary between the two. Scientists in the 1960s recognized that it was unlikely for two different complex macromolecules to have arisen abiotically and then started cooperating. Instead, it was simpler to envision a single molecule fulfilling both roles of catalyst and genome (13). Compared to proteins, RNA was considered a better candidate as an information molecule, given that Watson-Crick base pairing provides a plausible mechanism for replication. It was also hypothesized that RNA molecules might have had catalytic properties and that primitive ribosomes could have been composed entirely of RNA (13).

Support for this hypothesis came later when Harry F. Noller and colleagues demonstrated that RNA is the primary catalytic component of modern ribosomes. They discovered that even after the removal of most ribosomal proteins using treatments with proteinase K and SDS, significant peptidyl transferase activity persisted, particularly in ribosomes from *Thermus aquaticus* (14). This resistance to protein extraction, coupled with the loss of activity upon ribonuclease treatment, indicated that peptidyl transferase activity is primarily RNA-based and not dependent on ribosomal proteins, which appear to play only a minor role, if any, in the enzyme's function (14). When the ribozyme structure was first resolved, it was revealed that no amino acid side chain was within 18 nm of the peptide bond being formed, confirming that the ribosome is indeed a ribozyme (15). The fact that the protein-synthesis machinery, common across all species, is a ribozyme supports the idea that there was an earlier stage of life where RNA was the sole catalytic macromolecule and that, at some point, coded proteins took over this role. In other words, since an RNA ribozyme produces all the protein enzymes that facilitate life's processes, it is reasonable to hypothesize that RNA enzymes existed before protein enzymes.

The next piece of evidence for the RNA world hypothesis is the fact that many contemporary protein enzymes use cofactors that are nucleotides, such as ATP and NADH, or nucleotide derivatives like Coenzyme A and S-adenosyl methionine. Harold White suggested that these coenzymes are remnants of a primordial RNA world and were originally the catalytic active sites of ancient ribozymes (16). As translation evolved, the ribozyme scaffolds surrounding these active sites were gradually replaced by protein apoenzymes, leading to the coenzymes' current forms.

Assuming the RNA world hypothesis is accurate, and there was a stage in life's evolution where RNA acted as both genome and catalyst, it remains unclear how these RNA molecules initially formed. In what has been called the "Molecular Biologist's Dream", one can imagine that there was a prebiotic pool with all the nucleotide components present and these components formed RNA polymers *de novo*, and some of these polymers gained the ability to self-replicate (17). Once any self-replicating polymer came into being, Darwinian evolution can explain the emergence of further complexity. However, before the emergence of a system capable of self-replicating and undergoing Darwinian evolution, formation of RNA can be studied under the field of prebiotic chemistry. RNA is a complex molecule, and it is uncertain how it could have formed de novo under prebiotically plausible conditions. One possible route by which RNA molecules could have been formed includes the synthesis of ribose sugar, nucleoside bases, joining of the ribose sugar to the nucleoside base, activation of nucleoside bases through a mechanism such as phosphorylation, and finally polymerization of nucleotides (17). However, the formation of nucleotides could have happened through other routes such as "assembly of the base on a preformed ribose or ribose phosphate, or the coassembly of the base and sugar-phosphate" (17).

One of the early proposed pathways for prebiotic ribose synthesis involves the polymerization of formaldehyde, commonly known as the formose reaction (18), first discovered by Aleksandr Butlerov in 1861 (19). The formose reaction takes place in alkaline conditions and in the presence of some catalyst, usually calcium hydroxide or other divalent metal ions (13). There are several issues with the plausibility of the formose reaction as a prebiotic mechanism for ribose synthesis. The formose reaction produces a complex mixture of products, with ribose being a minor product. Also, the reaction requires high pH and reactant concentrations (20).

The synthesis of purine nucleoside bases is possible through different mechanisms. Synthesis of adenine was first reported by Oro, through heating a solution of NH4CN (21). Related mechanisms of HCN polymerization continued to be investigated and showed promising yields of adenine formation (13), and guanine was also detected in these reactions, but at lower yields compared to adenine (22). One of the concerns with the prebiotic plausibility of HCN polymerization is that HCN might not have been present in necessary concentrations in the ocean of early earth, although some potential solutions such as localized freezing of water have been proposed (23).

Synthesis of pyrimidines through HCN polymerization is not possible. However, pyrimidines have been synthesized through mechanisms that generally involve reactions between cyanoacetylene (IV) or cyanoacetaldehyde (V), and cyanate ions, cyanogen or urea (13). More recent work has used prebiotically plausible reactants including cyanamide, cyanoacetylene, glycolaldehyde, glyceraldehyde and inorganic phosphate, to form activated ribonucleotides through a mechanism involving arabinose amino-oxazoline and anhydronucleoside intermediates (24). Overall, despite the progress made in prebiotic chemistry, significant challenges remain in establishing the prebiotic plausibility of nucleotide synthesis. Given the crucial role of proteins as catalysts in modern organisms, it is also important to consider the prebiotic synthesis of peptides.

Role of Peptides:

In contemporary organisms, proteins carry out almost all enzymatic processes. While the RNA world hypothesis posits that ribozymes preceded the template-directed synthesis of peptides, this does not rule out the possibility that non-coded peptides, synthesized abiotically, existed before the advent of ribosomes and contributed to the evolution of life (25). Amino acids

are simpler to synthesize abiotically under prebiotic conditions than nucleotides, as demonstrated by the famous Miller-Urey experiment (26). In the Miller-Urey experiment, induction of electric discharge in a mixture of methane, ammonia, and hydrogen gasses in a reaction vessel resulted in the detection of biological amino-acids including Gly, Ala and Asp and two non- biological amino acids, β-alanine and α-amino-n-butyric acid (26). In later re-analysis of the reaction mixture of the Miller-Urey experiment, 23 amino acids were detected (27). The Miller-Urey experiment was done under a highly reducing condition, which at the time was thought to be representative of the early earth's atmosphere. However, evidence now suggests that the early Earth's atmosphere was likely not as reducing as initially thought (28). A proposed solution for this problem is the use of prebiotically plausible oxidation inhibitors such as Fe^{2+} to achieve similar synthesis of amino acids under less reducing conditions (28).

Therefore, the early earth likely contained amino acids, which could have formed simple peptides through several prebiotically plausible mechanisms. Rode and colleagues showed that amino acids formed short peptides during wet–drying cycles with acidic pH in the presence of NaCl and Cu2+ (29). Forsythe and colleagues showed a mechanism for peptide bond formation that can generate longer peptides. The mechanism is enabled by α-hydroxy acids, which react with α-amino acids to form oligomers containing a combination of ester and amide linkages (30). More recently, a mechanism involving prebiotically plausible minerals and mechanochemical activation has been proposed for a solvent free route of oligopeptides formation (31).

Therefore, peptides likely existed in the prebiotic earth. Some scientists have suggested a possible role for peptides in the RNA world (32). It is possible that peptides and RNA molecules could have co-operated. The goal of my thesis is to investigate whether prebiotically plausible peptides benefit ribozymes.

References

1. Dalrymple, G. B. The Age of the Earth in the Twentieth Century: A Problem (Mostly) Solved. *Geological Society, London, Special Publications* **190** (1), 205–21 (2001**)**.

2. Henry Norris Russell and James Hopwood Jeans. A superior limit to the age of the earth'scrust. *Proceedings of the Royal Society of London. Series A, Containing Papers of aMathematical and Physical Character*, 99(696):84–86, (1921).

3. Rutherford, E. Origin of Actinium and Age of the Earth. *Nature* **123** (3096), 313–314 (1929).

4. Clair Patterson. The isotopic composition of meteoritic, basaltic and oceanic leads, and the age of the earth. In *Proceedings of the Conference on Nuclear Processes in Geologic Settings*, *Williams Bay, Wisconsin*, (1953).

5. Fouad Tera. Aspects of isochronism in pb isotope systematics application to planetary evolution. *Geochimica et Cosmochimica Acta*, 45(9):1439–1448, (1981).

6. Schopf, J. W. Fossil Evidence of Archaean Life. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 361 (1470), 869–885, (2006).

7. Abigail C. Allwood, Malcolm R. Walter, Balz S. Kamber, Craig P. Marshall, and Ian W.Burch. Stromatolite reef from the early archaean era of australia. *Nature*, 441:714–8, (2006).

8. Preiner, M.; Asche, S.; Becker, S.; Betts, H. C.; Boniface, A.; Camprubi, E.; Chandru, K.; Erastova, V.; Garg, S. G.; Khawaja, N.; Kostyrka, G.; Machné, R.; Moggioli, G.; Muchowska, K. B.; Neukirchen, S.; Peter, B.; Pichlhöfer, E.; Radványi, Á.; Rossetto, D.; Salditt, A. The Future of Origin of Life Research: Bridging Decades-Old Divisions. *Life 10* (3) (2020**)**.

9. A. Rich, On the problems of evolution and biochemical information transfer. *Horizons in biochemistry* **New York, NY**, 103–126 (1962).

10. F. H. C. Crick, The origin of the genetic code. *J. Mol. Biol.* **38**, 367–379 (1968).

11. C. R. Woese, The genetic code. The molecular basis for genetic expression. *Harper & Row, New York* (1967).

12. L. E. Orgel, Evolution of the genetic apparatus. *J. Mol. Biol.* **38**, 381–393 (1968).

13. Leslie E., O. Prebiotic Chemistry and the Origin of the RNA World. *Critical Reviews in Biochemistry and Molecular Biology **39*** (2), 99–123, (2004**)**.

14. Noller, H.; Hoffarth, V.; Zimniak, L. Unusual Resistance of Peptidyl Transferase to Protein Extraction Procedures. *Science **256*** (5062), 1416–19 (1992**)**.

15. Nissen, P. The Structural Basis of Ribosome Activity in Peptide Bond Synthesis. *Science* **289** (5481), 920–30 (2000**)**.

16. H. B. White 3rd, Coenzymes as fossils of an earlier metabolic state. *J Mol Evol* **7**, 101–4 (1976).

17. Robertson, M. P.; Joyce, G. F. The Origins of the RNA World. *Cold Spring Harbor Perspectives in Biology 4* (5), (2012**)**.

18. Mizuno, T.; Weiss, A. H. Synthesis and Utilization of Formose Sugars. *Advances in carbohydrate chemistry and biochemistry*, 173–227, (1974**)**.

19. A. Boutlerow "Formation synthétique d'une substance sucrée" (Synthetic formation of a sugary substance), *Comptes rendus ...* 53: 145–147 (1861).

20. Omran, A.; Menor-Salvan, C.; Springsteen, G.; Pasek, M. The Messy Alkaline Formose Reaction and Its Link to Metabolism. *Life **10*** (8), 125 (2020**)**.

21. Oró, J. Synthesis of Adenine from Ammonium Cyanide. *Biochemical and Biophysical Research Communications **2*** (6), 407–12, (1960**)**.

22. Levy, M.; Miller, S. L.; Oró, J. Production of Guanine from NH4CN Polymerizations. *Journal of Molecular Evolution **49*** (2), 165–168 (1999**)**.

23. Kitadai, N.; Maruyama, S. Origins of Building Blocks of Life: A Review. *Geoscience Frontiers 9* (4), 1117–53 (2018**)**.

24. Powner, M. W.; Gerland, B.; Sutherland, J. D. Synthesis of Activated Pyrimidine Ribonucleotides in Prebiotically Plausible Conditions. *Nature **459*** (7244), 239–42 (2009**)**.

25. Fine, J. L.; Pearlman, R. E. On the Origin of Life: An RNA-Focused Synthesis and Narrative. *RNA*, **29** (8), 1085-98 (2023**)**.

26. Miller, S. L. A Production of Amino Acids under Possible Primitive Earth Conditions. *Science*, **117** (3046), 528–9 (1953**)**.

27. Lazcano, A.; Bada, J. L. The 1953 Stanley L. Miller Experiment: Fifty Years of Prebiotic Organic Chemistry. *Origins of Life and Evolution of the Biosphere*, *33* (3), 235–42 (2003**)**.

28. Cleaves, H. J.; Chalmers, J. H.; Lazcano, A.; Miller, S. L.; Bada, J. L. A Reassessment of Prebiotic Organic Synthesis in Neutral Planetary Atmospheres. *Origins of Life and Evolution of Biospheres*, *38* (2), 105–15 (2008**)**.

29. Schwendinger, M. G.; Rode, B. M. Salt-Induced Formation of Mixed Peptides under Possible Prebiotic Conditions. *Inorganica Chimica Acta*, ***186*** (2), 247–51 (1991**)**.

30. Forsythe, J. G.; Yu, S.-S.; Mamajanov, I.; Grover, M. A.; Krishnamurthy, R.; Fernández, F. M.; Hud, N. V. Ester-Mediated Amide Bond Formation Driven by Wet-Dry Cycles: A Possible Path to Polypeptides on the Prebiotic Earth. *Angewandte Chemie International Edition*, ***54*** (34), 9871–5 (2015**)**.

31. Stolar, T.; Saša Grubešić; Nikola Cindro; Meštrović, E.; Krunoslav Užarević; Hernández, J. G. Mechanochemical Prebiotic Peptide Bond Formation**. *Angewandte Chemie 60* (23), 12727–12731 (2021**)**.

32. van der Gulik, P.; Speijer, D. How Amino Acids and Peptides Shaped the RNA World. *Life **5*** (1), 230–246 (2015**)**.

Chapter 1: Weak Effect of Prebiotically Plausible Peptides on Ribozyme Function

Abstract:

Catalytic RNAs (ribozymes) were central to early stages of life on Earth. The first ribozymes probably emerged in the presence of prebiotically generated peptides because amino acids are generated under conditions that lead to the formation of nucleotides, and amino acids can oligomerize into peptides under prebiotically plausible conditions. Here we tested whether the presence of prebiotically plausible peptides could have aided the emergence of ribozymes, by an in vitro selection of self-triphosphorylation ribozymes from random sequence in the presence of ten different octapeptides. These peptides were composed of ten different, prebiotically plausible amino acids, each as mixture of D- and L-stereoisomers. After five rounds of selection and high throughput sequencing analysis, ten ribozymes that appeared most promising for peptide benefits were tested biochemically for possible benefits from each of the ten peptides. The strongest peptide benefit enhanced ribozyme activity by 2.6-fold, similar to the effect from an increase in the pH by one-half unit. Four arbitrarily chosen ribozymes from a previous selection without peptides showed no significant change in their activity from the ten peptides. These results suggest that prebiotically plausible peptides - peptides without evolutionarily optimized sequence - did not provide a strong benefit for ribozyme activity during the RNA world. This finding stands in contrast to previously identified polycationic peptides, conjugates between peptides and polyaromatic hydrocarbons, and modern mRNA encoded proteins, all of which can strongly increase ribozyme function. The results are discussed in the context of origins of life.

Introduction:

Early stages of life on Earth likely relied on catalytic RNAs to promote self-replication and evolution (1–4). This idea is supported by the finding that many of today's enzyme cofactors are nucleotides or variants thereof (5, 6), and by the idea that RNA's ability to store information as well as catalyze reactions could have been sufficient for an early RNA-dominated life form that was later overtaken by the contemporary, interdependent DNA-RNA-protein systems (1–4). The strongest evidence for this scenario is the ribosome, which is a highly evolved ribozyme and synthesizes all genetically encoded proteins in today's life forms (7, 8). This means that the last universal common ancestor (LUCA) - which contained a complex ribosome (9) - was preceded by an evolving molecular system with sophisticated catalytic RNAs. However, it is unclear how a system based on sophisticated RNAs emerged in a prebiotic environment (10): Based on current prebiotic model reactions for nucleotide synthesis and non-enzymatic RNA polymerization (11–13) and highly controlled experiments that select ribozymes from random sequence (14, 15), it appears unlikely that sophisticated ribozymes emerged de novo in a prebiotic environment. One approach to address this challenge is to investigate the possible help of molecules that likely existed on prebiotic Earth but don't have that function in today's biochemistry (16). An example for this approach lies in □-hydroxy acids, which can catalyze peptide bond formation in wet-dry cycling conditions (17). Analogously, we explore here whether prebiotically plausible peptides containing five biological and five non-biological amino acids, as mixtures of D- and L-isomers could have aided the function of catalytic RNAs. Such peptides are fundamentally different from ribosome-synthesized peptides and proteins because they contain a set of prebiotically plausible amino acids that only partially overlap with biological amino acids, they contain both (biological) L and (non-biological) D stereoisomers,

and their amino acid sequence was not encoded and therefore not optimized by Darwinian evolution.

Amino acids and peptides can be generated under prebiotically plausible conditions, much easier than the building blocks of RNA (compare (17–20) to (12, 21–23)). Since amino acids appear as side products in reactions that may have given rise to nucleotides (24), it is reasonable to assume that the first catalytic RNAs emerged in the presence of amino acids and peptides. Such peptides might have benefitted early ribozymes: Peptides can increase the thermodynamic stability of duplex RNA (25) and improve ribozyme interactions with their RNA substrate either by direct interactions (26) or by co-localizing the ribozyme and its RNA substrate within RNA/peptide coacervates (27). However, these previous studies used L-amino acids with a high content of arginine and lysine, and they tested the effect of peptides on individual, existing ribozymes that were selected in the absence of peptides. A recent in vitro selection study of new ribozymes in the presence of octapeptides composed of the 20 biological L-amino acids, found that conjugates of octapeptides with polyaromatic hydrocarbons can lead to ~900-fold enhancement of ribozyme activity (28). However, these peptides also consisted exclusively of L-amino acids, and included amino acids that would probably not have existed before a macromolecule-catalyzed metabolism. To directly address the question whether prebiotically plausible peptides could have aided the emergence of the first ribozymes, it is necessary to probe RNA sequence space for the existence of ribozymes - for example by in vitro selection from random sequence - in the presence of prebiotically plausible peptides that consist of prebiotically plausible amino acids, and include the chiral amino acids as D- and L-stereoisomers.

The model system used in this study to probe RNA sequence space for the existence of ribozymes is an in vitro selection for self-triphosphorylation ribozymes (29). This in vitro selection system previously yielded more than 300 different ribozyme clusters from 1.7 • 1014 different sequences (30), and has performed robustly in five different selection contexts (28, 29, 31–33). Self-triphosphorylation ribozymes are relevant for prebiotic scenarios because they catalyze the reaction between their 5'-hydroxyl group and the prebiotically plausible energy source 'cyclic trimetaphosphate' (cTmp) (34, 35) to generate 5'-triphosphates. The analogous nucleoside 5'-triphosphates (NTPs) are used today as energy currency in all biological organisms, and were thus likely at the core of life forms even before LUCA. Together, the robust performance and prebiotic relevance make this in vitro selection an ideal model system to study what molecules can increase the frequency of ribozymes in RNA sequence space, in different chemical environments.

Here we describe the in vitro selection of self-triphosphorylation ribozymes in the presence of ten octapeptides, each composed of D/L mixtures of ten different, prebiotically plausible amino acids. After high throughput sequencing analysis of the selected RNA sequences, ten promising ribozymes were tested biochemically with each of the ten peptides individually. The strongest peptide benefit to a ribozyme was a 2.6-fold increase in activity, less than the benefit from increasing the pH by one unit. Similarly, four ribozymes that were previously selected in the absence of peptides were not strongly affected by the same ten peptides, suggesting that prebiotically plausible peptides did not strongly influence the activity of early ribozymes.

Results:

To test whether prebiotically plausible peptides could help the emergence of ribozymes, we performed an in vitro selection of self-triphosphorylating ribozymes in the presence of

peptide octamers containing ten different prebiotically plausible amino acids, covering five biological and five non-biological amino acids (figure 1.1). The frequency of these ten amino acids mirrored their abundance in prebiotic model reactions (*18*, *19*): 22 x glycine, 14 x alanine, 9 x aspartate, 8 x b-alanine, 6 x a-amino butyric acid, 6 x b-amino butyric acid, 5 x valine, 5 x a-amino isobutyric acid, 3 x g-amino butyric acid, 2 x serine. All amino acids with a chiral center were represented as mixtures of D- and L-isomers. The presence of a-, b-, and g-amino acids generated heterogeneous backbones, with different spacing between peptide bonds. Each of the peptides contained between three and six chiral amino acids both as D- and L-isomers, with a total of 232 stereoisomers in the ten peptides. The length of eight amino acids per peptide was chosen because this length is accessible via wet-dry cycling reactions in the presence of a-hydroxy acids (*16*, *17*).

The in vitro selection was conducted as described earlier (29, 33) (figure 1.1), with an effective complexity of 4.2 • 1014 sequences in the first selection round. The RNAs were challenged to catalyze the nucleophilic attack of their 5'-hydroxyl group to the phosphorus of cyclic trimetaphosphate (cTmp), thereby generating a 5'-triphosphate. The incubation with cTmp was done in the presence of the described ten peptides, each at 1 mM concentration. Importantly, there was no selection pressure to utilize these peptides: the procedure only selected for self-triphosphorylation catalysis. Therefore, if a large fraction of the selected ribozymes would depend on peptides, it would show that evolving RNAs adopt to the presence of peptides and suggest that prebiotic peptides may increase the frequency of ribozymes in sequence space. RNAs that successfully catalyzed the reaction during the incubation with cTmp and the ten peptides, and therefore carried a 5'-triphosphate, were isolated via ligation to a biotinylated oligonucleotide and streptavidin-coated magnetic beads, reverse transcribed, and PCR amplified

to complete one cycle of selection. After four rounds of selection, the RNA library was dominated by catalytically active sequences as judged by the decrease of PCR cycles required after reverse transcription (figure S1).

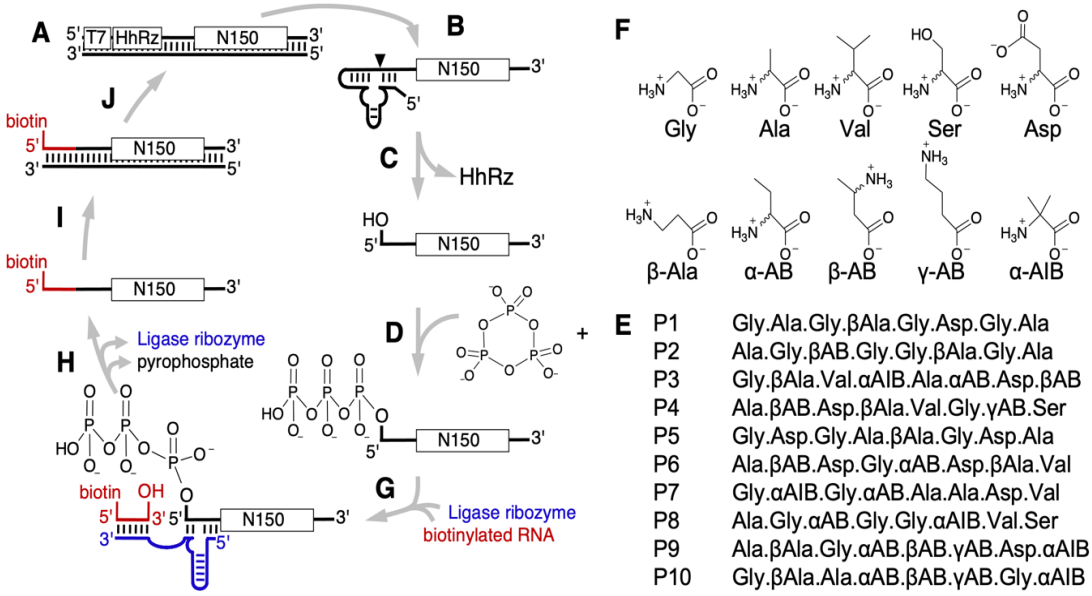

Figure 1: Schematic of the in vitro selection procedure for ribozymes that triphosphorylate their 5'-hydroxyl group in the presence of prebiotically plausible octapeptides.

(A) The DNA library contained a promoter for transcription by T7 RNA polymerase (T7), the sequence encoding a hammerhead ribozyme (HhRz), and 150 nucleotides of randomized sequence (N150) flanked by constant regions (bold lines). (B) Transcription into RNA generated a 5'-terminal hammerhead ribozyme, which (C) self-cleaved co-transcriptionally and thereby generated a 5'-hydroxyl terminus at the RNA library. (D) The RNA library was incubated in the presence of 50 mM Tris/HCl pH 8.3, 100 mM MgCl2, 50 mM Na3cTmp, and each 1 mM of (E) ten different octapeptides that were composed of (F) ten different, prebiotically plausible amino acids, each as mixtures of their D- and L- isomers. (G) Library RNAs that triphosphorylated their 5'-hydroxyl group were ligated to a 5'-biotinylated oligonucleotide (red) using the R3C ligase ribozyme (blue). This allowed their (H) capture and isolation via streptavidin coated magnetic beads. (I) Reverse transcription to DNA and (J) PCR amplification re-generated the DNA library, now enriched for sequences that were able to self-triphosphorylate in the presence of the ten peptides.

High throughput sequencing (HTS) analysis showed that more than 200 sequence clusters were enriched in the selection protocol (figure 2A). To identify whether individual RNAs were

benefitting from the presence of peptides, a fifth selection round was appended, in 22 different variants that differed only in the presence of peptides. The first variant was in the presence of all ten peptides, each at 1 mM concentration. The second variant was done in the absence of peptides. The remaining 20 variants were done in the presence of each one of the ten peptides, at either 1 mM or 10 mM concentration. After HTS analysis of these 22 sub-libraries and the four preceding selection rounds, the enrichment of sequence clusters in the presence of individual peptides was determined (figure S2). To identify ribozymes that may benefit from peptides in their activity, the increase of reads in the presence of individual peptides was determined. No cluster with at least 20 reads showed a stronger peptide-mediated read enhancement than 4-fold, suggesting that none of the ribozyme clusters benefitted more than 4-fold from peptides (the 9.4-fold peptide benefit for cluster 408 could be a result of statistical error because this cluster contained only 16 reads). To test the outcome with the ribozymes that appeared most likely to have a peptide benefit, 30 sequences with an elevated score on both parameters were chosen for further analysis (colored symbols in figure 1.2B; sequences in figure S3).

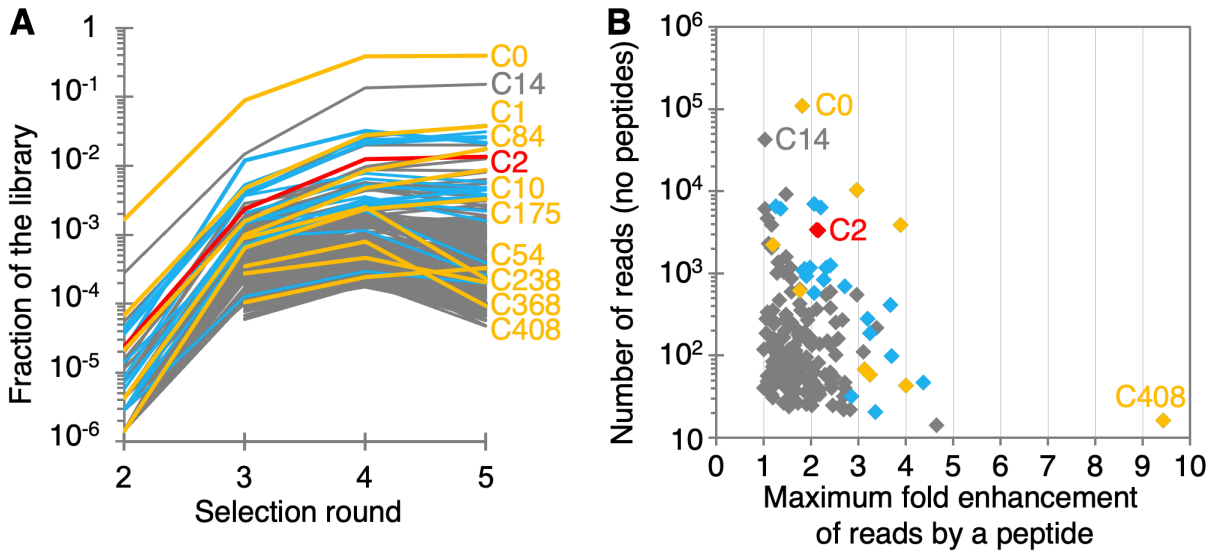

Figure 2: High Throughput Sequencing analysis of selection rounds 2-5.

(A) Enrichment of sequence clusters over multiple rounds of selection. Each line represents one sequence cluster. The number of sequence reads is shown for each of the 212 most abundant clusters with at least 500 reads in the 25 sub-libraries of selection round 5. Lines in blue denote clusters that appeared promising for peptide benefit and were screened in a biochemical assay. Lines in yellow denote ten clusters that were then tested with attention to the individual peptide's influence on reaction pH. The red line denotes cluster C2, which was analyzed in more detail. (B) Predicted peptide benefits for the same sequence clusters as in (A). Each diamond represents one sequence cluster, using the same color coding as in (A), with four clusters labeled to allow correlation with (A). The vertical axis shows the number of reads in the selection sub-round 5 without peptides, while the horizontal axis shows the fold enhancement by the most beneficial peptide. This plot was used to choose sequences for biochemical analysis.

To identify peptide/ribozyme pairs where the peptide may increase ribozyme activity, the 30 chosen sequences were tested for catalytic activity under selection conditions, in the absence and presence of the specific peptides that were suggested as beneficial by HTS analysis (figure S4A). The strongest peptide benefit for ribozyme activity was 1.65-fold, consistent with the idea that under selection conditions, these peptides mediated no, or little benefit to the ribozymes. To test whether peptides could benefit selected ribozymes under suboptimal reaction conditions, the pH in the assay was decreased to pH 7.0, which facilitated about 2-fold peptide benefit for ribozyme S2 (figure S4B). For additional insight into ribozyme S2, it was truncated at its 3'-terminus, reducing its length from 201 nucleotides in the original isolate to 180 nucleotides while

increasing ribozyme activity (figure S5), and its secondary structure was probed using SHAPE (figure S6).

Because the HTS data appeared unreliable to identify small peptide benefits, the biochemical assays were expanded to a larger number of ribozyme / peptide combinations. The activity of ten different ribozymes was measured in the presence of the ten individual peptides, in 100 ribozyme / peptide combinations (figure 1.3A). The effect of the ten peptide preparations on the activity of the ten ribozymes was largely described by their influence on the reaction pH. This pH influence was mostly due to an unintended carryover of ammonium hydrogen carbonate from the purification procedure: NH4HCO3 leads to a pH increase in solution because CO2 has a higher vapor pressure than NH3 in aqueous solution. For nine of the ten ribozymes, the logarithm of ribozyme activity was well-correlated with the reaction pH with a slope of 1.0. This slope is indicative of a single deprotonation event as the rate limiting step of catalysis and was observed previously for a self-triphosphorylation ribozyme (29). This deprotonation is likely the deprotonation of the 5'-hydroxyl group, which is rate limiting for its nucleophilic attack on a phosphorus of cyclic trimetaphosphate. However, ribozyme S2 showed significant deviations from this correlation (yellow symbols in figure 1.3A), with increases of activity in the presence of peptides P4 and P2. Therefore, ribozyme S2 was studied in more detail.

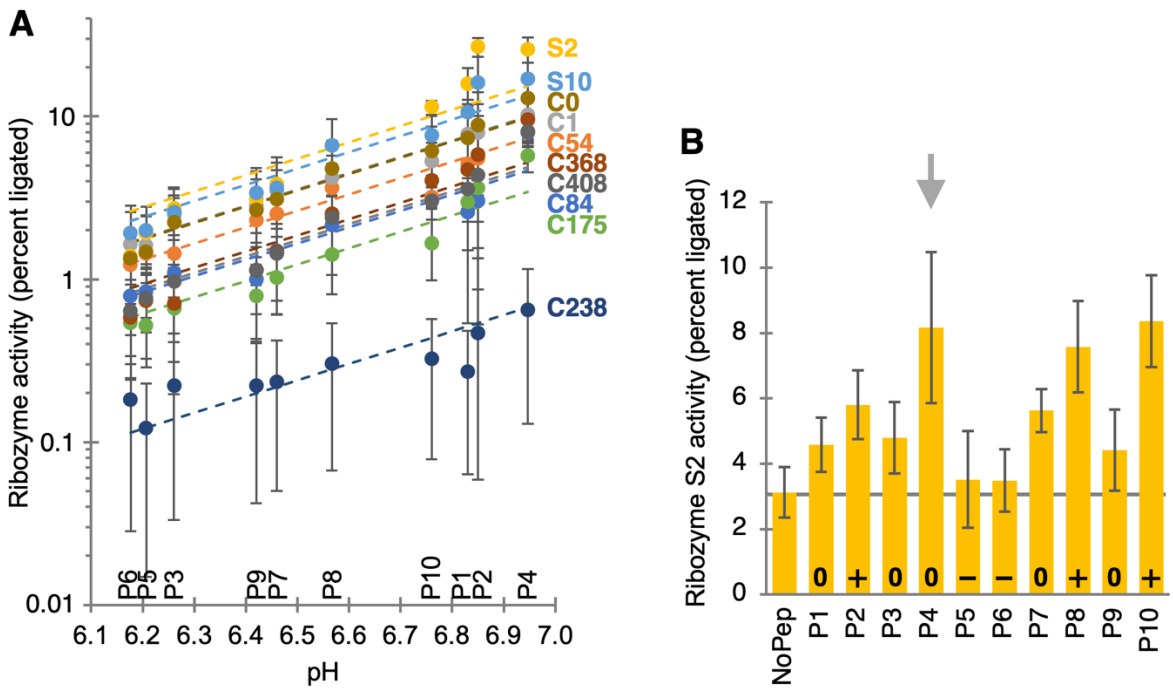

Figure 3: Biochemical analysis of ribozyme activity dependent on the presence of 10 different peptides.

(A) Ribozyme activity plotted on a logarithmic scale as function of the pH shift caused by peptide preparation for each of 10 peptides at 10 mM concentration. The color of each ribozyme's data is annotated with the color of its label on the right. The dashed lines are fitted to each ribozyme's data with a slope of 1.0. Error bars are standard deviations of triplicate experiments. (B) Activity of ribozyme S2 in the presence of 10 mM of each individual peptide, after removing traces of ammonium carbonate from the peptide preparations. The remaining overall charge of each peptide is indicated at the bottom of each column. The peptide without overall charge and the strongest ribozyme benefit (P4) is indicated with a grey arrow. Error bars are standard deviations of triplicate experiments.

To measure the peptide influence on the activity of ribozyme S2 without the confounding influence of ammonium hydrogen carbonate (figure 3B), all peptide preparations were subjected to an additional, 12-hour long lyophilization procedure. The remaining influence of peptide preparations on the reaction pH was therefore caused by the influence of the peptide's amino and carboxy groups, reflected in the peptide's overall charge (the carboxy terminus was used as amide and therefore uncharged). When the ribozyme S2 was then incubated with cTmp in the presence or absence of the re-purified peptides, the three peptides with an overall positive charge

(P2, P8, P10) resulted in the expected 2- to 3-fold increase in ribozyme activity. Peptides 5 and 6 carried an overall negative charge and did not increase ribozyme activity. In contrast, peptide 4 mediated a 2.6-fold increase in ribozyme activity. Peptides P1, P3, P7 and P9 also did not carry an overall charge and also appeared to increase ribozyme activity by an average of 1.6 □ 0.2-fold. Together, this means that uncharged peptides may mediate a less than 2-fold, nonspecific benefit for ribozyme activity, and that peptide 4 increased the activity of ribozyme S2 slightly more, due to specific interactions of peptide 4 with the ribozyme.

All ribozymes tested above had been selected in the presence of the ten studied peptides, and might therefore have adopted to the presence of these peptides. To test whether ribozymes that were selected in the absence of peptides would be influenced by the peptides, four ribozymes that were selected in the presence of only 50 mM MgCl2, 100 mM Na3cTmp, and 50 mM Tris/HCl pH 8.3 (29, 30) were tested in the same buffer but in the presence of the peptides used in this study (figure 1.4). None of the four ribozymes showed a peptide-mediated increase or decrease in activity. Together, these data suggest that most octapeptides consisting of prebiotically plausible amino acids do not significantly increase, or decrease the frequency of ribozymes in sequence space.

Figure 4: Biochemical analysis of the effect of prebiotically plausible peptides on four ribozymes that were selected in the absence of peptides.

Ribozyme activity is shown as a function of the different ribozyme constructs (Ap2, Ap4, Ap5, R8), in the absence (light grey) or presence (dark grey) of the same peptides as used in the selection of the current study. Error bars are standard deviations of triplicate experiments.

Discussion:

To test whether prebiotically plausible peptides can increase the existence and activity of ribozymes in RNA sequence space, an in vitro selection for self-triphosphorylation ribozymes was conducted in the presence of ten prebiotically plausible peptides. This selection identified hundreds of ribozyme clusters, and 10 sequences were analyzed biochemically with each individual peptide for possible benefits. Most peptide effects on ribozyme activity were explained by their effect on reaction pH. A less than 2-fold, non-specific benefit for ribozyme activity was found for four uncharged peptides, while the uncharged peptide P4 mediated a 2.6-fold benefit for ribozyme S2.

The chemical structure of the ten peptides (figure 5) allows to speculate what features of the peptides may be responsible for the effect of peptides on ribozyme S2. However, the small differences between the benefits of uncharged peptides could be caused by interactions weaker

than a single hydrogen bond, therefore the current data set is insufficient to pin-point the feature of peptide P4 mediating the 2.6-fold benefit for ribozyme S2.

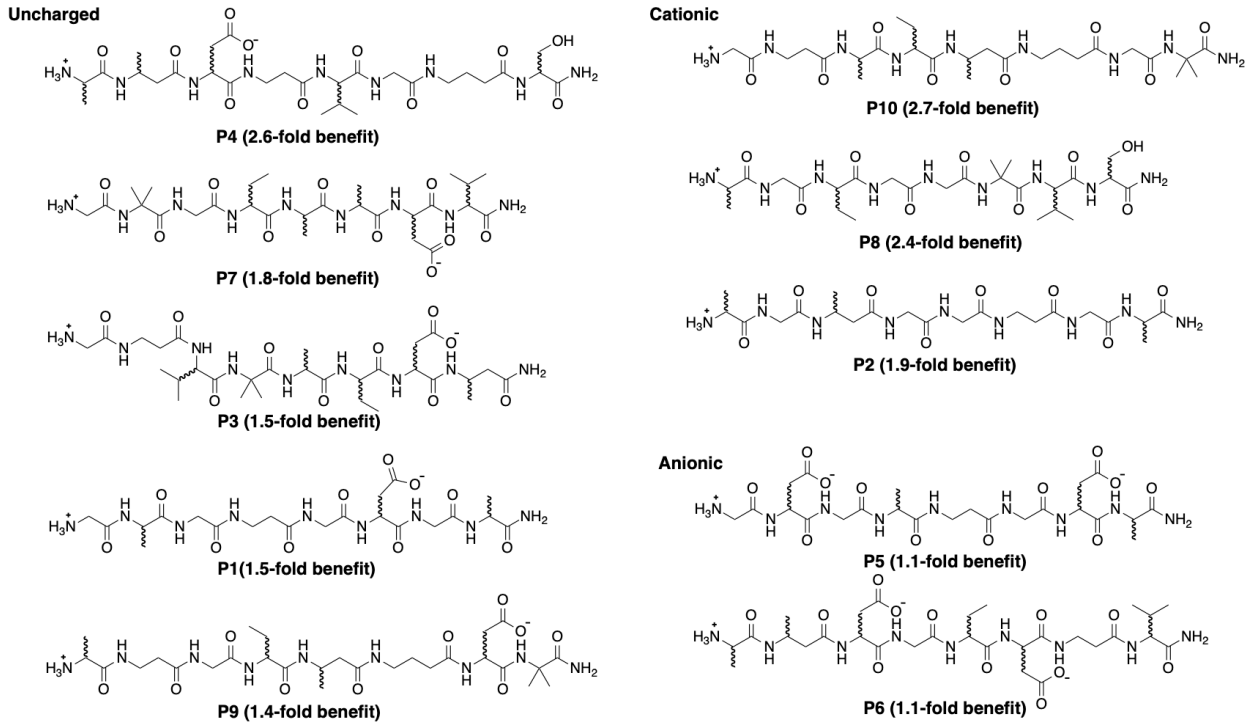

Figure 5: Biochemical analysis of the effect of prebiotically plausible peptides on four ribozymes that were selected in the absence of peptides.

Ribozyme activity is shown as a function of the different ribozyme constructs (Ap2, Ap4, Ap5, R8), in the absence (light grey) or presence (dark grey) of the same peptides as used in the selection of the current study. Error bars are standard deviations of triplicate experiments.

These results of this study suggest that peptides composed of the used ten amino acids did not strongly support ribozyme activity during the RNA world. While this conclusion may sound too broad based on only ten octapeptides, it needs to be considered that prebiotically plausible peptides were not synthesized by a ribosome, and did not have an encoded sequence. If prebiotically plausible peptides could have a genetically encoded sequence, their sequence could have evolved and been optimized for very specific surface and charge complementarity to ribozymes. In the absence of such sequence encoding, the representation of the ten used prebiotically amino acids by the ten used peptides - which neither strongly benefitted nor

inhibited ribozymes - provides an 'average view' that suggests that prebiotically plausible peptides based on these ten amino acids were unlikely to have strongly affected the emergence of catalytic RNAs.

If cationic amino acids such as ornithine and 2,4-diamino butyric acid could have been abundant in prebiotic peptides then the corresponding peptides may have changed the outcome of this study. These amino acids were not included in the current study because their abundance in prebiotic model reactions is far below that of the ten used amino acids (19). However, it is well documented that peptides that are rich in the cationic amino acids arginine or lysine can strongly improve ribozyme activity by increasing the thermodynamic stability of RNA/RNA duplexes (25), stabilizing the sequence non-specific binding of an RNA substrate by a ribozyme (26), or by forming coacervates with the RNA that lead to better RNA/RNA interactions (27). All of these studies required multiple cationic amino acids to benefit ribozyme function, and the requirement of multiple rare amino acids further reduces the peptide's expected abundance in a prebiotic setting. However, very rare peptides could have become enriched and co-localized with RNA by partitioning of polycationic peptides in coacervates with polyanionic RNA. It remains to be seen whether polycationic peptides can be generated, and enriched in prebiotic model reactions in sufficient abundance to provide their benefit to catalytic nucleic acids.

A different type of peptide derivatives may have been important for early ribozymes: Conjugates between polyaromatic hydrocarbons and peptides are able to increase the activity of some ribozymes by 100-fold or more (28). Since insoluble organic matter is the dominant form of carbon in meteorites, and rich in polyaromatic hydrocarbons (36, 37), it is possible that the prebiotic environment was rich in polyaromatic hydrocarbons, and present in prebiotic peptides. However, it is unclear whether this usefulness of PAHs for ribozymes is a rare phenomenon or

whether it may have been wide-spread and important during an RNA world. Future work will show what kind of prebiotically plausible molecules could have aided the emergence and function of an RNA world organism.

Materials and Methods:

Synthesis and purification of peptides:

Peptides were synthesized on rink amide PS resin (0.77 mmol/g, Novabiochem) and Rink Amide ChemMatrix (0.45 mmol/g, Biotage) following the Fmoc/tBu strategy using a microwave-assisted peptide synthesizer (Alstra, Biotage). Side chain protecting groups for amino acids were as follows: Ser(tBu), His(Trt), Lys(Boc), Asp(OtBu), Glu(OtBu), Trp(Boc), Arg(Pbf), AsN(Trt), Thr(tBu), Gln(Trt), Tyr(tBu), Cys (Trt). Typical coupling reactions were performed at 0.200 - 0.300 mmol scale with 3.0 equiv. of Fmoc-protected amino acids, 2.94 equiv. of HBTU and 6.0 equiv. of NMM for 5 min at 75 ºC. All reagents were pre-dissolved in DMF at 0.5 M. Fmoc deprotections were performed with 20% piperidine in DMF for 10 min at rt. Washing was performed after every deprotection and coupling step using DMF. For the syntheses of some sequences, the deprotection and coupling steps described above were either performed twice or with increased reagent equivalence to ensure reaction completion. Cleavage of the peptide sequences from the solid support with concomitant side chain deprotection was accomplished by placing the resin in fritted SPE tube and treating with TFA cleavage cocktail (~20mL/g) containing 90:5:2.5:2.5 TFA:dimethoxybenzene:H2O:TIS for 2 hr. For peptide sequences containing cysteine and methionine, EDT was added at 90:4:2.5:2.5:1.0 TFA:dimethoxybenzene:H2O:EDT:TIS. Cleaved peptides were then precipitated in cold ether, centrifuged, dissolved in methanol and reprecipitated in cold ether (3x).

HPLC characterization and purification of peptides were carried out at rt on analytical (Jupiter C18 5 μm, 300 Å, 150 x 4.6 mm by Phenomenex, Torrance, CA) and semi-preparative columns (Aquasil C18 5 μm, 100 Å, 150 × 10 mm by Keystone Scientific Inc., Waltham, MA) with Prostar 325 Dual Wavelength UV-Vis Detector from Agilent Technologies with Varian pumps (Santa Clara, CA) with detection set at 225 and 406 nm. Peptides were eluted from column following a gradient using mobile phases A: 0.1% TFA in H20 and B: 0.1% TFA in CH3CN. MS analyses were obtained on a LTQ ESI-MS spectrometer (San Jose, CA). Solutions were prepared in either methanol or methanol/water (formic acid 1%) with flow rate of 10 μL/min, spray voltage at 4.50 kV, capillary temperature at 300°C, capillarity voltage at 7.00 V, tube lens at 135.00 V. Purified peptides were characterized by analytical HPLC (with purity typically greater than 90%) and MS (either by direct injection or LC-MS).

To remove a possible carryover of TFA, peptides were dissolved in a totalvolume of 2 to 5 mL 100 mM (NH4)HCO3, and frozen as a thin film on the inside of a glass bulb cooled in a dry ice / isopropanol bath. After desiccating the frozen solution to dryness in oil vaccum (~2 mbar), the process was repeated once with 100 mM (NH4)HCO3, and once with water. The desiccated peptide was weighed and dissolved in water to a stock concentration of 10 mM.

In vitro selection:

A 183 nucleotide (nt) long DNA oligomer that included 150 positions with randomized sequence and flanked by 5'- and 3'-constant regions was purchased from Integrated DNA Technologies (IDT). 5'-GAGACCGCAAGAGAC(N150)CATGGTTCAGACTACAAC-3' A 82 nt long oligomer was used to add the promoter sequence of T7 RNA polymerase and the sequence of the hammerhead ribozyme (underlined) to the 5’ end of the 183-mer. 5’-

AATTTAATACGACTCACTATAGGGAGCGGTCTCCTGACGAGCTAAGCGAAACTGCG GAAACGCAGTCGAGACCGCAAGAGAC-3'. Amplification with the 67-mer and a short 3' primer generated the double stranded DNA pool. qPCR was used to determine the amount of amplifiable DNA and this number was used to estimate that the starting DNA pool consisted of 4.2 x 10^14 sequences.

The DNA pool was transcribed into RNA using T7 RNA Polymerase under standard conditions. 100 nM of purified RNA sub-pool was incubated with 50 mM Tris/HCl, 3.3 mM NaOH, 100 mM MgCl2, 50 mM of freshly dissolved trimetaphosphate and 1 mM each of ten prebiotically plausible peptides (10 mM total). This mixture was allowed to incubate for 3 hours at room temperature. The reaction was quenched by ethanol precipitation. The RNA pellet was washed with chilled water and desalted using P30 Tris/HCl spin columns (Bio-Rad). Active ribozymes were ligated to a biotinylated oligomer (biotin-d(GTAACATAATGA)rU) using the R3C ligase ribozyme whose arms were designed to anneal to the 5' constant region of the pool RNA and the biotinylated oligomer. A mixture containing 1.3 uM of desalted pool RNA, 1 uM ligase ribozyme, 1.2 uM biotinylated oligomer, 100 mM KCl, 100 mM Tris/HCl, and 60 pM triphosphorylated RNA was heated to 65°C for 2 minutes then cooled to 30°C at a rate of 0.1°C per second. The triphosphorylated RNA was used to generate a small amount of ligation product and thereby reduce the number of PCR cycles, and prevent artifacts during PCR amplification. After heat renaturation, an equal volume of a mix consisting of 40% (w/v) PEG8000, 4 mM Spermidine, and 50 mM MgCl2 was added to the ligase reaction mixture. This mixture was incubated at 30°C for 3 hours. After the incubation step, the reaction was quenched by adding final concentrations of 13.9 mM sodium/EDTA, 50 mM Tris/HCl, 50 mM KCl, 0.012% (w/v) Triton X-100, and 1.19 uM of a 61 nt oligomer that was complementary to the ligase ribozyme

(5’ AGTAACATAATGACTTCAACCCATTCAAACTGTTCTTACGAACAATCGAGCACAAGAGAC-3’). This mixture was heated at 50°C for 10 minutes. Streptavadin magnetic beads (Promega) were washed thrice with 20 mM HEPES/KOH pH7.2, 0.01% (w/v) Triton X-100 and 50 mM KCl. Biotinylated RNA was then captured by mixing the RNA with the magnetic beads and rotating end-over-end at room temperature for at least 30 minutes. A magnetic rack was used to focus the beads and the beads were washed twice with a solution containing 0.01% (w/v) Triton X-100 and 20 mM NaOH. Captured RNA was eluted from the beads by incubating the beads in a solution of 25 mM Tris/HCl pH 8.5, 1.56 mM EDTA and 96% formamide at 65°C for 3 minutes. The beads were removed by immediate centrifugation, and the supernatant was concentrated by ethanol precipitation. 1 ug of tRNA was used as precipitation carrier. Captured RNA was resuspended in 10 mM Tris/HCl, pH 8.3. The RNA was reverse transcribed using Superscript III (Invitrogen) and a reverse transcription primer (5’-GTTGTAGTCTGAACCATG-3’) complementary to the 3’ constant region. The products were then PCR amplified with 5’ and 3’ primers that could bind to the sequence of the biotinylated oligomer and the 3’ constant region of the pool respectively. A second PCR amplification was performed to add the T7 promoter sequence and the hammerhead ribozyme sequence to the DNA.

HTS analysis:

Illumina sequencing primers were added to the ends of rounds 2-5 pools. Pools were sequenced on MiSeq with 10% PhiX. FastQ files were processed using custom Python scripts for rounds 2-5 (including all sub-rounds in round 5). The usearch merge function was used to merge forward and revers reads. Sequences not containing both constant regions or short randomized regions

(<135 nt in length or 10% of 150) were first filtered out. The file was reformatted to Fasta for easier downstream processing. The total counts for each sequence were calculated. The usearch clustering algorithm was used to group similar sequences into clusters. Clusters were defined as sequences that shared at least 75% sequence identity and had at least 100 reads in a given round. Clusters and sequences were tracked through each round. 189 sequences were evaluated further. For each sequence, the fold enrichment in the round 5 sub-rounds were determined by normalizing to the no peptide sub-round.

Generation of 3' Truncations:

3' truncations were generated by PCR using Taq DNA polymerase and a series of 3' primers. The 3' primers annealed to different portions of the ribozyme sequence of interest to generate a series of truncations with desired deletions from the 3' end. DNA templates of each truncation were transcribed using T7 RNA polymerase, and purified by denaturing PAGE.

Self-triphosphorylation activity assays:

Purified ribozyme (5 uM) was incubated with 50 mM Tris/HCl, 100 mM MgCl2, 50 mM of freshly dissolved trimetaphosphate, and either a mixture of peptides, individual peptides or water as a control. This mixture was incubated for 3 hours at room temperature. The products were ligated to a radiolabeled oligomer (γ32P-d(GTAACATAATGA)rU) using an R3C ligase ribozyme whose arms were designed to anneal to the 5' constant region of the pool RNA and the oligomer. Following the incubation with buffer, Tmp and peptides, the mixture was diluted 1:10 in a solution that contained a final concentration of 0.5 uM ligase ribozyme, 0.5 uM biotinylated oligomer, 100 mM KCl, 100 mM Tris/HCl, and 15 mM sodium EDTA. This mixture was heat

renatured at 65°C for 2 minutes then cooled to 30°C at a rate of 0.1°C per second. After heat renaturation, an equal volume of a mix consisting of 40% PEG8000 (w/v), 4 mM Spermidine, and 50 mM MgCl2 was added to the ligase reaction mixture. This mixture was incubated at 30°C for 2 hours. After the incubation step, the reaction was quenched by ethanol precipitation. The reaction products were resolved on a denaturing 15% polyacrylamide denaturing gel.

SHAPE probing:

Selective 2'Hydroxyl Acylation Analyzed by Primer Extension (SHAPE) was performed on the 180-nucleotide long ribozyme isolated in this study, with 1-Methyl-7-nitro-2H-3,1-benzoxazine-2,4(1H)-dione (1M7) as the chemical probe. 50 pmol of ribozyme were incubated in buffer containing 50 mM MES/NaOH pH 6.0, 100 mM MgCl2, 3.3 mM NaOH, 50 mM Tmp, and 10 mM peptide 4 at room temperature for 2 minutes. This mixture was incubated at room temperature for 2 minutes or 2 hours. 1M7 was added to this solution such that its final concentration was 2 mM and the concentration of DMSO was 1% of the solution. A negative control was prepared with no 1M7 and 1% DMSO. Both samples were incubated at room temperature for 3 minutes. The reaction was quenched by ethanol precipitation and resuspended in 10 uL of 5 mM Tris/HCl pH 8.0. The products were reverse transcribed using Superscript III reverse transcriptase (Invitrogen) according to manufacturer's instructions and trace amounts of a radiolabeled primers were added. Three reverse transcription primers were used to probe different portions of the ribozyme. The sequences of these primers are (5'-TTAGTCTGAACCATGA-3', annealed to pos. 165-180); (5'-CGATGTTGACATTTTC-3', annealed to pos. 112-127); and (5'-AAACTCAAGTGGGTTT-3', annealed to pos. 53-68). The RNA template in each sample was degraded by alkaline hydrolysis, by incubating for 5 minutes

at 80°C in a solution containing 750 mM NaOH. The reaction was quenched by adding a two-fold stoichiometric excess of acetic acid to generate a NaOAc/HOAc buffer with a final concentration of 300 mM. The products were ethanol precipitated, resuspended in formamide loading buffer and resolved on a denaturing 10% polyacrylamide gel. The RNA secondary structure was predicted with help by the secondary structure prediction tool RNAstructure (ReuterMathews10). To do this, the SHAPE data was mapped onto an initial secondary structure prediction and manually adjusted to better represent the SHAPE data.

Acknowledgement:
We would like to acknowledge Kevin J. Sweeney for advice on performing the in vitro selection in the presence of peptides, and on the HTS analysis.

Chapter 1, in full, is a reprint of the material as it appears in Weak Effects of Prebiotically Plausible Peptides on Ribozyme Function. Arriola, Josh T., Poordian, Shayan, Martínez Valdivia, Estefanía, Le, Tommy, Leman, Luke J., Schellinger, Joan G., & Müller, Ulrich F. The Thesis author was the co-investigator and co-author of this paper and performed the experiments resulting in data shown in figures 1.3 and 1.4.

Funding:
This work was supported by the National Aeronautics and Space Association (NASA) with the award NNX16AJ27G to UFM.

References:

1. A. Rich, On the problems of evolution and biochemical information transfer. *Horizons in biochemistry* **New York, NY**, 103–126 (1962).

2. F. H. C. Crick, The origin of the genetic code. *J. Mol. Biol.* **38**, 367–379 (1968).

3. C. R. Woese, The genetic code. The molecular basis for genetic expression. *Harper & Row, New York* (1967).

4. L. E. Orgel, Evolution of the genetic apparatus. *J. Mol. Biol.* **38**, 381–393 (1968).

5. H. B. White 3rd, Coenzymes as fossils of an earlier metabolic state. *J Mol Evol* **7**, 101–4 (1976).

6. A. D. Goldman, B. Kacar, Cofactors are Remnants of Life's Origin and Early Evolution. *J Mol Evol* **89**, 127–133 (2021).

7. H. F. Noller, V. Hoffarth, L. Zimniak, Unusual resistance of peptidyl transferase to protein extraction procedures. *Science* **256**, 1416–9 (1992).

8. P. Nissen, J. Hansen, N. Ban, P. B. Moore, T. A. Steitz, The structural basis of ribosome activity in peptide bond synthesis. *Science* **289**, 920–30 (2000).

9. A. J. Crapitto, A. Campbell, A. Harris, A. D. Goldman, A consensus view of the proteome of the last universal common ancestor. *Ecology and Evolution* **12**, e8930 (2022).

10. P. Pavlinova, C. N. Lambert, C. Malaterre, P. Nghe, Abiogenesis through gradual evolution of autocatalysis into template-based replication. *FEBS Letters*, 1873-3468.14507 (2022).

11. S. Stairs, A. Nikmal, D. K. Bucar, S. L. Zheng, J. W. Szostak, M. W. Powner, Divergent prebiotic synthesis of pyrimidine and 8-oxo-purine ribonucleotides. *Nat Commun* **8**, 15270 (2017).

12. S. Becker, J. Feldmann, S. Wiedemann, H. Okamura, C. Schneider, K. Iwan, A. Crisp, M. Rossa, T. Amatov, T. Carell, Unified prebiotically plausible synthesis of pyrimidine and purine RNA ribonucleotides. *Science* **366**, 76–82 (2019).

13. F. R. Bowler, C. K. Chan, C. D. Duffy, B. Gerland, S. Islam, M. W. Powner, J. D. Sutherland, J. Xu, Prebiotically plausible oligoribonucleotide ligation facilitated by chemoselective acetylation. *Nat Chem* **5**, 383–9 (2013).

14. D. P. Bartel, J. W. Szostak, Isolation of new ribozymes from a large pool of random sequences. *Science* **261**, 1411–8 (1993).

15. A. Akoopie, J. T. Arriola, D. Magde, U. F. Müller, A GTP-synthesizing ribozyme selected by metabolic coupling to an RNA polymerase ribozyme. *Sci. Adv.* **7**, eabj7487 (2021).

16. R. Krishnamurthy, Giving Rise to Life: Transition from Prebiotic Chemistry to Protobiology. *Acc. Chem. Res.* **50**, 455–459 (2017).

17. J. G. Forsythe, S.-S. Yu, I. Mamajanov, M. A. Grover, M. Krishnamurthy, F. M. Fernandez, N. V. Hud, Ester-mediated amide bond formation driven by wet-dry cycles: A possible path to polypeptides on the prebiotic Earth. *Angew Chem Int Ed Engl* **54**, 9871–9875 (2015).

18. S. L. Miller, A production of amino acids under possible primitive earth conditions. *Science* **117**, 528–9 (1953).

19. E. T. Parker, H. J. Cleaves, J. P. Dworkin, D. P. Glavin, M. Callahan, A. Aubrey, A. Lazcano, J. L. Bada, Primordial synthesis of amines and amino acids in a 1958 Miller H2S-rich spark discharge experiment. *Proc Natl Acad Sci U S A* **108**, 5526–31 (2011).

20. C. S. Foden, S. Islam, C. Fernández-García, L. Maugeri, T. D. Sheppard, M. W. Powner, Prebiotic synthesis of cysteine peptides that catalyze peptide ligation in neutral water. *Science* **370**, 865–869 (2020).

21. M. W. Powner, B. Gerland, J. D. Sutherland, Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. *Nature* **459**, 239–42 (2009).

22. W. Huang, J. P. Ferris, Synthesis of 35-40 mers of RNA oligomers from unblocked monomers. A simple approach to the RNA world. *Chem Commun (Camb)*, 1458–9 (2003).

23. L. Li, N. Prywes, C. P. Tam, D. K. O'Flaherty, V. S. Lelyveld, E. C. Izgu, A. Pal, J. W. Szostak, Enhanced Nonenzymatic RNA Copying with 2-Aminoimidazole Activated Nucleotides. *J Am Chem Soc* **139**, 1810–1813 (2017).

24. S. Islam, M. W. Powner, Prebiotic systems chemistry: Complexity overcoming clutter. *Chem* **2**, 470–501 (2017).

25. M. Frenkel-Pinter, J. W. Haynes, A. M. Mohyeldin, M. C, A. B. Sargon, A. S. Petrov, R. Krishnamurthy, N. V. Hud, L. D. Williams, L. J. Leman, Mutually stabilizing interactions between proto-peptides and RNA. *Nat Commun* **11**, 3137 (2020).

26. S. Tagami, J. Attwater, P. Holliger, Simple peptides derived from the ribosomal core potentiate RNA polymerase ribozyme function. *Nat Chem* **9**, 325–332 (2017).

27. R. R. Poudyal, C. D. Keating, P. C. Bevilacqua, Polyanion-Assisted Ribozyme Catalysis Inside Complex Coacervates. *ACS Chem Biol* **14**, 1243–1248 (2019).

28. Sweeney, Le, T., Jorge, M.Z., Schellinger, J.G., Leman, L.J., Müller, U.F., Peptide conjugates with polyaromatic hydrocarbons can benefit the activity of catalytic RNAs. *Chem. Sci.* **14**, 10318–10328 (2023).

29. J. E. Moretti, U. F. Muller, A ribozyme that triphosphorylates RNA 5'-hydroxyl groups. *Nucleic Acids Res* **42**, 4767–78 (2014).

30. A. Pressman, J. E. Moretti, G. W. Campbell, U. F. Muller, I. A. Chen, Analysis of in vitro evolution reveals the underlying distribution of catalytic activity among random sequences. *Nucleic Acids Res* **45**, 8167–8179 (2017).

31. G. F. Dolan, A. Akoopie, U. F. Muller, A Faster Triphosphorylation Ribozyme. *PLoS One* **10**, e0142559 (2015).

32. J. T. Arriola, U. F. Müller, A combinatorial method to isolate short ribozymes from complex ribozyme libraries. *Nucleic Acids Research* **48**, e116–e116 (2020).

33. K. J. Sweeney, X. Han, U. F. Müller, A ribozyme that uses lanthanides as cofactor. *Nucleic Acids Research*, gkad513 (2023).

34. M. A. Pasek, T. P. Kee, D. E. Bryant, A. A. Pavlov, J. I. Lunine, Production of potentially prebiotic condensed phosphates by phosphorus redox chemistry. *Angew Chem Int Ed Engl* **47**, 7918–20 (2008).

35. M. A. Pasek, J. P. Harnmeijer, R. Buick, M. Gull, Z. Atlas, Evidence for reactive reduced phosphorus species in the early Archean ocean. *Proc Natl Acad Sci U S A* **110**, 10089–94 (2013).

36. G. D. Cody, E. Heying, C. M. O. Alexander, L. R. Nittler, A. L. D. Kilcoyne, S. A. Sandford, R. M. Stroud, Establishing a molecular relationship between chondritic and cometary organic solids. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 19171–19176 (2011).

37. C. M. O. Alexander, M. Fogel, H. Yabuta, G. D. Cody, The origin and evolution of chondrites recorded in the elemental and isotopic compositions of their macromolecular organic matter. *Geochimica et Cosmochimica Acta* **71**, 4380–4403 (2007).

Chapter 2 : Discussion and future directions

To test whether prebiotically plausible peptides benefit ribozymes, an in-vitro selection for triphosphrylating ribozymes was done in the presence of 10 prebiotically plausible peptides. 30 of the resulting ribozymes were selected for further analysis and were shown to benefit from the peptides. The original goal of my project was to test the activity of ten of these ribozymes in the presence of each of the ten peptides to see whether the pattern of benefit is the same for all of the ribozymes. However, toward the end of collecting data for activity of the 10 ribozyme and peptide pairs, it became apparent that most of the peptides' observed benefit could be explained by their effect on the pH.

An average of two-fold beneficial effects was observed for 4 of the peptides and peptide 4 increased the activity of ribozyme s2 by 2.6 fold. This result suggests that non-coded peptides might not have had a significant beneficial effect on ribozymes. However, it is important to note that this experiment only tested ten peptides, each consisting of eight residues. Longer peptides offer a larger combination of sequences, increasing the likelihood that some may be useful for ribozymes. Although there is no prebiotically plausible pathway for synthesizing peptides significantly longer than those explored in this project, it is worth considering that the abiotic synthesis of polynucleotides is also challenging, with the prebiotic plausibility of proposed pathways still under debate. Given these challenges, it has been suggested that RNA might have been preceded by simpler molecules, just as RNA preceded coded protein synthesis (1). The RNA world hypothesis does not exclude the possibility that simpler molecules existed before RNA, and that RNA world organisms evolved from these earlier self-replicating systems. This

line of speculation could also support the idea that a simpler system predating RNA could have facilitated the formation of longer non-coded peptides.

To further investigate the potential benefits of peptides to ribozymes, a similar in-vitro selection could be conducted in the presence of a larger and longer array of peptides. Despite the questionable prebiotic plausibility of long peptides, exploring whether ribozymes can benefit from a more extensive library of longer peptides could be informative. If it is found that ribozymes do not benefit from long peptides, this could further support the idea that peptides likely did not play a significant role in the evolution of catalytic ribozymes. Conversely, if some ribozymes form useful and specific interactions with peptides, it would indicate a different conclusion. However, designing such a study would be challenging, as it would be difficult to identify which peptides are beneficial to ribozymes when using a large peptide library. In conclusion, the results of this project suggest that prebiotically plausible peptides might not have played a significant role in the evolution of catalytic RNAs, but further studies are needed to solidify this conclusion.

References:

1. Robertson, M. P.; Joyce, G. F. The Origins of the RNA World. *Cold Spring Harbor Perspectives in Biology 4* (5), (2012**)**.