Paths to constructive processes in RNA-based early life

Dissertation der Fakultät für Biologie
zur Erlangung des Doktorgrades der
Naturwissenschaften (Dr. rer. nat.)
der Ludwig-Maximilians-Universität München

Yeonwha Song Wratil

München, 2022

Diese Dissertation wurde angefertigt
unter der Leitung von Prof. Dr. Ralf Jungmann
an der Ludwig-Maximilians-Universität München

Ersgutachter:	Prof. Dr. Ralf Jungmann
Zweitgutachter:	Prof. Dr. Heinrich Jung
Tag der Abgabe:	18.08.2022
Tag der mündlichen Prüfung:	22.12.2022

TABLE OF CONTENTS

STATUTORY DECLARATION

Eidesstattliche Erklärung

Ich versichere hiermit an Eides statt, dass die vorgelegte Dissertation von mir selbständig und ohne unerlaubte Hilfe angefertigt ist.

München, 11.01.2023
Yeonwha Song Wratil

Erklärung

Hiermit erkläre ich,

- dass die Dissertation nicht ganz oder in wesentlichen Teilen einer anderen Prüfungskommission vorgelegt worden ist.
- dass ich mich anderweitig einer Doktorprüfung ohne Erfolg **nicht** unterzogen habe.
- dass ich mich mit Erfolg der Doktorprüfung im Hauptfach Biologie bei der Fakultät für Biologie der Ludwig-Maximilians-Universität München unterzogen habe.

München, 11.01.2023
Yeonwha Song Wratil

ACKNOWLEDGEMENTS

First and foremost, I wish to express deep gratitude to my direct supervisor Prof. Dr. Hannes Mutschler who, whether from within the same building or from the opposite side of the country, has provided invaluable scientific, administrative, and moral support throughout my time in his group. I am most fortunate to have had the chance to learn and grow under his supervision.

Likewise, I thank my Doktorvater Prof. Dr. Ralf Jungmann and thesis advisory member Prof. Dr. Heinrich Jung for their guidance and encouragement.

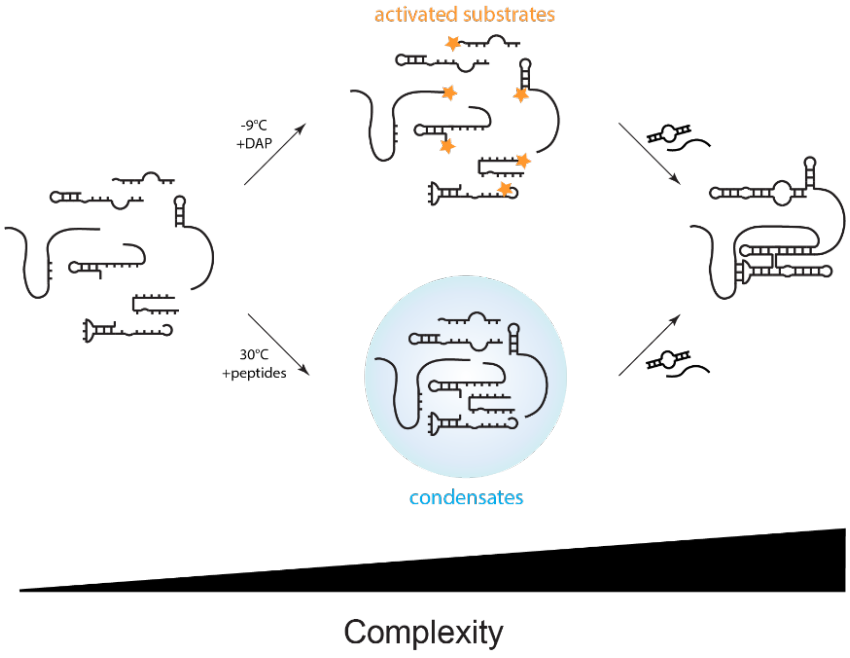
I acknowledge my graduate program CRC235 for the myriad opportunities, events, and funding. Thank you to Prof. Ram Krishnamurthy and his postdocs Huacan and Eddy for the interesting and productive scientific ventures during the better part of my doctoral research efforts, and to Paola Caselli, Michela Giuliano, and Max Winkler for the brief but eye-opening collaboration.

Instrumental to my wonderful experience in the last ~3 years were my fellow Mrunions: Laura Weise, Lena Stelzer, Alex Wagner, Kai Libicher, fellow CRClings Elia Salibi and Deni Szokoli, plus other past and present lab members. I only wish circumstances had allowed for even more game nights. Specially, many thanks to our beloved postdoc Kris through whom I got my regular dosages of cat news as well as fruitful scientific and culinary collabora-tions.

And what is the pursuit of (the origin of) life without love? The distance and time apart has not been easy, no thanks to a number of global crises, but my family remain closest to my heart. Appa and Umma, who have given everything and more for my education, growth, and outlook; Unni, who I have to admit has now advanced from my eternal tormentor to eternal tolerator (should she choose to allow it); Alex, Adeline, and Averie, the triple As of wonderful energy that fuel me. Lastly, I thank my husband Robin for his never-ending support, patience, and faith in me, and without whom I cannot imagine having been sufficiently confident or competent to complete my doctoral studies.

ABSTRACT

The work presented in this thesis aimed to bridge missing gaps in our knowledge of the emergence of primitive nucleic acid-based lifeforms. Specifically, I demonstrate two prebiotically plausible means by which longer and more complex RNA may have formed from pools of short oligomers. Firstly, I reconcile the remarkable utility of a phosphate source, diamidophosphate (DAP), in stably activating oligoribonucleotides under mild aqueous conditions amenable to the activity of the hairpin ribozyme. The reported approach provides long-lived 2',3'-cyclic phosphate oligomers whose hydrolysis products can be reactivated over weeks in a one-pot reaction. This stands in contrast to conventional laboratory methods that generate potent and short-lived activated RNA substrates to drive constructive processes such as polymerization and recombination. Secondly, while the candidacy of phase-separated droplets as a primitive protocellular environment has long been proposed, only destructive ribozyme catalysis such as cleavage of substrate RNA has thus far been shown therein. Here, I demonstrate robust and versatile constructive ribozyme activity within droplets formed from simple cationic peptides and short oligomers. In both investigations, I report the assembly of a complex >180 nt RNA polymerase ribozyme from simple <30 nt oligomers. Taken together, the described work advances our understanding of how constructive nucleic acid-based processes toward molecular evolution and – eventually – life may have occurred.



LIST OF PUBLICATIONS

Publications submitted toward this cumulative dissertation:

1. **Song, E. Y.**, Jiménez, E. I., Lin, H., Le Vay, K. K., Krishnamurthy, R., Mutschler, H. Prebiotically Plausible RNA Activation Compatible with Ribozyme-Catalyzed Ligation. ***Angew. Chem. Int. Ed.***, 60: pp. 2952-2957 (2021).

2. Le Vay, K. K.*, **Song, E. Y.***, Ghosh, B., Tang, D. T.-Y., Mutschler, H. Enhanced Ribozyme-Catalyzed Recombination and Oligonucleotide Assembly in Peptide-RNA Condensates. ***Angew. Chem. Int. Ed.***, 60: pp. 26096-26104 (2021).
   ** authors contributed equally*

Other publications resulting from this doctoral work:

3. Le Vay, K. K., Salibi, E.*, **Song, E. Y.***, Mutschler, H. Nucleic Acid Catalysis under Potential Prebiotic Conditions. ***Chem. Asian. J.*** 15: pp. 214-230 (2020).
   ** authors contributed equally*

1. INTRODUCTION

1.1 TOWARDS THE EMERGENCE OF LIFE OF EARTH

The pursuit of deciphering the origin of life has indisputably perplexed humanity for millennia. Life, defined by NASA as a "self-replicating chemical system capable of undergoing Darwinian Evolution"[1], can be observed thriving in extreme environments such as deep-sea vents[2], high-acidity geysers[3], and in many circumstances in between. The wonder of life's ubiquity on Earth aside, these observations may have somewhat clouded the investigation of narrowing down a consensus for the emergence of life on Earth.

Modern researchers in many fields such as chemistry, biology, geology, and astronomy alike have accordingly made great strides towards solving the complex puzzle; for example, our planet itself is postulated to have been a result of runaway accretion, a process that describes the accumulation of gas into protoplanetary bodies[4,5]. Investigations of fossilized microorganisms in ancient marine sedimentary rocks suggest that Earth's earliest life forms emerged at least 3.77 billion years ago in the Hadean-early Archean[6]. Further, whilst it is currently debated whether early Earth's atmosphere was oxidizing or reducing in these eras[7,8], one can generally assume access to a chemically active surface with frequent release of gaseous water and other simple abiotic molecules therein[9].

The first experimental production of organic molecules without modern metabolic intervention was the synthesis of urea by Wöhler in 1828[10]. Following this monumental report, Oparin and Haldane put forth the hypothesis that slow chemical evolution may have produced molecules of life from abiotic raw materials[11,12]. Prebiotic organic molecules such as carbohydrates and amino acids are thought to have been at least transiently accessible on early Earth, from examples including carbonaceous chondritic meteorites[13] and hydrated mineral surfaces[14,15]. In the early 1950s, Urey and Miller aimed to simulate proposed early Earth atmospheric conditions and recreate the moment that sparked simple gasses into life with electrical discharges[7]. The reported resulting mixture of amino acids ignited a wealth of investigation and schools of thought – whether genetics, compartmentalization, or metabolism was the first lifelike phenomenon to emerge. Whatever the route, an unlikely yet evidently sufficient combination of physicochemical parameters is assumed to have led eventually to what we now recognize as the universal molecules of life: DNA, RNA, and proteins. A paradigm shift away from attempting to unearth the

exact history of life's origin to exploring the possible scenarios for life's origin was perhaps helpful in incorporating concepts from diverse fields to continually develop our understanding of abiogenesis.

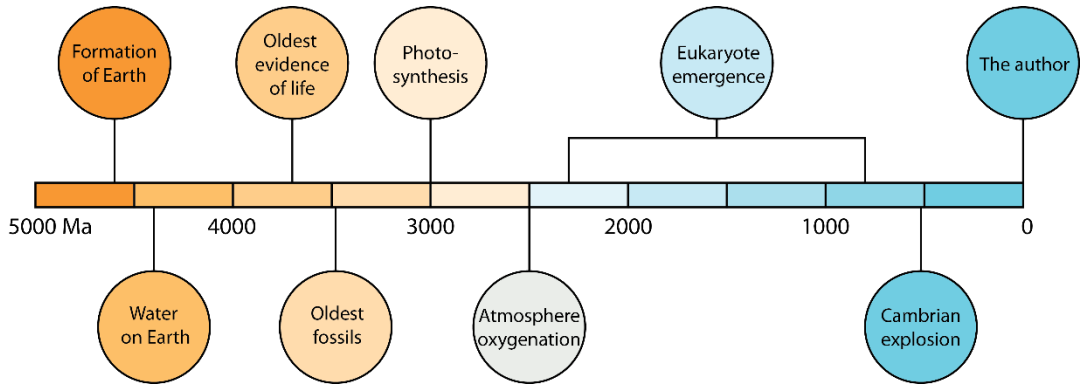


Figure 1. A humbling and simplified timeline of life on Earth with highlighted significant events. The relevant periods of interest for this thesis is between 3700 and 4400 Ma. Ma = Million years ago.

1.2 THE RNA WORLD

Modern cell-based life generally adheres to the central dogma of biology, wherein DNA acts as the informational carrier polymer that encodes genetic information transcribed into RNA, which is in turn translated into protein that carries out an array of functions. The complexities of the biochemistries involved in maintaining and driving these processes suggest a simpler ancestral system. In the hypothesis of the "RNA World", first posited in the 1960s by Alexander Rich[16–20], the corresponding primitive ancestral biochemistries instead find RNA in a central functional and genetic role before the advent of DNA and proteins, thereby providing an elegant solution to the long-standing question of whether proteins or DNA emerged first. Some 60 years after its conception, it is still one of the most widely accepted hypotheses for the emergence of life on Earth[21].

The discoveries of so-called "molecular fossils"[22,23] lend credence to many for this model of early biology; the significant role of RNA and RNA-like cofactors in modern cells that utilize nucleotides suggest a previous protein-free era[24]. Similarly, the ability of RNA to fold into protein-like structures point to its initial role in catalysis. A pivotal moment for the concept of the RNA World was the independent discoveries of the first ribozymes, or catalytic RNA, by Altman and Cech[25,26]. The significance of this finding

marked the beginning of a new field of research which led to another monumental "smoking gun" realization: that the catalytic core of the protein synthesis-driving ribosome present in all kingdoms of life is, in fact, a ribozyme[27–29]. Indeed, while valid opposing views contest the idea of the RNA World[14,30,31], "the primitive ribosome could have been made entirely of RNA"[18].

Although the notion of an RNA World may harbor various meanings and assumptions, it is generally understood that: replication of RNA ensured genetic heritability at some point in primordial biology; Watson-Crick base pairing drove replication; and catalytic proteins encoded by nucleic acids were not involved[32]. Further, central to this model is the emergence of a self-replicating RNA capable of open-ended evolution[33] whose emergence is depicted in Figure 2.

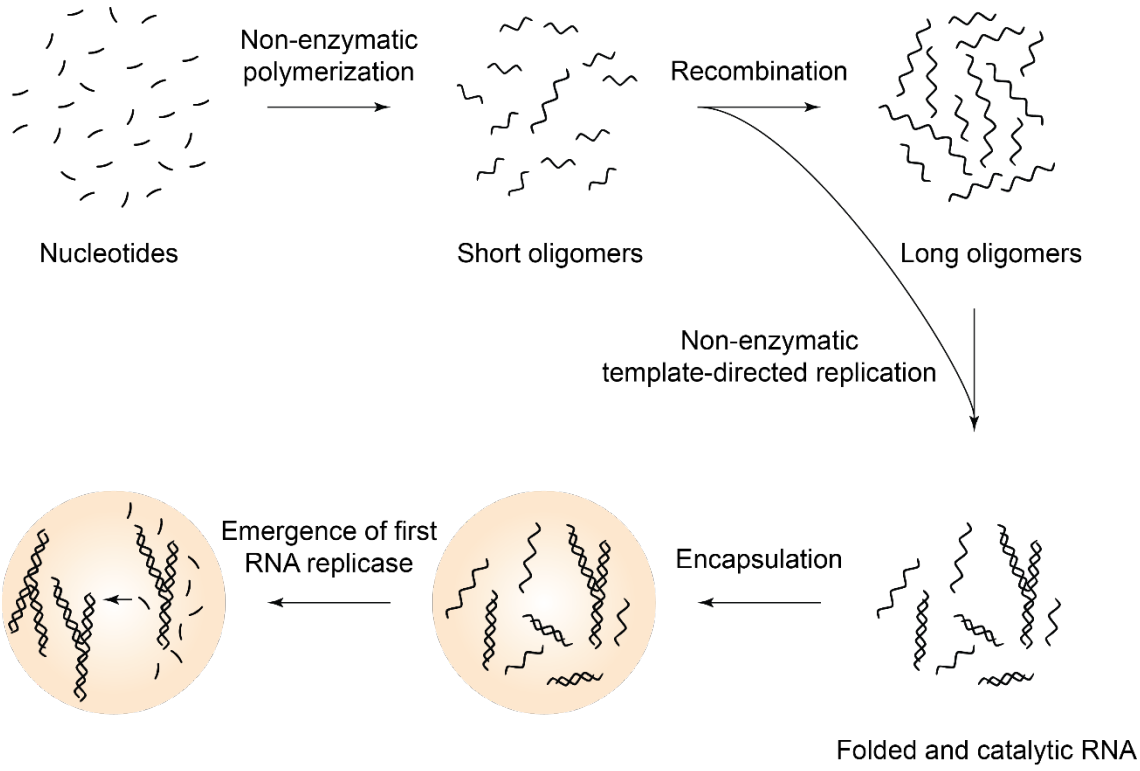


Figure 2. The classical RNA world hypothesis. Nucleotide and oligomer pools emerge randomly from prebiotic synthesis and polymerization. Non-enzymatic template-directed replication and recombination therein result in the emergence of functional RNA, assuming various three-dimensional structures. With increasing complexity, the first RNA replicase and encapsulation realize protocells capable of Darwinian evolution. Contrary to the simplified depiction above, these processes likely did not occur sequentially but rather at least partly in parallel. Adapted from Le Vay and Mutschler[34].

The tractability, rather than the historical accuracy, of the RNA World as a model system for studying prebiotic processes, perhaps provides common ground across varying views in the field of origin of life to further our understanding in the long-standing questions. For example, it is unclear exactly how the first relatively short and simple ribozymes may have been sufficiently fueled and compartmentalized to increase complexity towards the first self-replicator. This thesis experimentally explores how certain aspects of such processes may have occurred.

1.3 CATALYTIC RNA AS A MODEL TO STUDY EARLY LIFE

The significant discovery of ribozymes in ribosomes and in the thermophilic protozoon *Tetrahymena*[25] had deep implications for the scientific community and marked the beginning of many other ribozyme discoveries. A number of naturally occurring ribozyme motifs have since been extensively characterized, such as the hammerhead ribozyme in viroids[35] and hairpin[36]. Importantly, many such sequences catalyze ligation reactions. Ligation and polymerization are attractive potential reactions for the first replicating ribozymes because they inherently increase length and complexity of the nascent nucleic acid pool.

While naturally-occurring ribozyme activity is limited to phosphodiester transfer and hydrolysis reactions of RNA and DNA, the inherent catalytic potential of RNA far exceeds these limitations[37]. Significant efforts have been invested in *in vitro* evolution experiments to unearth novel ribozyme activities from populations of random RNA sequences. Therein, a selection pressure of e.g. catalyzing a ligation reaction is applied, then the sufficiently active candidate sequences are purified and amplified using RT-PCR. The process is repeated over multiple rounds to isolate active sequences specific for the catalytic activity of interest. These SELEX (systematic evolution of ligands by exponential enrichment) approaches have also fruitfully led to highly adapted sequences, or aptamers, for applications such as biosensing and therapeutics[38], as well as polymerases and ligases.

Interestingly, *in vitro* selected ligase ribozymes display lower threshold of sequence length for activity compared to their polymerase counterparts[39]. Short sequences have the intrinsic advantage of being copied more quickly and these observations raise curiosities around the role of simple ligating RNA sequences in the emergence of life. In this

work, I build on the well-characterized hairpin ribozyme as a model system to simulate processes that may have provided increased biochemical complexity on prebiotic Earth.

1.3.1 THE HAIRPIN RIBOZYME

First discovered in 1986 in the negative strand of the tobacco ringspot virus satellite RNA[40], the same genome that harbors a hammerhead ribozyme in its positive strand, the hairpin ribozyme motif belongs to the class of small, self-cleaving catalytic RNAs. It is extensively characterized and studied[36] in part due to its relative simplicity and structure consisting of 50 to 150 nt. In nature, hairpin ribozymes are involved in processing RNA replication intermediates[41]. In the laboratory, their tractability and robustness have led to extensive use in, and insights from, studying the nature of catalytic RNA[36]. Indeed, the hairpin motif may have provided sufficient polyribonucleotide length for the genetic information and catalytic capacities of an ancestral replicator[42]. Moreover, naturally existing ribozymes such as the hairpin ribozyme are thought to have retained some catalytic features of their predecessors and to be relics of the RNA world[43,44].

The hairpin ribozyme catalyzes the reversible sequence-specific cleavage of its RNA substrate phosphodiester backbone, which produces products containing 2',3'-cyclic phosphate and 5'-hydroxyl moieties (Figure 3a). Currently known observations suggest precise substrate orientation, preferential transition state binding, electrostatic catalysis, as well as general acid base catalysis for its mechanistic catalysis strategies[45]. Its structure comprises two domains, each consisting of two base-paired helices H1 to H4 separated by internal loops A and B (Figure 3b). The first domain includes the 2',3'-cyclic phosphate species residing within loop A and the larger second domain contains the primary catalytic residues – the two domains interact to form a complex for catalytic function[46,47]. Activity is unaffected by increasing H1 and H4 lengths up to 25 base pairs[36], whereas increasing base pairs from four to five in H2 inhibits catalysis[48]. Active complexes can be formed both from single and multiple separate RNAs. Metal ions are necessary to drive the docking transition to form the active catalytic complex, but unlike many ribozymes and protein enzymes, the hairpin ribozyme does not employ metal ions as catalytic cofactors, instead solely depending on internal functional chemical groups for its catalysis[49–51].

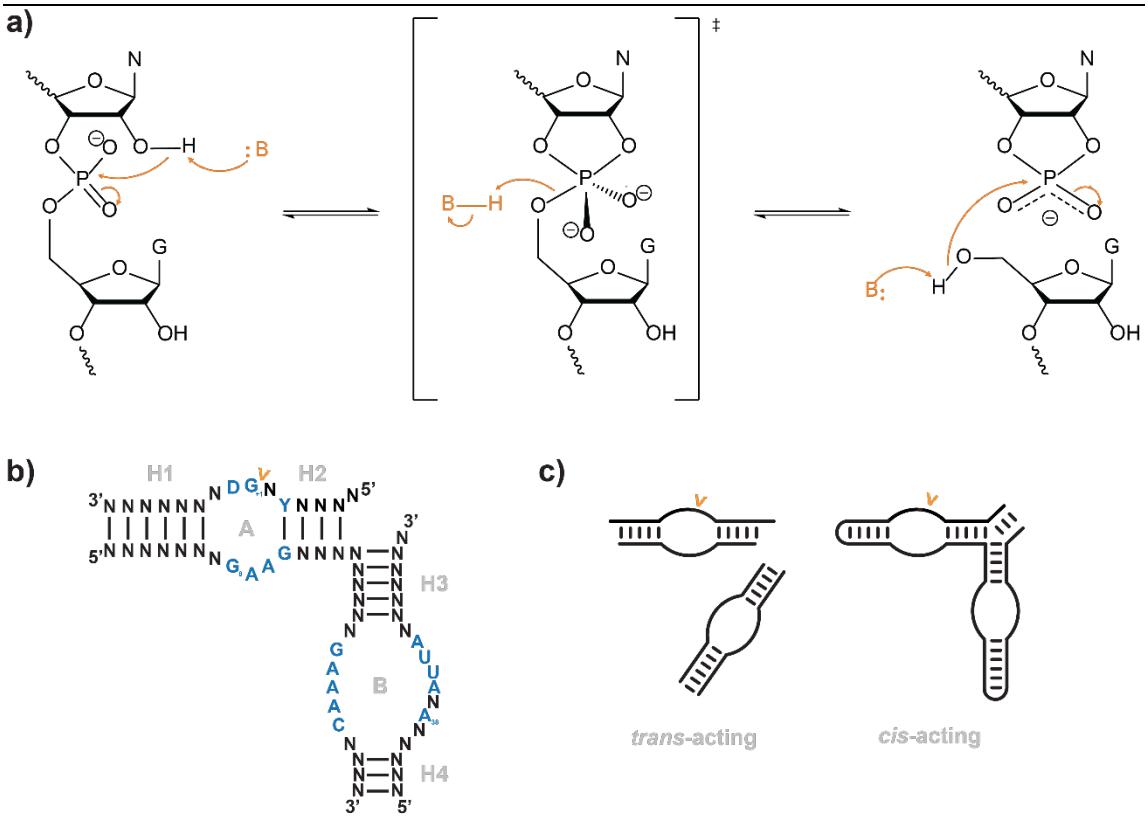


Figure 3. Sequence and mechanism of the hairpin ribozyme. a) Consensus secondary sequence of the minimal hairpin ribozyme motif. Nucleotides in blue are essential for catalytic function[52]. Site of cleavage/ligation is shown in orange (arrow). N = nucleotide identity inconsequential for activity if base pairing remains intact. D = not C (U or G or A). Y = pyrimidine (C or U). H1 to H4 represent base-paired helices. b) *trans*-acting and *cis*-acting engineered variants of hairpin ribozyme used in this thesis. c) The mechanism of hairpin ribozyme-catalyzed cleavage and ligation reactions. Cleavage begins with nucleophilic attack by the 2' oxygen of N_{+2} on the adjacent phosphorous via S_N2, resulting in a transition state stabilized by G_8 and A_{38}. Breaking the bond between the 5' oxygen of G_{+1} and the phosphate produces 2',3'-cyclic phosphate and 5' hydroxyl product termini, the latter of which acts as the nucleophile in the reverse ligation reaction[36].

The equilibrium between the cleavage and the reverse ligation reactions catalyzed by the hairpin ribozyme is largely affected by environmental conditions[36,53]. k_{cat} and K_M values ranging from 0.2 to 0.5 min^{-1} and 30 nM, respectively, have been reported for minimal hairpin ribozymes under similar standard conditions[54,55], representing approximately 10^6-fold increase relative to the uncatalyzed reaction. Increased salt and reduced temperatures were found to encourage equilibrium toward ligation[53]; indeed, in frozen solutions, the hairpin ribozyme displays enhanced ligation activity while its cleavage activity is inhibited[56]. This effect is enhanced in freeze-thaw cycles through which entrop-

ically disfavoured assembly of the hairpin ribozyme complex is overcome by steep temperature and concentration gradients[57]. Further, ligation in *cis* is most efficient at -10 °C, highlighting the important role of concentration by dehydration in driving constructive reactions[58]. Such findings lend further credence to the notion of primordial biology having found favourable conditions in at least transiently cold environments[59–61].

In this thesis, I employ the robust and tractable hairpin ribozyme as a model system to explore prebiotically plausible processes that allow for increased nucleic acid sequence complexity and length. Both *cis*- and *trans*-acting variants were employed; separate RNA strands were designed either with an intact ribozyme strand and separate substrate strand, or with five shorter separate strands that assembled into the active complex (Figure 3c). Its compact size, previous extensive characterization, and ligative catalytic activity filling a pivotal role towards larger RNA make it an attractive tool for this purpose.

1.4 CHALLENGES FOR THE EMERGENCE OF THE FIRST RNA SELF-REPLICATOR

The popular conjecture that Earth's lifeforms emerged from an RNA-dominant environment is not without objection[34]. Despite a considerable amount of evidence that suggests RNA was a key prebiotic molecule capable of self-replication and catalysis, several hurdles still plague our understanding of the advent of increasing complexity.[62] The relative complexities of RNA molecules are regarded by some as too extreme and "energy-hungry" to arise *de novo* from prebiotic chemistry[14]. Moreover, it is argued that functional RNA comprises only a small percentage of all RNA and requires prebiotically implausible lengths, limiting the suite of chemical reactions catalyzed by RNA[30]. Finally, RNA replication and evolution may have required higher concentrations of building blocks than is currently thought to have been prebiotically viable[63,64]. Naturally, the RNA world concept does not yet fully describe the origin of life, but perhaps more importantly, it has helped streamline experimental advances[65].

Further, central to the RNA world is the emergence and survival of a self-replicating RNA that functions as an RNA-dependent RNA polymerase, using itself to produce complementary RNAs which are then used to produce more copies of itself[20,21,66]. This view assumes the availability of activated nucleotides to fuel the propagative reactions. However, the efficiency of template RNA replication is limited by substrate degradation, and currently employed methods to re-activate hydrolyzed substrates lead to undesirable side reactions that compromise both templates[62].

The combination of physicochemical and geological parameters that shaped the nurturing environment from which the first RNA replicator would have arisen is also hotly debated[67]. First, the emergence of the self-replicator necessitates sufficient concentrations of substrate RNA. Additionally, because the RNA backbone degrades rapidly through transesterification under warm aqueous conditions in the presence of divalent metal ions[68,69], a prebiotic environment requires a mechanism by which this is prevented. Moreover, chemical evolution of increased RNA length and sequence complexity dictates more efficient RNA synthesis compared to degradation. Such concerns may be addressed by a microenvironment in which the nucleic acids can be sequestered. In this thesis, the hurdles of activation and compartmentalization for the self-replicator are addressed.

1.4.1 FUELING LIFE-FUELING REACTIONS

Modern cells rely heavily on nucleoside triphosphates to drive energy-requiring ligative processes. However, these high-energy molecules are generally unreactive without protein enzyme catalysis[70]. Moreover, despite the necessity of constructive ligation reactions for the emergence of the self-replicator RNA from pools of oligomer RNAs, such condensation reactions are not favoured in dilute aqueous environments[66]. In order to drive these reactions, a reliable source of substrate activation is paramount. Some have claimed that proton and thermal gradients between the hot alkaline deep-sea hydrothermal vent current and the surrounding cooler acidic waters may have provided the much-needed first sources of energy to drive the origin of life.[71–73] However, the instability of RNA under alkaline conditions seem to contradict the plausibility of such locations for RNA world emergence[74].

Approaches to simulate non-enzymatic activation of ribonucleotides in the laboratory towards bridging the gap between prebiotic chemistry and RNA-based lifelike forms have been explored[75,76]. Generally, a separate chemical reaction is performed to produce anhydronucleotides or monomers with a leaving group, which are then subsequently isolated and used in further experiments. These studies have shown the formation of RNA strands on mineral surfaces[75] and incorporation of all four nucleobases in heterogeneous media as well as longer polymer formation in eutectic phases[56]. The use of pre-activated nucleotides such as azolides also enable non-enzymatic template replication via primer extension[77–79]. However, the prebiotic viability of such activation chemistries is unclear. It may therefore be beneficial to consider activation chemistries that are not only prebiotically plausible but come from prebiotically plausible means of building block syntheses.

2',3'-cyclic phosphate ribonucleoside (>P) activated species are considered to be a feasible prebiotic basis for abiotic RNA polymerization[80]. >Ps themselves can arise from prebiotically plausible routes such as *de novo* pyrimidine nucleotide synthesis[81], RNA degradation[82], nucleoside phosphorylation[83], from recyclization of 2'/3' monophosphate nucleosides[84] (Figure 4a), as well as abundantly occurring in modern cells[85]. Due to the thermodynamically disfavoured polymerization of >Ps in aqueous conditions[86], concentrating environments such as evaporation or eutectic ice have been employed to drive RNA oligomer formation from >Ps[87,88]. Indeed, RNA extension under frozen aqueous conditions have been demonstrated using a variant of the hairpin ribozyme and >Ps[89], highlighting the likeliness of the utility of >Ps in prebiotic RNA oligomer formation from pools of nucleotides.

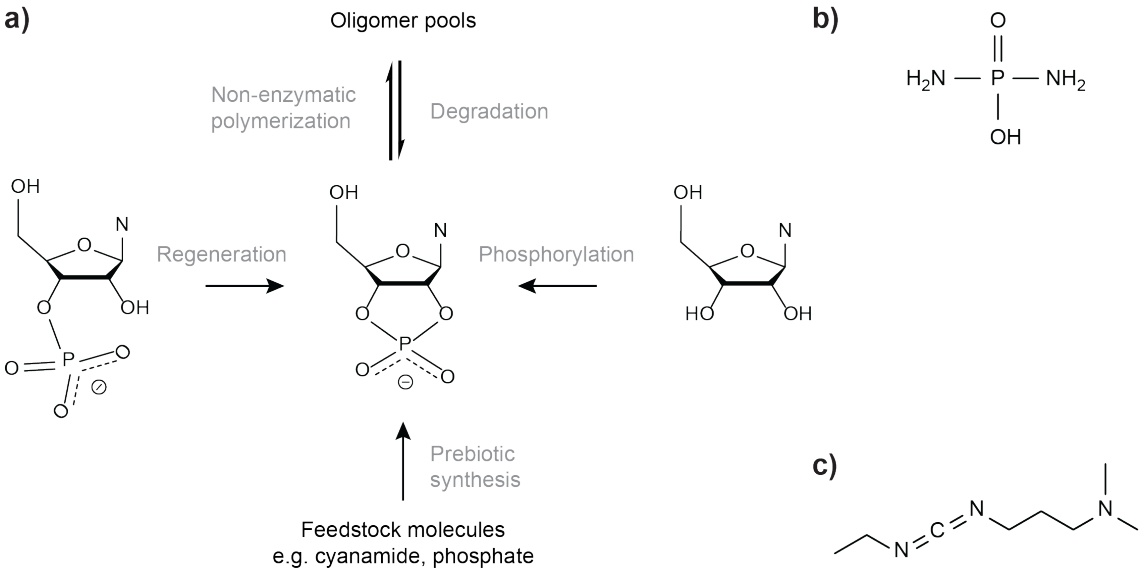


Figure 4. a) Prebiotically plausible routes towards 2',3'-cyclic phosphate nucleoside formation include oligoribonucleotide degradation by transesterification[82], nucleoside phosphorylation[83], *de novo* nucleotide synthesis[81], and regeneration from hydrolyzed 2'/3' monophosphate nucleosides[90]. Adapted from Mutschler and Holliger[89]. b,c) Chemical structures of diamidophosphate (DAP) and N-ethyl-N′-(3-dimethylaminopropyl)carbodiimide (EDC), respectively.

A class of widely-used >P-generating reagents are carbodiimides such as N-ethyl-N´-(3-dimethylaminopropyl)carbodiimide (EDC, Figure 4c). At mildly acidic aqueous conditions, the diimide moiety of EDC can robustly generate >Ps within 1 h in relatively simple one-pot reactions[91], in contrast to other means of RNA activation requiring organic solvents and purification. Although versatile and robust, their prebiotic plausibility is unclear

due to their instability in water[92]; indeed, EDC readily and irreversibly hydrolyzes to a urea byproduct due to its high reactivity. Further, the high reactivity can also nonspecifically react with other chemical groups of nucleic acids, thereby compromising their structural integrity[93]. Therefore, water-stable cyclic phosphate-generating reagents are needed in simulated prebiotic chemistry experiments. Further, those amenable to conditions that also host catalytic activity of >P-accepting prebiotically plausible RNA such as the hairpin ribozyme would strengthen our understanding of how the first replicators could have sustained constructive processes.

Diamidophosphate (DAP, Figure 4b) is one such promising water-stable and prebiotically plausible molecule[94]. It can be sourced from prebiotically plausible sources such as schreibersite minerals in aqueous ammonia[95], highlighting the possibility of a reliable feedstock of prebiotic phosphorylation reagents[96,97]. Indeed, DAP has been shown to phosphorylate biomolecules such as amino acids, lipid precursors, nucleic acids and their building blocks, as well as to initiate polymerization of DNA[95,98]. Importantly, DAP-activated RNAs produce >Ps in solution within days to weeks[98] – a rapid process in the geological time scale.

2',3'-cyclic phosphate ribonucleosides and their generation from monophosphate counterparts with DAP may comprise a brick in the incomplete bridge from prebiotic chemistry to the RNA World. However, the utility of this energy source in facilitating constructive ligative processes under dilute and mild aqueous conditions amenable to the activity of plausible prebiotic ribozymes remains unclear. As part of this thesis, I explore a stable RNA activation approach compatible with the activity of a simple RNA-ligating ribozyme model system.

1.4.2 A PROTOCELL TO CALL HOME

Compartmentalization in modern biology is crucial for proper cellular function. Similarly, in primordial biology, appropriate sequestration of relevant molecules is considered to have been vital in the evolution of lifelike systems. Especially for the primordial RNA self-replicase, the lack of compartmentalization would simply result in the products diffusing away in free solution, failing to benefit its self-replication and subsequent molecular evolution[99]. Moreover, in free solution, unfolded parasitic sequences that provide no catalytic activity but harbor better templating properties outcompete self-replicating

ribozyme activities that depend on specific folded structures[100,101]. The eventual population crash can be mitigated by even transient compartmentalization[100,101], highlighting the crucial nature of partitioning primitive lifelike processes.

The modern lipid bilayer cellular membrane, virtually impermeable to ions and molecules such as nucleotides, is composed of phospholipids and sterols, among other complex structures and can be observed in all levels of biological organization. In contrast to the sophisticated machinery observed in modern cells, protocells – primitive forms of compartmentalization – likely comprised simpler and more permeable borders, while still complying with physicochemical requirements necessary to host RNA chemistry. For example, the instability of fatty acid-based membranes at Mg^{2+} concentrations required for most ribozyme activity and nonenzymatic copying of RNA templates seems to suggest an inherent incompatibility, although the presence of citrate seems to mitigate this to an extent[102]. Accordingly, a variety of alternative prebiotically plausible forms of compartmentalization have been proposed.

The fundamental necessities to prevent accumulation of parasitic sequences and free diffusion of polymers in achieving functional molecular evolution can be met by several additional means. Porous rocks observed in deep-sea vents, a proposed microenvironment for the emergence of life[103], are thought to provide sufficient concentration gradients to concentrate oligonucleotides and assemble longer products over shorter ones[104–106]. Secondly, montmorillonite clay is known to facilitate nucleic acid polymerization; the strong binding of RNA to the mineral surface may have provided the sequestration of relevant molecules[75,107]. Third, while the prebiotic plausibility of water-in-oil emulsion droplets harbouring sufficient molecular evolution processes is unclear, their potential for cyclical growth and division provides a valuable model for protocell studies[108]. A fourth alternative to fatty acid-based vesicular compartmentalization of nucleic acid propagation is nucleic acids themselves. As exemplified by the growing field of nucleic acid origami structure assembly[109–111], the compartment may be encoded by the replicating nucleic acid strands within. Lastly, liquid-liquid phase separation (LLPS), a common phenomenon in aqueous solutions of macromolecules, presents an interesting mode of prebiotic compartmentalization. In a class of LLPS called complex coacervation, strong electrostatic interactions between polymers harboring opposite charges result in dynamic and concentrated liquid droplet condensates[112] (Figure 5).

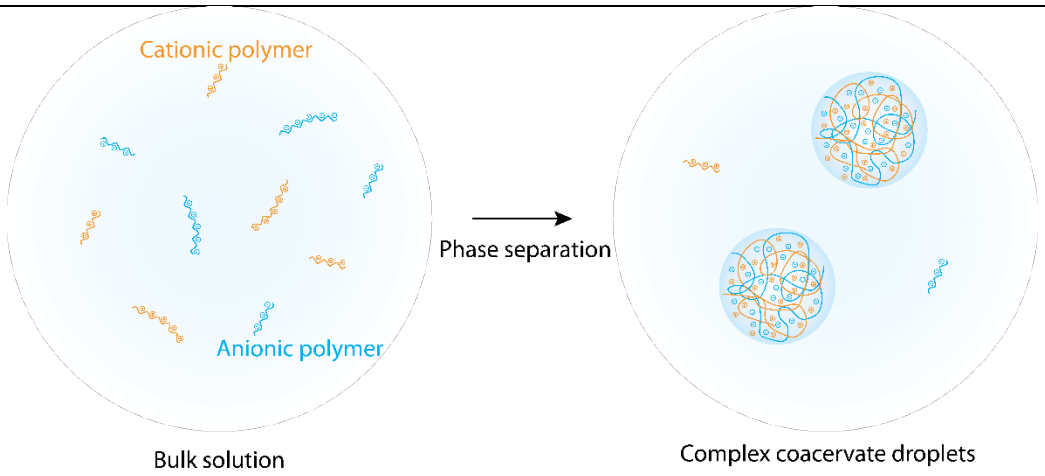


Figure 5. Schematic diagram of complex coacervation driven by electrostatic interactions between oppositely charged polymers. Entropic release of the polymers from water and counterions generate membraneless droplets via phase separation[113].

These dense, polymer-rich droplets exhibit remarkable fluidity and are subject to growth and fusion; indeed, their formation and disassembly are affected by many factors such as salt concentrations[114,115], pH[116], and temperature[117]. Moreover, due to their barrier-free exposure to the surroundings, uptake of charged species such as Mg^{2+} ions is unhindered relative to fatty acid-based protocellular environments[118,119].

Complex coacervation offers several advantages as a candidate for early protocell forms[11,12]. First and most obviously, this phenomenon offers continuity to modern cells that host a plethora of phase-separated membraneless organelles such as stress granules and signaling complexes[120,121]. Moreover, coacervate droplets can comprise a variety of polymers such as simple prebiotically plausible charged peptides. Indeed, acidic peptides facilitate RNA polymerization in nature by positioning Mg^{2+} ions for catalysis and may have been available on early earth by condensation on mineral surfaces[75,122,123]. Further, prebiotic formation of individual nucleosides may have been facilitated by the presence of simple peptides[124]. However, it is currently unclear whether phase-separated droplet systems can host constructive, rather than destructive, nucleic acid catalysis.

While ribozyme activity has been demonstrated within complex coacervate droplets, such studies require fine-tuning of environmental conditions to prevent ribozyme misfolding and have only shown cleavage reactions[125,126]. Therefore, experimentally illustrating constructive ribozyme activity such as ligation and recombination within phase-

separated systems would provide a valuable bridge between potential protocellular compartmentalization and molecular evolution. As part of this thesis, I leveraged dynamic but sufficiently isolating liquid-liquid phase separated condensate droplets to drive the ligation of long and complex RNA under a range of mild aqueous conditions.

1.5 AIMS OF THE THESIS

As depicted, the core aim of this thesis was to investigate whether constructive RNA-based processes could be hosted in prebiotically plausible environments. First, I addressed the open question of how primitive RNA-based systems could have sourced a reliable feedstock of activated nucleotide substrates to drive ligation. To this end, using a prebiotically plausible ribozyme system, I investigated the potential role of the water-stable phosphorylation reagent diamidophosphate in stably activating RNA substrates in ribozyme-catalyzed RNA ligation. Secondly, I asked whether complex coacervate droplets could host similar propagative reactions, because only nucleic acid cleavage processes have been demonstrated therein thus far. I investigated the role of droplets formed from poly-*L*-lysine and a modified hairpin ribozyme formed from short prebiotically plausible oligomers in driving the reaction equilibrium towards ligation to produce longer and more complex products.

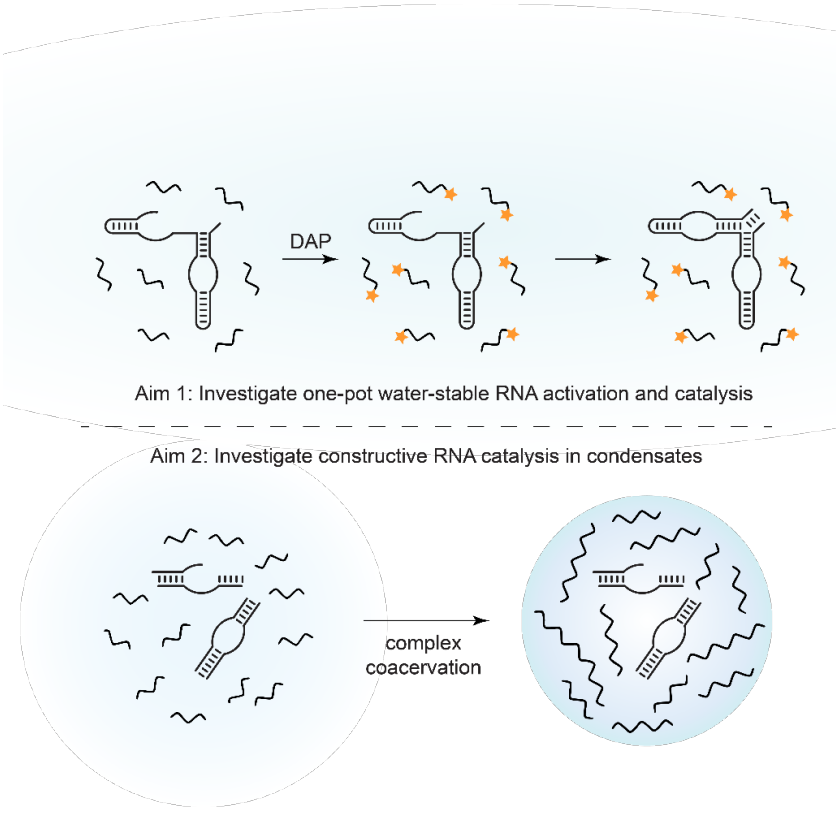


Figure 6. Aims of this thesis.

2. PUBLISHED WORKS

2.1 PUBLICATION I

Prebiotically Plausible RNA Activation Compatible with Ribozyme-Catalyzed Ligation

AUTHORS: **Emilie Yeonwha Song**, Dr. Eddy Ivanhoe Jiménez[+], Dr. Huacan Lin[+], Dr. Kristian Le Vay, Prof. Dr. Ramanarayanan Krishnamurthy[*], Prof. Dr. Hannes Mutschler[*]

[+] *authors contributed equally.* [*] *corresponding authors.*

JOURNAL, YEAR, VOLUME, AND PAGE NO.:

Angew. Chem. Int. Ed., **2021**, *60*, 2952-2957

2.1.1 SUMMARY

A key step in primordial biology to form increasingly larger and complex molecules is the ligation of RNA substrates by catalytic RNA, or ribozymes. However, it is unclear how these prebiotic systems generated and sustained pools of activated building blocks required for such constructive processes. Typical laboratory routes towards activating RNA to form activated phosphates, such as with carbodiimides, require separate synthetic steps and are incompatible with ligase ribozyme catalysis.

In this publication, we describe the first example of prebiotically plausible *in situ* RNA activation in conjunction with ribozyme-catalyzed RNA ligation. We show that diamidophosphate (DAP), a plausible prebiotic phosphate activating reagent, is able to activate the 3'-phosphate of RNA under dehydrating conditions to form the corresponding 2',3'-cyclic phosphate, which is then available for robust ligation by a natural ribozyme. The long-term stability of activation reagents weighs more than their reactivity for advancing molecular evolution. While the highly reactive carbodiimides are irreversibly hydrolyzed, DAP is stable under these conditions and is persistently available for ligand activation. Additionally, the laborious process of generating and purifying short-lived pre-activated RNA is unnecessary with DAP due to its ability to reactivate hydrolyzed phosphates back to cyclic phosphates.

Moreover, we demonstrate DAP activation-driven generation of complex RNA polymerase enzymes from multiple short oligonucleotides. These results provide a much-needed plausible scenario for the generation of higher-energy substrates required to fuel ribozyme-catalyzed RNA ligation in the absence of modern-day metabolism.

Check for updates

How to cite: *Angew. Chem. Int. Ed.* **2021**, *60*, 2952–2957
International Edition: doi.org/10.1002/anie.202010918
German Edition: doi.org/10.1002/ange.202010918

Prebiotic Chemistry

Prebiotically Plausible RNA Activation Compatible with Ribozyme-Catalyzed Ligation

Emilie Yeonwha Song, Eddy Ivanhoe Jiménez[+], Huacan Lin[+], Kristian Le Vay, Ramanarayanan Krishnamurthy, and Hannes Mutschler**

Abstract: *RNA-catalyzed RNA ligation is widely believed to be a key reaction for primordial biology. However, since typical chemical routes towards activating RNA substrates are incompatible with ribozyme catalysis, it remains unclear how prebiotic systems generated and sustained pools of activated building blocks needed to form increasingly larger and complex RNA. Herein, we demonstrate in situ activation of RNA substrates under reaction conditions amenable to catalysis by the hairpin ribozyme. We found that diamidophosphate (DAP) and imidazole drive the formation of 2′,3′-cyclic phosphate RNA mono- and oligonucleotides from monophosphorylated precursors in frozen water-ice. This long-lived activation enables iterative enzymatic assembly of long RNAs. Our results provide a plausible scenario for the generation of higher-energy substrates required to fuel ribozyme-catalyzed RNA synthesis in the absence of a highly evolved metabolism.*

Modern cells have evolved elaborate metabolic networks in mild aqueous conditions to ensure their self-preservation and to sustain their pool of activated building blocks. In contrast, it remains unclear how analogous (re-)activation of building blocks in primitive precursor systems could have been possible without access to sophisticated protein enzymes. Given a likely central role of RNA catalysts (ribozymes) during the Origin of Life,[1] robust prebiotic processes must have provided pools of activated mono- and oligonucleotides for activities such as RNA self-replication.

In the laboratory, RNA synthesis by ribozymes is achieved through the ligation of pre-activated RNA substrates. For example, some artificial ribozymes obtained by in vitro selection utilize 5′-triphosphate activation chemistry for phosphodiester formation.[2–4] Their substrates are typically obtained from in vitro transcription reactions involving RNA polymerase proteins and retain the 5′-triphosphate of the first nucleotide. In vitro selection experiments also yielded ribozymes that can triphosphorylate specific RNA substrates using trimetaphosphate.[5] However, credible non-enzymatic pathways for high-yielding 5′-triphosphate activation of RNA are currently missing.[6,7] Other ligase ribozymes utilize short-lived 5′-phosphoramidate-activated RNA. These substrates are obtained through dedicated pre-activation of 5′-phosphorylated RNA with reagents such as imidazole derivatives,[8–10] and either labile and nonspecific carbodiimides or prebiotically implausible 2,2-dipyridyl disulfide and triphenyl phosphine.[11] Thus, robust chemical pathways that yield activated RNA under conditions that also enable ribozyme activity remain elusive.

Diamidophosphate (DAP) was recently identified as a promising water-stable and prebiotically plausible reagent to phosphorylate biological building blocks such as nucleic acids, amino acids, and lipid precursors.[12,13] Of particular interest for the activation of RNA is the ability of DAP to produce nucleoside 2′,3′-cyclic monophosphates (>P), which can subsequently polymerize into short RNA oligonucleotides without the need for additional activating reagents. Enthalpically, >P-dependent formation of RNA phosphodiester bonds is favored due to the small amount of energy stored in the strained 2′,3′-cyclic phosphate, but is disfavored entropically.[14,15] However, stable substrate binding and/or low temperatures can compensate for the entropic costs and lead to a strong shift in reaction equilibrium towards ligation.[16,17] Consequently, >P-activated mono- and oligoribonucleotides may act as potent building blocks for primitive ribozyme-catalyzed[17–20] and even spontaneous RNA ligation reactions that increase the diversity and length distribution of RNA polymer chains.[21,22] The mild conditions required for DAP-dependent phosphorylation provide an attractive approach to activate free 3′-termini of RNA for ligation reactions in the context of primitive metabolism. Under solution conditions and high millimolar concentrations of DAP, nucleoside, and metal chloride, quantitative >P formation is observed within days to weeks.[12]

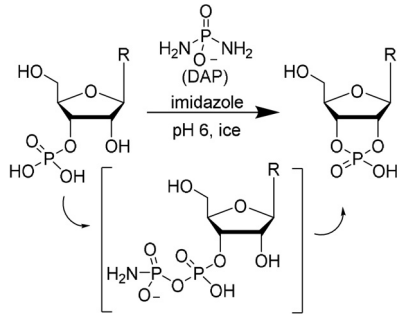
However, high levels of metal ions such as Mg^{2+}, which are required for efficient >P formation under aqueous conditions, are incompatible with half-lives of ribozymes in water.[23]

[*] E. Y. Song, Dr. K. Le Vay, Prof. Dr. H. Mutschler
Max Planck Institute of Biochemistry
Am Klopferspitz 18, 82152 Martinsried (Germany)
E-mail: mutschler@biochem.mpg.de
Dr. E. I. Jiménez,[+] Dr. H. Lin,[+] Prof. Dr. R. Krishnamurthy
Department of Chemistry, The Scripps Research Institute
10550 North Torrey Pines Road, La Jolla, CA 92037 (USA)
E-mail: rkrishna@scripps.edu
Prof. Dr. H. Mutschler
Technical University Dortmund
Otto-Hahn-Strasse 4a, 44227 Dortmund (Germany)
[+] These authors contributed equally to this work.
Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under: https://doi.org/10.1002/anie.202010918.

© 2020 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Here, we demonstrate RNA activation by DAP under conditions that are amenable to the catalytic activity of >P-dependent ribozymes. In particular, we found that the concentrating and preserving environment of water-ice enables efficient formation of >P-activated RNA. The activated RNA can be used in situ for RNA ligation by derivatives of the naturally occurring hairpin ribozyme (HPz), which serves as a versatile model system for prebiotic RNA ligation.[17–19]

To explore the potential of water-ice in combination with DAP to enable >P RNA formation, we incubated individual canonical ribonucleoside 3′-monophosphates (3′-NMPs) with DAP and imidazole (or its derivatives 2-amino- and 2-methylimidazole) at −20 °C and monitored the reactions using ion-exchange liquid chromatography (Figures S25–S72). In solution, DAP reacts with phosphates to form the corresponding amidopyrophosphate. This intermediate forms the 2′,3′-cyclophosphate in the presence of a 2′-OH group, with the amidophosphate as the leaving group (Scheme 1).[12] Intriguingly, we observed moderate to good yields for the U, C, A, and G >Ps after 28 days (Figures 1a, S1–S12) confirmed by ^{31}P NMR (Figures S73,74). Among the three activators tested in this study, imidazole-containing reactions were faster (Figures S1–S12) and yielded the highest amounts of >Ps (Figures 1a, S13–S16) under the four pH conditions tested. We also note remarkable differences in the efficiency of >P formation depending on the nucleobase (Figure 1b). After 7 days at pH 6 in the presence of imidazole, approximately 50 % of both pyrimidine nucleotides 3′-UMP and 3′-CMP were converted into their respective >Ps. In contrast, conversion of the purine nucleotides under the same conditions was much slower with circa 33 % for 3′-AMP and only <3 % for 3′-GMP. After 28 days, the conversion yield for 3′-UMP, 3′-CMP, and 3′-AMP reached 72–88 % while it approached 30 % for 3′-GMP. See below for further mechanistic explanations for the observed influence of the nucleobase moieties.



Scheme 1. Diamidophosphate-mediated formation of 2′,3′-cyclic phosphate ribonucleotides under conditions compatible with ribozyme catalysis.

Having shown that DAP can be used as an efficient activation agent for N>P formation in frozen water-ice systems, we sought to explore whether DAP can also activate oligonucleotides to fuel ribozyme-catalyzed ligation reactions. Previous studies have shown that HPzs are capable of efficient >P-dependent in-ice ligation of RNA substrates.[18] Thus, we probed whether in-ice activation of RNA substrates by DAP could enable direct downstream HPz-dependent ligation reactions.

To monitor the coupled activation–ligation reactions, we developed a reporter electrophoretic mobility shift assay (EMSA) based on a *cis*-ligating version of the HPz (**cHPz**, Figures 2, S17). We confirmed the self-ligation activity of **cHPz** using >P RNA (**sub>P**) generated with ethyl-3-(3′-dimethylaminopropyl)-carbodiimide (EDC), with which circa 88 % self-ligation was observed via urea-PAGE in reaction conditions used in previous HPz in-ice studies (Figure S18).[18] No band shift was observed in the presence of **sub-P**, the non-activated 3′-monophosphate RNA.

Following the implementation of the reporter EMSA, we sought to test the efficacy of the coupled DAP activation–ligation reactions under various conditions. Specifically, we tested a range of concentrations of DAP, Mg^{2+}, and imidazole for their potential to enable coupled activation–ligation reactions. We identified the inclusion of 5 mM DAP, 5 mM Mg^{2+}, and 5 mM imidazole as optimal conditions thus far for the combined activation and ligation of the RNA substrate **sub-P** (Figure S19), and used these conditions in all further experiments. We also noted an increased ligation yield at

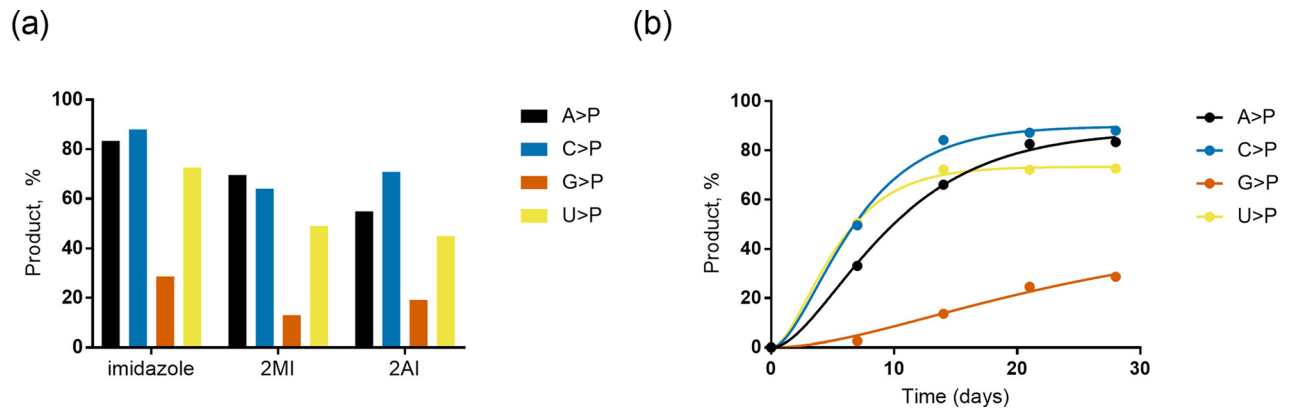


Figure 1. a) Comparison of activator performance in the formation of N>Ps from their respective 3′-NMPs (1 mM) after 28 days at pH 6 and −20 °C in the presence of 1 mM DAP, 5 mM $MgCl_2$, and 5 mM of either imidazole, 2-methylimidazole (2MI), or 2-aminoimidazole (2AI); similar trends were observed at pH 5, 7, and 8 (Figures S13–S16). b) Percentage conversion to >P as a function of time with imidazole as activator at pH 6.

 © 2020 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH

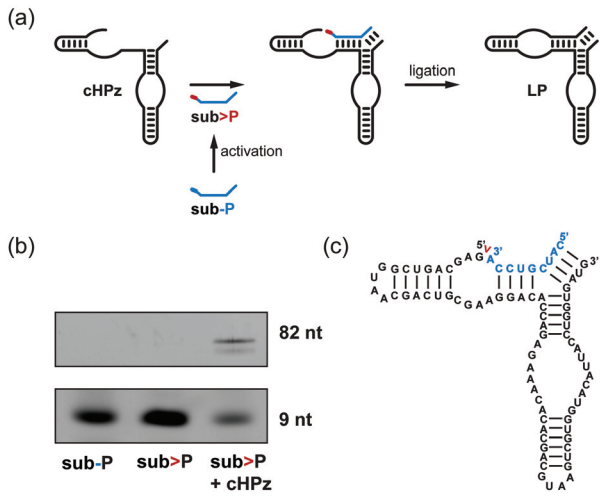


Figure 2. A reporter electrophoretic band shift assay based on a modified hairpin ribozyme (**cHPz**) which accepts a 2′,3′-cyclic phosphate substrate (**sub > P**) for its catalytic activity. a) A simplified schematic of the reaction during which the substrate (**sub-P**) is activated to yield the 2′,3′-cyclic phosphate RNA oligomer. LP = ligation product. b) Denaturing EMSA of the species shown in (a). **sub-P** is 5′-tagged with a fluorescein derivative to aid in analysis. c) A detailed secondary structure of the substrate–ribozyme complex prior to ligation. The location of the 2′,3′-cyclic phosphate and site of ligation are indicated in red.

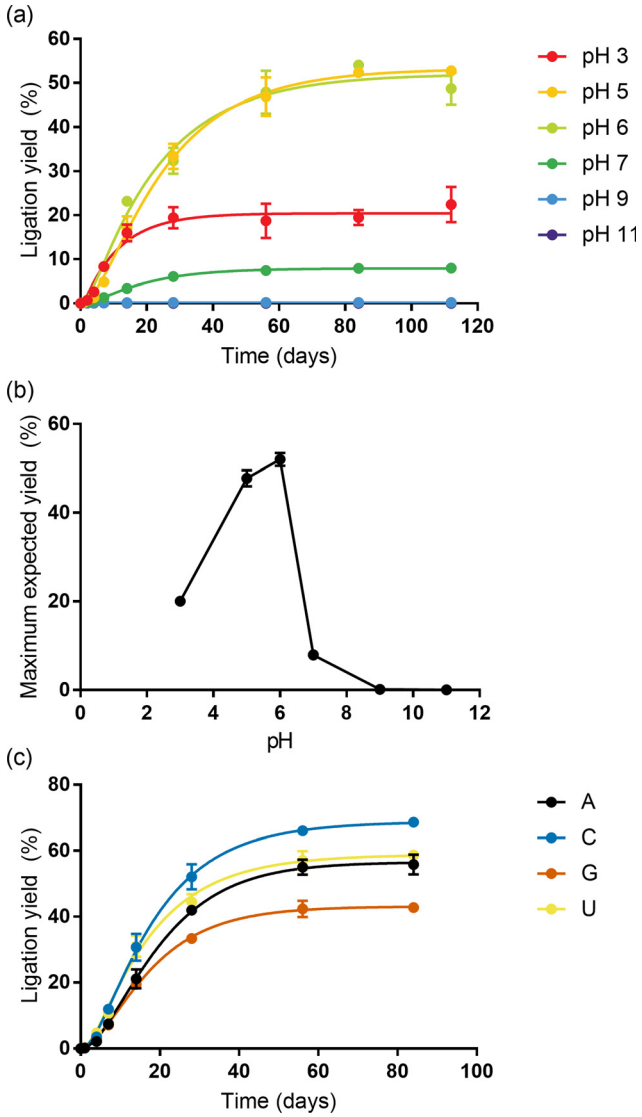


Figure 3. DAP-mediated formation of 2′,3′-cyclic phosphate RNA and its subsequent ligation in the presence of imidazole occur under various conditions compatible with ribozyme catalysis and are optimal at pH 5–6. a) Time course analysis of in situ activation–ligation reactions at different pH values. Reaction mixtures containing 10 mM of glycine·HCl buffer (pH 3), MES buffer (pH 5–7), Tris·HCl buffer (pH 9), or carbonate/bicarbonate buffers (pH 11) were analyzed via TBE-urea PAGE. Data points were fitted assuming a simplified 2-step reaction mechanism to extrapolate maximum ligation yields. b) pH dependency of the maximum ligation yield using the best-fit parameters from (a). Error bars indicate the 95% confidence intervals. c) Dependency of coupled activation–ligation reactions on the 3′-terminal base of **sub-P**. Reactions were performed using four 2′/3′-monophosphorylated variants and analyzed via TBE-urea PAGE.

50 mM DAP in the absence of imidazole. We presumed that the increased concentration of DAP, especially in eutectic ice, compensates for the absence of imidazole. However, increasing both DAP and imidazole concentrations to 50 mM resulted in a lower ligation yield compared to 5 mM (Figure S20). This result may be due to the high DAP and imidazole concentrations that reduce the concentration-by-freezing efficiency for the RNA components. The higher initial concentration of the two components reduces the amount of water-ice necessary to reach the molal equilibrium concentrations of the unfrozen phase.[24] Thus, the resulting cHPz and substrate concentrations in the aqueous phase may be lower than with decreased DAP and imidazole concentrations, leading to less efficient activation and/or ligation.

We then explored the pH dependency of the activation–ligation reaction. We speculated that a highly acidic environment may decrease the availability of the DAP-activating agent imidazole ($pK_a = 6.95$[25]), whose nucleophilicity decreases upon protonation. Moreover, the activity of HPz decreases below pH 5[26] while the stability of RNA is optimal at pH 4–5. Indeed, we did not detect significant RNA degradation during the experiment (Figure S21). Consistent with these considerations and with our results with monomeric ribonucleotides, we observed the highest ligation yields and reaction rates at pH 5–6 (Figures 3 and S22). We postulate that a nucleophilic pool of unprotonated imidazole is maintained at this mildly acidic pH while preserving the activity of cHPz. The decrease in yields at pH above 6 likely reflects that > P formation requires protonation of the -NH_2 group of DAP ($pK_a \approx 5.5$).[27] This assumption is supported by the pH dependence of the activation of 3′-NMPs (Figures S1–S12). While we note that it is typically difficult to estimate the pH of the interstitial brine in water-ice accurately, we expect that the pH in our buffered samples is only increased by about 0.5 pH units at −9 °C.[28]

We initially observed a considerable influence of the base moiety on DAP-mediated phosphate activation for the four mononucleotides (Figure 1b). Therefore, we wondered whether a similar underlying nucleobase dependency might govern the DAP activation of RNA oligonucleotides in the

 © 2020 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH

EMSA analysis. Indeed, we observed increased ligation yields for 3′-terminal pyrimidines over purines (Figure 3c). The difference was most obvious between substrates with a 3′-terminal C and G: After 84 days, the 3′-terminal pyrimidine resulted in 68% ligation yield compared to the purine at 43%. The yield differences are unlikely to be the result of a preference of the cHPz for different 3′ ends, because the influence of the 3′ base identity was only marginal in control experiments with pre-activated substrates (Figure S23). While the solubilities of monophosphates (CMP, 16.3 $g L^{-1}$; UMP, 12 $g L^{-1}$; AMP, 8 $g L^{-1}$; GMP, 8 $g L^{-1}$)[29] may be a factor for the different activation yields at the monomer level, this is unlikely with oligonucleotides. pK_a values of the 2′-OH group, between 13.22 and 13.47,[30] would predict minimal differences in the nucleophilicity of the 2′-OH groups of the four nucleotides for the formation of >Ps. However, it is possible that a larger fraction of the nucleobases A and C is protonated at pH 5–6 and becomes more electron-withdrawing, thereby increasing the acidity of the 2′-OH group.[31] For U, the 2-keto-oxygen may coordinate with the 2′-OH group, thus facilitating its deprotonation.[32]

To further test the generality of the reaction, we investigated the influence of the co-activator imidazole on substrate activation. 2-methylimidazole (2MI) and 2-aminoimidazole (2AI) are close analogues of imidazole (IMI) that are considered as superior 5′-leaving groups for non-enzymatic copying of RNA templates.[33,34] We repeated our EMSA assay with 5 mM 2MI or 2AI instead of IMI (Figure S25) and found that ligation yields were considerably lower (approximately 25% after 84 days) in the presence of 2AI or 2MI compared to IMI (approximately 55%). The lower pK_a of IMI (6.95 compared to 8.46 for 2AI[35] and 7.86 for 2MI[36])—and therefore its larger unprotonated nucleophilic pool—coupled with less steric hindrance may amplify the formation of the activated imidazole-amidophosphate that is responsible for phosphorylation. Nevertheless, our results show that different imidazole derivatives can serve as activators for the DAP-dependent generation of cyclic phosphates.

Next, we examined the performance of DAP compared to the prebiotically implausible carbodiimide EDC, which is typically used to generate >Ps. Intriguingly, the long-term kinetics of RNA activation under in situ ribozyme catalysis-compatible conditions showed higher ligation yields with DAP compared to EDC (Figure 4a). While one-pot in situ activation with EDC caused higher initial yields during the first 10 days of incubation, likely due to higher reactivity of the carbodiimide moiety, the yields in the DAP-based reactions significantly exceeded those incubated with EDC during long-term incubation, reaching >50% more RNA ligation after 28 days. A plausible explanation for this observation is the overall low stability of EDC due to hydrolysis, which hampers reactivation of hydrolyzed cyclic phosphates. In contrast, the high stability of DAP may allow repeated reactivation of hydrolyzed 3′-terminal phosphates back to >Ps. Furthermore, EDC modifies nucleic acid base moieties at rates comparable to those of RNA-catalyzed ligation.[37–39] In contrast, we observed no indication of similar undesirable side reactions with DAP.

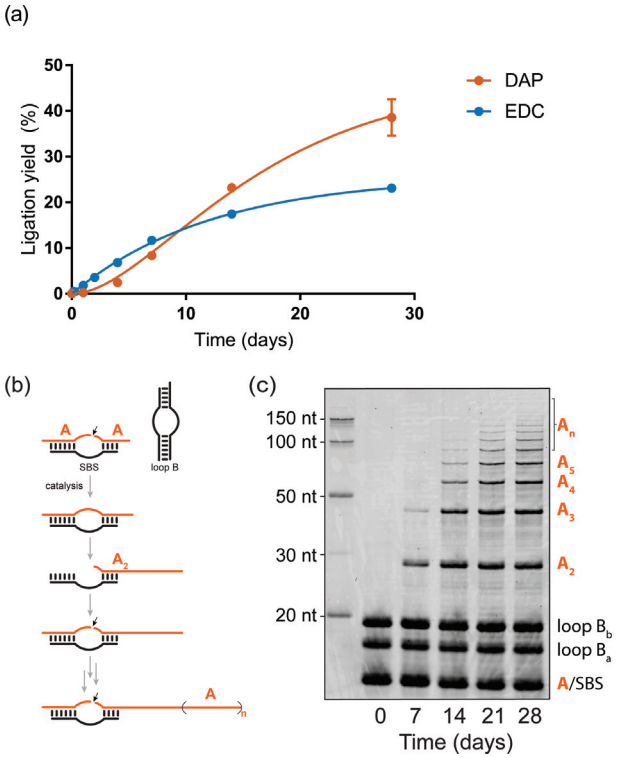


Figure 4. a) Comparison of DAP and EDC under in situ activation–ligation conditions. Reaction mixtures differing only in the inclusion of either EDC or DAP were analyzed via TBE-urea PAGE. The data points were fitted assuming a simplified irreversible 2-step reaction mechanism. b) Simplified schematic of 16 nt RNA oligomer concatenation. SBS = substrate-binding strand. c) DAP-mediated RNA concatenation. Time course analysis via SYBR Gold-stained TBE-urea PAGE of reaction mixtures containing a fragmented ribozyme and 16 nt 3′-monophosphorylated substrates. We note that the RNA ladder displays lower electrophoretic mobility, presumably due to different salt concentrations in the loading buffer and/or conjugated fluorophores.

Finally, we investigated whether DAP activation also allows multiple ligations to produce long RNAs, which is a prerequisite for the formation of more complex RNA molecules. To this end, we probed our one-pot activation/ligation scheme for concatenation reactions of RNA oligomers (Figures 4b,c and S26), including the assembly of an RNA polymerase ribozyme from seven ≤30 nt fragments (RPR7, Figure S26) catalyzed by fragmented hairpin ribozymes.[19] Strikingly, for both reactions we observed six or more successive ligation events after 21 days of in-ice incubation.

The data herein present DAP-mediated formation of activated RNA substrates under conditions compatible with ribozyme catalysis. DAP is an attractive candidate reagent for the prebiotic activation of RNA due to its long half-life and reactivity in aqueous environments.[40] Moreover, its potential in the primordial activation of lipids, peptides, and 5′ ends of nucleic acids—and therefore as a key reagent in the origin of all life on Earth—has been documented.[12,13] Furthermore, the compatibility of the reaction with frozen water-ice matrices that upconcentrate solutes allows low initial DAP and imidazole concentrations to achieve high-yielding activation of mono- and oligonucleotides. DAP is therefore

 © 2020 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH

a powerful reagent to mitigate quasi-irreversible hydrolysis of 2′,3′-cyclic phosphates that occurs over time in the presence of M^{2+} ions,[41] and to sustain a pool of activated >P RNA molecules that can serve as starting material for both enzymatic and non-enzymatic ligation reactions.

>P RNA can also be formed by enzymatic or non-enzymatic RNA cleavage. In combination with strand dissociation, consecutive cleavage–ligation reactions may therefore lead to the formation of longer products.[19,22,42–44] While offering an intriguing alternative for RNA synthesis in the absence of direct chemical activation, such recombination reactions have several disadvantages. The extra cleavage and dissociation steps slow product formation and decrease final yields due to a higher number of reversible steps required for product formation.[19,22] Moreover, ribozyme-catalyzed recombination requires the presence of additional sequence elements on substrates which reduces the pool of compatible oligonucleotides in randomized libraries. Finally, recombination does not regenerate hydrolyzed >Ps. Thus, while both direct ligation and recombination are capable of exploring a large sequence space through repeated ligation reactions,[22] both reactions can ultimately benefit from the mild re-activating properties of DAP.

The resulting high-diversity products could provide the starting material for evolutionary processes in the form of nascent ribozymes.[22] Additionally, continuous activation of RNA substrates provides the possibility of maintaining cross-catalytic reaction networks consisting of several ribozyme components[19] and may enable new strategies for continuous RNA evolution under prebiotically plausible conditions. In conclusion, our work underlines the importance of identifying ribozyme catalysis-compatible RNA activation reagents in enabling one-pot processes, which has implications for abiotic molecular evolution of nucleic acids.

Acknowledgements

H.M. and E.Y.S. acknowledge funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—Project-ID 364653263—TRR 235. H.M. gratefully acknowledges support by the MaxSynBio consortium, which is jointly funded by the Federal Ministry of Education and Research of Germany and the Max Planck Society. H.M. and K.L.V acknowledge support by the Volkswagen Foundation. Funding jointly from the NSF and the NASA Exobiology Program, under the NSF-Center for Chemical Evolution, CHE-1504217, and an award from the Simons Foundation (327124FY19) to R.K. is gratefully acknowledged. Open access funding enabled and organized by Projekt DEAL.

Conflict of interest

The authors declare no conflict of interest.

Keywords: diamidophosphate · early Earth · prebiotic chemistry · ribozymes · RNA

[1] J. W. Szostak, *Angew. Chem. Int. Ed.* **2017**, *56*, 11037 – 11043; *Angew. Chem.* **2017**, *129*, 11182 – 11189.
[2] D. P. Bartel, J. W. Szostak, *Science* **1993**, *261*, 1411 – 1418.
[3] W. K. Johnston, P. J. Unrau, M. S. Lawrence, M. E. Glasner, D. P. Bartel, *Science* **2001**, *292*, 1319 – 1325.
[4] D. P. Horning, G. F. Joyce, *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 9786 – 9791.
[5] J. E. Moretti, U. F. Müller, *Nucleic Acids Res.* **2014**, *42*, 4767 – 4778.
[6] M. A. Pasek, M. Gull, B. Herschy, *Chem. Geol.* **2017**, *475*, 149 – 170.
[7] K. Gao, L. E. Orgel, *Origins Life Evol. Biospheres* **2000**, *30*, 45 – 51.
[8] R. Lohrmann, *J. Mol. Evol.* **1977**, *10*, 137 – 154.
[9] H. Sawai, L. E. Orgel, *J. Am. Chem. Soc.* **1975**, *97*, 3532 – 3533.
[10] T. Inoue, L. E. Orgel, *J. Mol. Biol.* **1982**, *162*, 201 – 217.
[11] T. Walton, S. DasGupta, D. Duzdevich, S. S. Oh, J. W. Szostak, *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 5741 – 5748.
[12] C. Gibard, S. Bhowmik, M. Karki, E. K. Kim, R. Krishnamurthy, *Nat. Chem.* **2018**, *10*, 212 – 217.
[13] C. Gibard, I. B. Gorrell, E. I. Jiménez, T. P. Kee, M. A. Pasek, R. Krishnamurthy, *Angew. Chem. Int. Ed.* **2019**, *58*, 8151 – 8155; *Angew. Chem.* **2019**, *131*, 8235 – 8239.
[14] J. A. Gerlt, F. H. Westheimer, J. M. Sturtevant, *J. Biol. Chem.* **1975**, *250*, 5059 – 5067.
[15] S. Loverix, G. Laus, J. C. Martins, L. Wyns, J. Steyaert, *Eur. J. Biochem.* **1998**, *257*, 286 – 290.
[16] S. M. Nesbitt, H. A. Erlacher, M. J. Fedor, *J. Mol. Biol.* **1999**, *286*, 1009 – 1024.
[17] A. V. Vlassov, B. H. Johnston, L. F. Landweber, S. A. Kazakov, *Nucleic Acids Res.* **2004**, *32*, 2966 – 2974.
[18] H. Mutschler, P. Holliger, *J. Am. Chem. Soc.* **2014**, *136*, 5193 – 5196.
[19] H. Mutschler, A. Wochner, P. Holliger, *Nat. Chem.* **2015**, *7*, 502 – 508.
[20] W. G. Scott, A. Szöke, J. Blaustein, S. M. O'Rourke, M. P. Robertson, *Life* **2014**, *4*, 131 – 141.
[21] A. V. Lutay, E. L. Chernolovskaya, M. A. Zenkova, V. V. Vlassov, *Biogeosciences* **2006**, *3*, 243 – 249.
[22] H. Mutschler, A. I. Taylor, B. T. Porebski, A. Lightowlers, G. Houlihan, M. Abramov, P. Herdewijn, P. Holliger, *eLife* **2018**, *7*, e43022.
[23] J. Attwater, A. Wochner, V. B. Pinheiro, A. Coulson, P. Holliger, *Nat. Commun.* **2010**, *1*, 76.
[24] P. A. Monnard, H. Ziock, *Chem. Biodiversity* **2008**, *5*, 1521 – 1539.
[25] H. Walba, R. W. Isensee, *J. Phys. Chem.* **1961**, *26*, 2789 – 2791.
[26] S. Kath-Schorr, T. J. Wilson, N. S. Li, J. Lu, J. A. Piccirilli, D. M. J. Lilley, *J. Am. Chem. Soc.* **2012**, *134*, 16717 – 16724.
[27] K. Ashe, *Studies towards the Prebiotic Synthesis of Nucleotides and Amino Acids*, University College London, London, **2018**.
[28] P. Kolhe, E. Amend, S. K. Singh, *Biotechnol. Prog.* **2010**, *26*, 727 – 733.
[29] D. S. Wishart, C. Knox, A. C. Guo, S. Shrivastava, M. Hassanali, P. Stothard, Z. Chang, J. Woolsey, *Nucleic Acids Res.* **2006**, *34*, D668 – D672.
[30] I. Velikyan, S. Acharya, A. Trifonova, A. Földesi, J. Chattopadhyaya, *J. Am. Chem. Soc.* **2001**, *123*, 2893 – 2894.
[31] R. Krishnamurthy, *Acc. Chem. Res.* **2012**, *45*, 2035 – 2044.
[32] A. Bibillo, M. Figlerowicz, K. Ziomek, R. Kierzek, *Nucleosides Nucleotides Nucleic Acids* **2000**, *19*, 977 – 994.
[33] T. Inoue, L. E. Orgel, *J. Am. Chem. Soc.* **1981**, *103*, 7666 – 7667.
[34] T. Walton, J. W. Szostak, *J. Am. Chem. Soc.* **2016**, *138*, 11996 – 12002.
[35] B. T. Storey, W. W. Sullivan, C. L. Moyer, *J. Org. Chem.* **1964**, *29*, 3118 – 3120.

 © 2020 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH

[36] D. Perrin, *Dissociation Constants of Organic Bases in Aqueous Solution: Supplement 1972*, Butterworths, London, **1972**.
[37] C. Obianyor, G. Newnam, B. Clifton, M. Grover, N. V. Hud, *ChemBioChem* **2020**, *21*, 3359 – 3370.
[38] E. Edeleva, A. Salditt, J. Stamp, P. Schwintek, J. Boekhoven, D. Braun, *Chem. Sci.* **2019**, *10*, 5807 – 5814.
[39] P. Y. Wang, A. N. Sexton, W. J. Culligan, M. D. Simon, *RNA* **2019**, *25*, 135 – 146.
[40] R. Krishnamurthy, S. Guntha, A. Eschenmoser, *Angew. Chem. Int. Ed.* **2000**, *39*, 2281 – 2285; *Angew. Chem.* **2000**, *112*, 2369 – 2373.
[41] A. Dallas, A. V. Vlassov, S. A. Kazakov, *Artificial Nucleases*, Springer, Berlin, **2004**, pp. 61 – 88.
[42] S. Petkovic, S. Müller, *FEBS Lett.* **2013**, *587*, 2435 – 2440.
[43] S. Petkovic, S. Badelt, S. Block, C. Flamm, M. Delcea, I. Hofacker, S. Müller, *RNA* **2015**, *21*, 1249 – 1260.
[44] R. Hieronymus, S. P. Godehard, D. Balke, S. Müller, *Chem. Commun.* **2016**, *52*, 4365 – 4368.

Manuscript received: August 9, 2020
Revised manuscript received: October 29, 2020
Accepted manuscript online: October 30, 2020
Version of record online: December 10, 2020

 © 2020 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH

2.2 PUBLICATION II

Enhanced Ribozyme-Catalyzed Recombination and Oligonucleotide Assembly in Peptide-RNA Condensates

AUTHORS: Dr. Kristian Le Vay[+*], **Emilie Yeonwha Song**[+], Dr. Basusree Ghosh, Dr. T.-Y. Dora Tang[*], Prof. Dr. Hannes Mutschler[*]

[+] *authors contributed equally (shared 1st authorship).* [*] *corresponding authors.*

JOURNAL, YEAR, VOLUME, AND PAGE NO.:

Angew. Chem. Int. Ed., **2021**, *60*, 26096-26104

2.2.1 SUMMARY

The generation of nucleic acid polymers with sufficient length to fold into functional motifs is a critical requirement for the emergence of self-replicating systems in early molecular evolution. While liquid-liquid separated condensed phases have long been proposed to have hosted primordial processes, it is currently unclear whether they can support and enhance constructive catalytic processes. Thus far, only reactions that lead to shorter nucleic acid lengths have been demonstrated therein.

In this publication, we describe the first example of phase separation significantly enhancing the constructive catalytic activity of a *trans*-acting hairpin ribozyme, shifting the reaction equilibrium to ligation rather than cleavage. The robust activity is maintained over a broad range of temperatures and magnesium concentrations, even in solution conditions where activity is otherwise not observed. The condensates formed from short fragmented ribozyme strands and poly(lysine) oligopeptides enable the generation of long RNA chains and a model RNA polymerase ribozyme from short fragments. Further, strong ribozyme activity is observed in both gel-like and coacervate condensed phases, finally reconciling constructive ribozyme activity with an important protocellular model.

These results experimentally demonstrate how catalytic nucleic acids and protocellular environments may have synergized to facilitate the assembly of complex informational polymers at the origin of life.

2.2.2 PERSONAL CONTRIBUTIONS

- Performed $(Lys)_{5-24}$ and $(Lys)_{19-72}$ ratio, temperature, $[Mg^{2+}]$, and kinetics screens independently and together with K. K. Le Vay
- Performed RPR4 assembly reactions independently and together with K. K. Le Vay
- Prepared figures independently and together with K. K. Le Vay (main figures, supplementary figures 4 and 5)
- Composed, edited, and revised the manuscript and supporting information together with K. K. Le Vay, B. Ghosh, T.-Y. D. Tang, and H. Mutschler

2.2.3 AUTHORSHIP DECLARATION

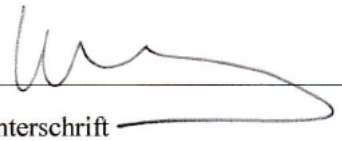
Declaration on the substantial contribution to a publication submitted as part of a cumulative dissertation

I hereby declare that, for the publication titled "Enhanced Ribozyme-Catalyzed Recombination and Oligonucleotide Assembly in Peptide-RNA Condensates", both co-first authors made substantial and equal contributions, involving experimental design, execution, and preparation of the published manuscript.

09.08.2022 München

Ort und Datum

Place and date



Unterschrift

Signature

10.08.2022 Munich

Ort und Datum

Place and date

Unterschrift

Signature

Check for updates

Peptides Hot Paper

How to cite: *Angew. Chem. Int. Ed.* **2021**, *60*, 26096–26104
International Edition: doi.org/10.1002/anie.202109267
German Edition: doi.org/10.1002/ange.202109267

Enhanced Ribozyme-Catalyzed Recombination and Oligonucleotide Assembly in Peptide-RNA Condensates

Kristian Le Vay[+], Emilie Yeonwha Song[+], Basusree Ghosh, T.-Y. Dora Tang,* and Hannes Mutschler**

Abstract: *The ability of RNA to catalyze RNA ligation is critical to its central role in many prebiotic model scenarios, in particular the copying of information during self-replication. Prebiotically plausible ribozymes formed from short oligonucleotides can catalyze reversible RNA cleavage and ligation reactions, but harsh conditions or unusual scenarios are often required to promote folding and drive the reaction equilibrium towards ligation. Here, we demonstrate that ribozyme activity is greatly enhanced by charge-mediated phase separation with poly-L-lysine, which shifts the reaction equilibrium from cleavage in solution to ligation in peptide-RNA coaggregates and coacervates. This compartmentalization enables robust isothermal RNA assembly over a broad range of conditions, which can be leveraged to assemble long and complex RNAs from short fragments under mild conditions in the absence of exogenous activation chemistry, bridging the gap between pools of short oligomers and functional RNAs.*

Introduction

The generation of RNA or other nucleic acid strands with sufficient length and sequence diversity to fold into functional catalysts and replicators is a key step in many models of early biology. Plausible pathways exist for the generation of activated ribonucleotides[1–3] and the formation of short oligomers by non-enzymatic polymerization,[4,5] but the untemplated polymerization of long RNA oligomers containing all four nucleobases remains inefficient.[6] The formation of long and complex RNAs from pools of oligonucleotides can potentially bridge this gap,[7] as demonstrated by the assembly of a >200 nt RNA polymerase ribozyme from short oligomers by a fragmented version of the small nucleolytic hairpin ribozyme (HPz).[8]

HPz is a small self-cleaving RNA motif that catalyzes the cleavage and ligation of specific RNA sequences.[9,10] The ability of HPz to fold into active conformations[11,12] is sensitive to environmental conditions such as ionic strength[13,14] and temperature.[15] In solution, the reaction equilibrium is shifted towards cleavage, whilst in dehydrating conditions (e.g. eutectic phases in ice, ethanol and drying) efficient ligation can be observed.[16–18] The ability of HPz to assemble from short oligonucleotides and to process a wide range of RNA junctions makes it an attractive model system for an early RNA catalyst.[8,16] The reversibility of the HPz-catalyzed reaction allows for the possibility of recombination (Figure S1): First, RNA cleavage produces a strained 2′,3′-cyclic phosphate functionalized substrate. Next, strand exchange then ligation can occur with other strands of a compatible sequence. Although relieving strain in the 2′,3′-cyclic phosphate provides a limited enthalpic driving force for ligation, the associated entropic penalty results in cleavage being thermodynamically favored in solution for some systems.[19] These reactions are prebiotically appealing as they can occur independently from an exogenous RNA-activation chemistry and have been shown to increase structural and informational diversity in pools of RNA oligomers.[20] Researchers have exploited recombination and ligation by the HPz ribozyme to assemble complex functional RNAs from short fragments,[8,21] to combine RNA enzymes and aptamers into larger, more complex aptazyme systems,[22] and to produce simple self-replicators.[23] In combination with a polymerization chemistry capable of generating short oligomers, the realization of replicating and evolving systems may be possible. Ribozyme-catalyzed recombination can require high magnesium concentrations to occur in solution, with minimal or fragmented HPz variants reliant on dehydration or freezing to drive ligation.[16,18] These requirements render recombination-based RNA assembly incompatible with compartmentalization in some protocellular systems and narrow the range of viable environmental conditions. For example, high Mg^{2+} concentrations have a destabilizing effect on vesicles formed from single chain amphiphiles,[24] while

[*] Dr. K. Le Vay,[+] E. Y. Song[+]
Biomimetic Systems, Max Planck Institute of Biochemistry
Am Klopferspitz 18, 82152 Martinsried (Germany)
E-mail: levay@biochem.mpg.de
Dr. B. Ghosh, Dr. T.-Y. D. Tang
Max-Planck Institute of Molecular Cell Biology and Genetics
Pfotenhauerstraße 108, 01307 Dresden (Germany)
E-mail: tang@mpi-cbg.de
Dr. K. Le Vay,[+] E. Y. Song,[+] Prof. Dr. H. Mutschler
Department of Chemistry and Chemical Biology, TU Dortmund University
Otto-Hahn-Str. 4a, 44227 Dortmund (Germany)
E-mail: hannes.mutschler@tu-dortmund.de
[+] These authors contributed equally to this work.
Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under: https://doi.org/10.1002/anie.202109267.

© 2021 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution Non-Commercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

phospholipid-based vesicles are prone to fragmentation[25,26] and content loss[27] when exposed to freezing conditions.

Phase separation by oppositely charged polymers such as peptides and nucleic acids is a ubiquitous process that is both widely observed in modern biology[28–30] and exploited in synthetic biology.[31,32] Charged peptides are appealing components in origin of life scenarios, potentially supporting the function of early catalytic nucleic acids[33,34] and forming protocellular compartments such as coacervate droplets by phase separation.[35] Although liquid coacervate droplets are the most widely studied system within the context of the origin of life, other phases such as coaggregates[36,37] and gels[38] are also of interest. Spatial confinement, as well as the ability to localize reactions and concentrate key components, would have been valuable in early biology and is a prerequisite for open-ended evolution.[39] The ability of coacervate phases to strongly concentrate divalent metal ions and RNA suggests a highly favorable environment for nucleic acid catalysis,[40] and interactions with polycationic peptides have been shown to enhance ribozyme activity in some cases.[41,42]

Despite this, catalysis by ribozymes has only recently been reconciled with coacervation,[43,44] and specific conditions are required to prevent loss of function due to misfolding. Previously reported enhancements in catalytic activity by condensed phases typically rely on the concentration and rescue of highly dilute systems below the ribozyme / substrate dissociation constant (K_D),[44] and so far only ribozyme cleavage reactions have been demonstrated.[44] In order to realize complex, functional phase-separated systems that may eventually be capable of open-ended evolution, constructive processes such as the polymerization, ligation or recombination of nucleic acids must be supported.

In this work, we utilize poly-L-lysine ($(Lys)_n$) and a split HPz ribozyme as a model for interactions between early charged protopeptides and nucleic acid catalysts. Although cationic amino acids are crucial factors in protein-nucleic acid interactions, and thus are essential components in any prebiotic scenario that leads to the emergence of protein synthesis,[45] basic cationic amino acids such as arginine and lysine only rarely emerge from model prebiotic syntheses.[46–49] However, ligation of these amino acids into peptides has been demonstrated in prebiotic conditions,[50] assuming sufficient feedstocks. The synthesis of non-proteinaceous analogues such as ornithine and 2,4-diamino-butyric acid, potential precursors to extant cationic amino acids,[51] is more plausible.[51–53] Peptides composed of these analogues can act as functional substitutes to poly-L-lysine and have provided similar enhancements in ribozyme activity.[42] Depsipeptides, protopeptides which contain a mixture of ester and amide linkages, can be formed by dry-wet cycles,[54] which selectively incorporate proteinaceous amino acids over non-proteinaceous precursors, thus providing a potential mechanism for the enrichment of poorly abundant cationic amino acids in early polypeptides.[55] Unlike the model system used here, peptides or protopeptides emerging from prebiotic processes are likely not to have been homopeptides, or enantiomerically pure.

We show that charge-based phase separation between $(Lys)_n$ and the HPz ribozyme RNA both enhances and modulates ribozyme activity, shifting the HPz reaction equilibrium from cleavage to ligation at ambient temperatures. This enables the production of both long and complex functional RNA from short fragments under isothermal conditions. Unlike under solution conditions, poly-L-lysine renders the HPz ribozyme functional over a broad range of temperatures and magnesium concentrations, even with varying lengths, charge ratios, and phase behaviors. These findings reconcile HPz-catalyzed RNA recombination with a far broader range of environmental conditions, and furthermore suggest the ability of simple peptides or other more diverse polycations to support and enhance nucleic acid catalysis in heterogeneous prebiotic scenarios.[56]

Results

To investigate the recombination activity of the HPz ribozyme, we developed an assay that allows the independent characterization of both cleavage and ligation, based on a fragmented ribozyme system (Figure 1a).[16] The ribozyme system comprises four strands forming two catalytic loops: the HPz loop B (two strands, 18 and 21 nt) and loop A, formed from a 14 nt substrate binding strand (SBS) and 3′-Cy5 tagged 20 nt substrate (subC). Initially the ribozyme may cleave subC into two 10 nt products, one possessing the 3′-Cy5 tag (5fragC), the other activated with a 2′,3′-cyclic phosphate and capable of ligation (3frag). The reaction mixture also contains a 10 nt fragment identical in sequence to the 5fragC produced in the previous step but bearing a 3′-FAM tag (5fragF). This fragment can compete for the substrate binding strand and forms a new FAM-tagged 20 nt strand (subF) when ligated to the activated 3frag. The mechanism of recombination is described in Figure S1, and an example gel showing the activity of this system under various conditions is shown in Figure S2. In solution (Tris·HCl pH 8, 1 mM $MgCl_2$, 30 °C), the cleavage of subC was slow, yielding only 5% of 5fragC after 24 h (Figure 1b, dashed line). No ligation of 5fragF to form subF was observed under these conditions (Figure 1c, dashed line).

We then investigated the effect of poly-L-lysine on the activity of the HPz system. Turbidity was observed after the addition of $(Lys)_{19–72}$ (charge ratio = 0.67:1 $(Lys)_{19–72}$:RNA), as well as a drastic shift in ribozyme activity. Cleavage of subC proceeded rapidly compared to solution conditions, with approximately 60% cleavage over 24 h ($k_{cl} = 0.42 \pm 0.02\ h^{-1}$). Strikingly, ligation of 5fragF, forming subF, was observed after just 2 h of reaction, with a final yield at 24 h of approximately 50%. Whilst it is not possible to directly determine the rate of ligation in the recombination system due to the initial cleavage step followed by a presumably rate-limiting strand exchange reaction, the ligation of a pre-activated 3frag intermediate to 5fragF proceeded rapidly under these conditions, with an apparent ligation rate of $k_{lig} = 0.076 \pm 0.004\ min^{-1}$ (Figure S3).

We then investigated the effect of polycation length and concentration on HPz ribozyme activity (Figure 2, Figure S4). Increasing concentrations of either $(Lys)_{19–72}$ (Figure 2a) or $(Lys)_{5–24}$ (Figure 2b) were titrated into a fixed concentration

 © 2021 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH

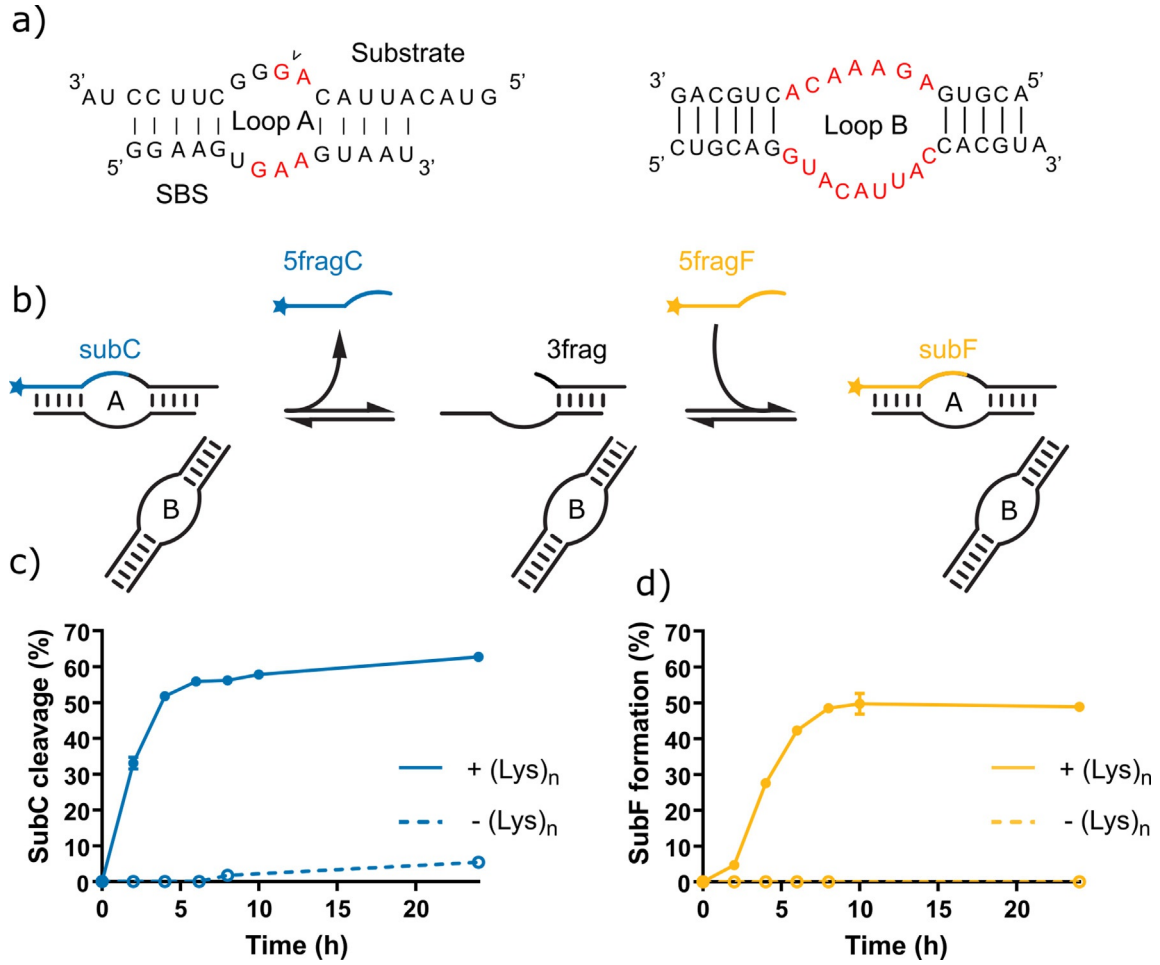


Figure 1. The design and function of a simple split hairpin ribozyme recombination assay. a) The structure of the split HPz ribozyme, which is composed of loop A (substrate and substrate binding strand, or SBS) and loop B. The cleavage site is indicated with an arrow, and conserved bases are shown in red. b) Simplified diagram showing hairpin ribozyme-catalyzed cleavage of a Cy5-tagged 20 nt substrate (subC) (blue), strand exchange, and re-ligation of a FAM-tagged 10 nt fragment (5fragF) (yellow). c) Cleavage of the subC and d) ligation of the 5fragF over time by the HPz ribozyme with $(Lys)_{19–72}$ (0.67:1 $(Lys)_{19–72}$:RNA, 1 mM $MgCl_2$, Tris·HCl pH 8, 30 °C) (solid line) and in solution (1 mM $MgCl_2$, Tris·HCl pH 8, 30 °C) (dashed line).

of RNA, and the degree of phase separation at the various $(Lys)_n$:RNA ratios was measured by absorbance at 500 nm (Figure 2a, b). The critical peptide concentrations for phase separation (CPs) were determined by the onset of turbidity. Phase separation occurred at a sub-stoichiometric concentration of the longer peptide ($CP_{19–72}$ ca. 0.7:1 $(Lys)_{19–72}$:RNA), but not until a concentration ratio of near unity for the shorter peptide ($CP_{5–24}$ ca. 1:1 $(Lys)_{5–24}$:RNA). Activity measurements at $t = 24$ h showed that $(Lys)_{19–72}$ enhances both HPz cleavage and ligation at sub-stochiometric ratios, but that both of these activities are suppressed at high ratios (Figure 2c). The addition of $(Lys)_{5–24}$ (Figure 2d) led to high HPz cleavage activity and relatively weaker ligation compared to the longer $(Lys)_{19–72}$ at all ratios tested, although the enhancement over solution conditions remained considerable. Optimal recombination of 5fragF to form subF was observed at charge ratios of 0.85:1 $(Lys)_{19–72}$:RNA, and 2:1 $(Lys)_{5–24}$:RNA. This enhancement of activity was found to be independent of the presence of the cationic Cy5 fluorophore on 5fragC (Figure S5). To determine the distribution of $(Lys)_n$ oligomers in our samples, we performed electrospray ionization mass spectrometry (ESI-MS) (Figure S6). Surprisingly, the distributions were heavily biased towards shorter oligomers in both cases: for $(Lys)_{5–24}$, only oligomers with $n = 3$ to $n = 9$ were detected, whilst for $(Lys)_{19–72}$, only oligomers between $n = 16$ to $n = 46$ were observed (Figure S7). Whilst oligomers with lengths up to the specified upper limit are likely to be present in the samples, based on the manufacturer's characterization, these species appear to occur at such low concentrations that they fall below the detection limit of ESI-MS.

Brightfield and fluorescence microscopy of samples at each tested $(Lys)_n$:RNA ratio revealed the formation of various separated phases with increasing charge ratio (Figures 2e,f, S8 and S9). Increasing ratios of both $(Lys)_{5–24}$ and $(Lys)_{19–72}$:RNA led to the formation of particles that accumulated without coalescence into liquid droplets, and in which both labelled RNAs were strongly colocalized. This implies strong partitioning of the labelled substrate RNA into the condensed phase, which is likely to be responsible for the shift

 © 2021 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH

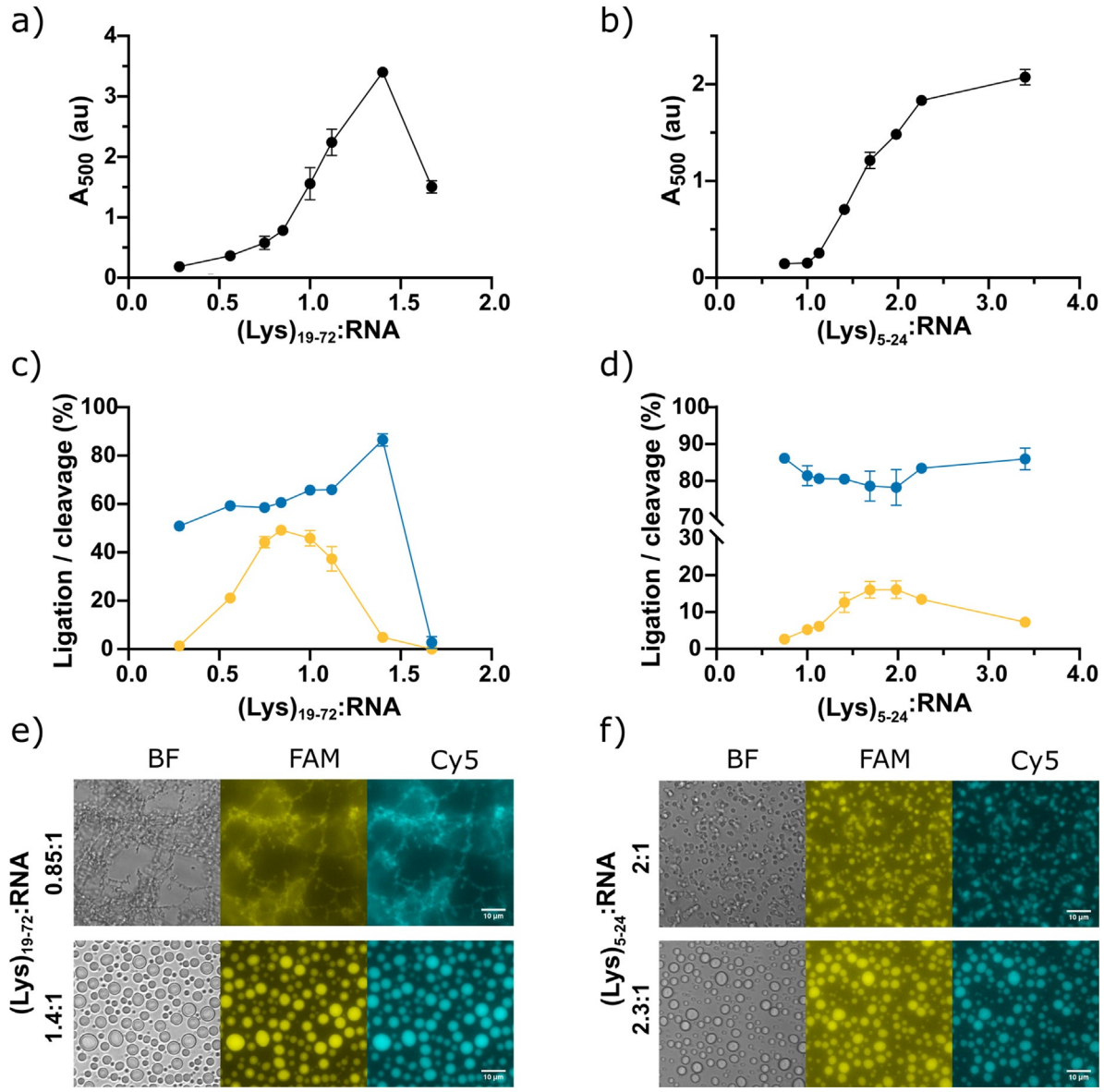


Figure 2. Influence of $(Lys)_n$ length and $(Lys)_n$:RNA mixing ratio on phase separation and HPz activity. a) Variation in absorbance at 500 nm with varying ratios of $(Lys)_{19–72}$:RNA. The minimal ratio for the onset of phase separation CP_{19-72} ca. 0.7:1 $(Lys)_{19–72}$:RNA. b) Variation in absorbance at 500 nm with varying ratios of $(Lys)_{5–24}$. The minimal ratio for the onset of phase separation CP_{5-24} ca. 1:1 $(Lys)_{5–24}$:RNA. c) Endpoint cleavage and ligation activity of the HPz ribozyme with varying $(Lys)_{19–72}$:RNA. d) Endpoint cleavage and ligation activity of the HPz ribozyme with varying $(Lys)_{5–24}$:RNA. Endpoint ligation of 5 fragF is shown in yellow, endpoint cleavage of subC is shown in blue. e) Example fluorescence microscopy images of $(Lys)_{19–72}$:RNA condensates before (0.85:1 $(Lys)_{19–72}$:RNA) and after (1.4:1 $(Lys)_{19–72}$:RNA) the transition to liquid droplets. f) Example fluorescence microscopy images of $(Lys)_{5–24}$:RNA condensates before (2:1 $(Lys)_{5–24}$:RNA) and after (2.3:1 $(Lys)_{5–24}$:RNA) the transition to liquid droplets. Brightfield imaging is shown in grey, FAM fluorescence in yellow and Cy5 fluorescence in cyan. Scale bars = 10 μm. Full imaging of all datapoints is shown in Figures S8 and S9.

to ligation. At $(Lys)_{19–72}$:RNA ratios near unity (0.85-1:1), a web-like structure was observed (Figures 2e, Figure S9). For both poly-L-lysine lengths, a transition to liquid coacervate droplets was observed at excess charge ratios ($(Lys)_{5–24}$:RNA $\geq$ 2.3:1, $(Lys)_{19–72}$:RNA $\geq$ 1.1:1) (Figure 2e,f). The phase separated particles initially formed as an unstable dispersion, eventually settling onto the bottom of the slide without wetting the passivated surface. Ligation activity was optimal at ratios immediately preceding the transition to coacervate droplets for both $(Lys)_n$ lengths ($(Lys)_{5–24}$:RNA = 2:1, $(Lys)_{19–72}$:RNA = 0.85:1). Although the formation of liquid droplets was associated with a reduction in activity, robust ligation activity was still observed in this phase at lower ratios (37.3 % yield at $(Lys)_{19–72}$:RNA = 1.1:1, 13.4 % yield at $(Lys)_{5–24}$:RNA = 2.3:1).

We sought to further characterize the system by studying endpoint recombination activity at the optimal $(Lys)_{19–72}$:RNA ratio for ligation (0.75:1) at various temperatures and magnesium concentrations. The endpoint activities at various temperatures are shown in Figure 3a. Optimal

 © 2021 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH

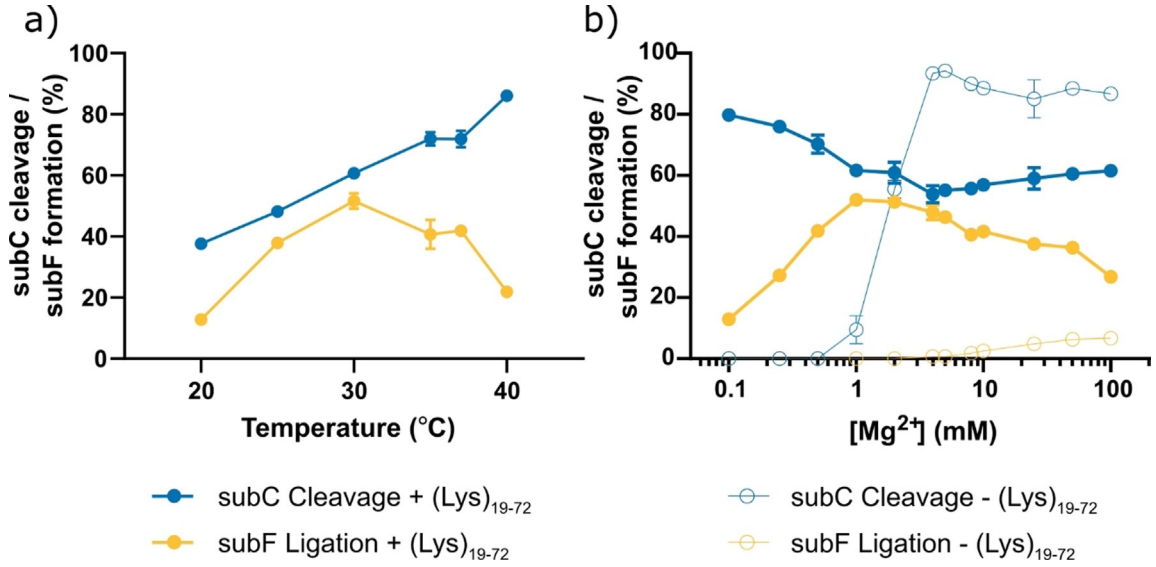


Figure 3. Influence of temperature, magnesium concentration, and poly-L-lysine length and mixing ratio on recombination yield. a) The effect of reaction temperature on the endpoint recombination yield (0.75:1 $(Lys)_{19–72}$:RNA, 1 mM $MgCl_2$, Tris·HCl pH 8). The reaction was quenched at either 144 h (20 °C), 96 h (25 °C) or 24 h (30–40 °C). b) The effect of magnesium concentration on the HPz ribozyme recombination yield with $(Lys)_n$ (0.75:1 $(Lys)_{19–72}$:RNA, 1 mM $MgCl_2$, Tris·HCl pH 8, $t = 24$ h) (solid circles) and in solution (Tris·HCl pH 8, 30 °C, $t = 24$ h) (hollow circles). For both panels, cleavage of subC is shown in blue, whilst ligation of the 5 fragF is shown in yellow.

recombination yields were obtained at 30 °C (1 mM Mg^{2+}, Tris·HCl pH 8, 52 % ligation after 24 h). Similar yields were observed at lower temperatures, although the time required to reach equilibrium increased ($t_{25\,°C} = 96$ h, $t_{20\,°C} = 144$ h). Increasing temperature shifted the reaction equilibrium towards cleavage, as expected due to the entropic cost of ligation inherent in the system.[57] Recombination activity was also supported across a broad range of Mg^{2+} concentrations (0.1–100 mM) in the presence of $(Lys)_{19–72}$ (0.75:1), notably at concentrations as low as 0.1 mM Mg^{2+} (80 % subC cleavage, 13 % 5 fragF ligation) (Figure 3 b). Optimal activity was observed between 1–5 mM Mg^{2+}. Reduced 5 fragF ligation yields above this point may be in part due to magnesium-catalyzed hydrolysis of the 2′,3′-cyclic phosphate, which is expected to slowly deplete the amount of activated 3 frag in the system over time.

To separate the effects of magnesium concentration[40] from other possible causes of recombination enhancement, we also measured activity under identical conditions but in the absence of peptide. Here, the reaction equilibrium was strongly shifted towards cleavage. No cleavage of subC was observed until 1 mM Mg^{2+}, at which point the yield of 5 fragC increased, reaching a maximum at 5 mM Mg^{2+} (94 % cleavage). The change in cleavage activity with magnesium concentration is sigmoidal and reminiscent of a ribozyme folding curve. Indeed, the measured midpoint of the cleavage data, 1.8 mM Mg^{2+}, is similar to previously reported Mg^{2+} induced folding midpoints for the HPz ribozyme,[58,59] although differences in structure and fragmentation make a direct comparison difficult. In the absence of peptide, no ligation of 5 fragF to form subF was observed until 4 mM Mg^{2+} (0.7 % yield), reaching a maximum yield of 7 % at 100 mM Mg^{2+} with a concomitant reduction in the measured proportion of 5 fragC. This suggests that concentration of magnesium ions within the condensed phases does not alone account for the enhanced recombination activity observed in the presence of $(Lys)_n$ in this study.

Having determined optimal conditions for ligation activity (Tris·HCl pH 8, 30 °C, 0.75:1 $(Lys)_{19–72}$:RNA), we exploited the enhancing effect of $(Lys)_{19–72}$ on the HPz ribozyme to form long RNA chains and complex structures from short fragments. The following ribozyme systems contain the same *in trans* loop B domain used in the previously described recombination assay (Figure 1 a), but combined with a range of substrate binding strands (SBSs) and substrates. First, we tested a recombination-based RNA ladder system (Figure 4 a): the fragmented HPz ribozyme cleaves a short 3'-tail from a 22 nt fragment (St), leaving a 16 nt 2′,3′-cyclic phosphate functionalized fragment (S), which can then be concatenated. After 24 h incubation (Tris·HCl pH 8, 30 °C, 0.7:1 $(Lys)_{19–72}$:RNA), the substrate tail was completely cleaved, and a ladder of concatenated products (S_n) was observed (longest product observed: $n = 15$, 240 nt). The lack of uncleaved substrate is expected, as the cleaved 3′-tail shares only three complementary bases with the substrate binding strand, and so should easily be displaced by the 5′ end of another substrate, making re-ligation unfavorable. In addition, the 6 nt fragment may only be poorly concentrated in the condensed phase due to its short length.[43]

RNA polymerase ribozymes (RPRs), which catalyze the templated synthesis of RNA from nucleotide triphosphates, are considered analogues of an early RNA-only replicator.[60] Although these ribozymes are capable of synthesizing long strands, their activity is not sufficient to quantitatively synthesize sequences with complex secondary structure beyond ca. 50 nt,[61,62] and as such self-replication has not been demonstrated. These limitations may be overcome by stepwise modular assembly, in which the large and complex

 © 2021 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH

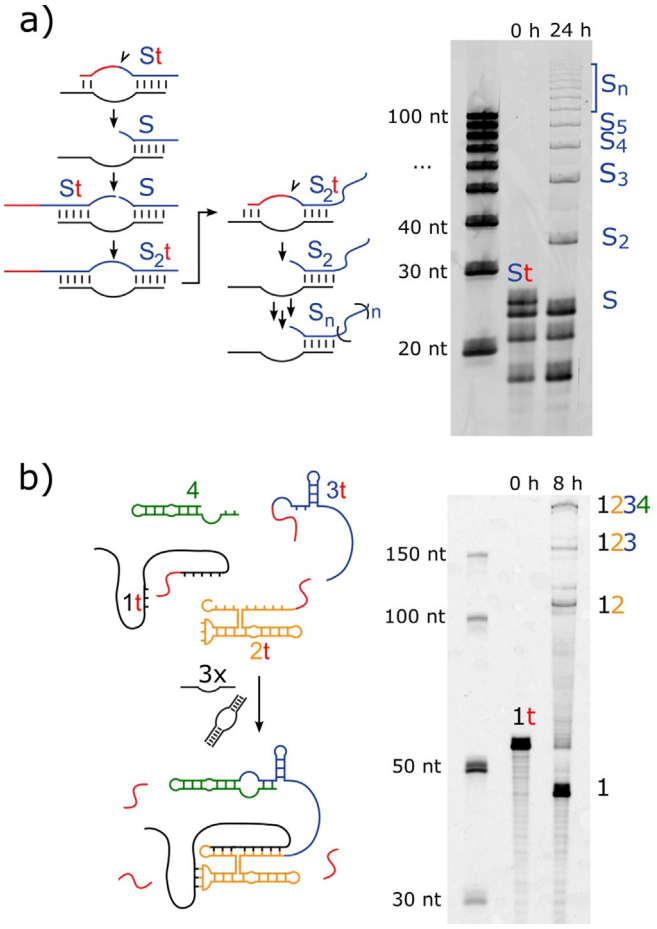


Figure 4. Formation of long RNA chains and complex RNA by HPz recombination under optimized conditions. Example urea-PAGE gels showing a) Concatenation of long RNA chains (>200 nt) by the HPz ribozyme with poly-L-lysine (0.75:1 $(Lys)_{19–72}$:RNA, 1 mM $MgCl_2$, Tris·HCl pH 8, $t = 24$ h). The fragmented HPz ribozyme first binds the 22 nt substrate fragment (St), forming an A loop between substrate and binding strand. This docks with a B loop, then a 6 nt tail (t, shown in red) is cleaved from the substrate leaving a 2′,3′-cyclic phosphate on the 16 nt reaction product (S). The cleavage site is indicated with an arrow. Another substrate (St or S) then displaces the cleaved tail, which has only three bases complementary to the substrate binding arm. In poly-L-lysine coacervates, or other concentrating conditions, the HPz ribozyme can then ligate the two fragments, forming a concatenated (S_2) product. This process repeats, resulting in long chains of up to $n > 13$ substrate fragments. The urea-PAGE gel was imaged using SYBER Gold staining. b) Assembly of the RPR4 ribozyme (198 nt) from four fragments by HPz with poly-L-lysine (0.8:1 $(Lys)_{19–72}$:RNA, 8 mM $MgCl_2$, Tris·HCl pH 8, $t = 8$ h, full length product yield = 8%). In this reaction, three different substrate binding strands bind pairs of oligomers from a set of four substrate strands. A short tail (t) is cleaved from fragments 1, 2 and 3, leaving 2′,3′-cyclic phosphate functionalized fragments that be ligated together with fragment 4 to form the full-length product. The assembly products are visualized using a fluorescent 5′-FAM tag on fragment 1.

functional sequence is constructed from shorter fragments. As a proof of concept, we assembled the RPR4 ribozyme from short fragments in the presence of long poly-L-lysine (Figure 4b), as previously demonstrated in-ice.[8] Indeed, full length RPR4 was produced with a final yield of 7.8% after 8 h (Tris·HCl pH 8, 30 °C, 0.8:1 $(Lys)_{19–72}$:RNA). The activity of the assembled product was not tested in this study. These yields are comparable to those observed in long-term Mg^{2+}-free reactions driven by repeated 12 h freeze-thaw cycles (10% after 24 days). The similarity in the product yields observed in-ice and in the condensed phase is both unexpected and notable: the absence of Mg^{2+} in previously reported freeze-thaw cycle experiments allows the system to be driven towards the ligated state, as no cleavage reaction can occur during the thawing phase. In this study, the presence of Mg^{2+} means that the cleavage reaction is always active but is presumably counteracted by high RNA concentrations.

Discussion

In this study, we have shown that model ribozyme-peptide interactions can drastically enhance the activity of a small nucleolytic ribozyme over a wide range of conditions. In many cases, phase separation with poly-L-lysine yielded activity in conditions where the ribozyme is otherwise completely inactive (<1 mM Mg^{2+} for cleavage and <4 mM Mg^{2+} for ligation). In conditions where the ribozyme was active in solution, the addition of $(Lys)_{19–72}$ led to enhancements of up to 11-fold for 5FragF ligation (4 mM Mg^{2+}) and 65-fold for subC cleavage (1 mM Mg^{2+}). Importantly, phase separation shifted the equilibrium of the reversible transesterification reaction towards RNA ligation rather than cleavage, likely due to increased RNA concentration in the condensates. Both effects combined allow efficient and robust assembly of long RNA molecules under mild conditions in the absence of an exogenous activation chemistry. To the best of our knowledge, such a strong and rugged enhancement of ribozyme activity or shift in behavior by phase separation has not been previously reported.

Efficient recombination-based RNA assembly by fragmented HPz systems is typically only observed in dehydrating environments, for example in eutectic ice,[16] alcoholic solutions,[63] or drying.[18] The shift from cleavage to ligation induced by $(Lys)_n$ suggests that interactions with cationic peptides could have greatly broadened the scope of recombination-based RNA assembly on prebiotic Earth, perhaps providing comparable ligation yields to dehydrating environments with enhanced kinetics. Indeed, the rate of direct ligation with $(Lys)_{19–72}$ was more than ten times faster than the rate of freezing-induced ligation by a fragmented HPz ribozyme similar to that used here ($k_{lig} = 0.006\ min^{-1}$, $T = -10$ °C).[16] Poly-L-lysine-supported HPz ribozyme catalysis is robust, with similar ligation yields observed across a range of temperatures below 35 °C. Similarly, the system is able to support ligation across a broad range of magnesium concentrations, even at levels far below that typically required for catalysis under solution conditions. This is especially relevant for fragmented ribozymes, which are prebiotically appealing due to their reduced complexity, but which require higher magnesium concentrations ($[Mg^{2+}]_{1/2} = 3$ mM) for folding and activity than more complex species with additional stabilizing loops ($[Mg^{2+}]_{1/2} = 20–40$ μM).[59]

In the present study, enhancements in ribozyme activity beyond typical solution behavior were observed across a range

 © 2021 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH

of charge ratios, with the greatest enhancements occurring before the formation of liquid droplets. Fluorescence microscopy showed strong concentration and colocalization of both tagged RNAs within all separated phases (Figures S8 and S9). This likely accounts for the observed enhancements in activity, analogous to hairpin ribozyme ligation in freezing and drying environments.[18] Assays performed in the absence of peptide suggest that increased Mg^{2+} concentrations, whilst probable in this system,[40] are alone insufficient to drive RNA ligation (Figure 3). Beyond simple concentration of reactants, increased hybridization, folding and thermal stability as a result of charge interactions with the peptide may contribute to the observed enhancements of activity. Interactions between lysine-containing peptides and nucleic acids in amyloid gel fibers have previously been shown to promote the hybridization of nucleic acids below their K_D,[37] and lysine containing copolymers have been reported to enhance the rate of DNA hybridization by over 200-fold.[64] Even in the absence of phase separation, RNA polymerase ribozyme (RPR) holoenzyme assembly and activity is enhanced by interactions with simple lysine containing peptides and non-proteinaceous analogues,[42] spermidine has been shown to stabilize the HPz ribozyme-substrate complex and enhance substrate cleavage,[41] and cationic proto-peptides have been shown to increase the thermal stability of folded RNA.[65] Such enhancements in hybridization as a result of charge interactions extend beyond peptides, and are a key aspect of the proposed role of clay surfaces in early molecular evolution.[66]

It is perhaps surprising that direct condensation of $(Lys)_n$ with catalytic RNA can support (let alone enhance) activity, given that ribozyme activity can be inhibited by strong polycation-RNA interactions, which may induce misfolding.[67,45] It is also of interest that the transition to liquid coacervate droplets at high $(Lys)_n$:RNA ratios was associated with a reduction in recombination activity. Similar phase transitions have been previously reported: poly-L-lysine has been shown to form precipitates when directly mixed with double-stranded RNA, but liquid droplets with single stranded RNA.[68] We speculate that observed phase transition in the $(Lys)_n$:HPz system may originate from reduced hybridization at excess charge ratios, which could also lead to a reduction in activity.

Previous studies have demonstrated that ribozyme-catalyzed RNA cleavage is indeed supported when ribozymes are hosted within coacervate droplets.[43,44] Below the apparent ribozyme / substrate dissociation constant (K_D), concentration by coacervation can rescue hammerhead and hairpin ribozyme activity, providing relative enhancements over dilute reactions in aqueous solution.[44] However, above the K_D, hammerhead ribozyme activity in coacervate droplets is inhibited.[43] For RNA in coacervate droplets, this effect is most prominent in charge neutral conditions, or when the polycation is in excess. The activity of the hammerhead ribozyme when hosted in oligoarginine (R_{10})/ oligoaspartic acid (D_{10}) coacervate droplets is inhibited at a 1:1 charge ratio of R_{10}:D_{10}.[44] However, in coacervate compositions with excess negative charge, activity can be enhanced over concentration-limited solution conditions.[69] The suppression of activity at excess $(Lys)_n$:RNA ratios reported here may be analogous, with strong lysine-RNA interactions leading to misfolding when the peptide is in charge excess. Even so, the inhibition of activity in the coacervate phase is not total: strong recombination activity is observed in droplets at charge ratios near the transition point, suggesting that at least a portion of the RNA is able to remain folded and active. This explanation also accounts for the relative differences in activity suppression by $(Lys)_{5–24}$ and $(Lys)_{19–72}$ at excess peptide:RNA ratios, as longer polyions have a greater tendency to interact with their oppositely charged partners, and therefore may lead to greater misfolding and suppression of activity when in excess.[70–72] Peptide length dependent helicase-like activity has been demonstrated for lysine/aspartic acid coacervates, in which droplets formed from shorter oligopeptides permitted hybridization of a short RNA duplex to a greater degree than droplets comprised of longer polyions.[73] Such effects have also been reported in extant biological systems, in particular Ddx4 protein coacervates, which melt double-stranded DNA and stabilize single stranded species.[74] The suppression of activity with excess peptide may be overcome in the future by the use of heteropeptides or other homopeptides with varying charge density.

Conclusion

The observation that direct phase separation between short model peptides and ribozymes provides robust enhancements to catalytic activity under a broad range of conditions strengthens the case for both an early coevolution of RNA and peptides, and the argument that long and functional RNA emerged from pools of short oligomers via the action of small catalytic motifs. It is noteworthy that recombination is supported in this environment, as this implies that the key processes of strand release and reannealing also take place in addition to simply cleavage and ligation. Even the short poly-L-lysine used in this study, which is predominantly composed of oligomers of less than 9 residues, was able to phase separate and modulate the catalytic activity of the ribozyme. This suggests that short cationic protopeptides formed, for example, by dry-wet cycling may also be capable of such enhancements. Further investigation into the interactions between heterogeneous proto-peptides and catalytic nucleic acids is therefore of great interest. The greater enhancements provided by the longer model peptide perhaps emerged later, with the selection or synthesis of longer polycationic species.

Acknowledgements

We wish to thank Martin Spitaler, Markus Oster, Giovanni Cardone, and the MPIB Imaging Core Facility for providing excellent guidance throughout the project, as well as subsidized equipment access. We would also like to thank Maria Victoria Sanchez Caballero, Barbara Agnes Steigenberger, Evelyn Stieger and the MPIB Mass Spectrometry Core Facility for their valuable assistance in characterizing the peptides used in this work. H.M. and E.Y.S. were supported

 © 2021 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH

by funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation, Project-ID 364653263-TRR 235). H.M. gratefully acknowledges support by the European Research Council (ERC) under the Horizon 2020 research and innovation programme (grant agreement ID: 802000, RiboLife). K.L.V., B.G., T.Y.D.T. and H.M. were supported by the Volkswagen Foundation with funding from the initiative "Life?—A Fresh Scientific Approach to the Basic Principles of Life" (grant number: 92772). Open access funding enabled and organized by Projekt DEAL. Open Access funding enabled and organized by Projekt DEAL.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: catalysis · coacervates · ligases · peptides · ribozymes

[1] S. Becker, J. Feldmann, S. Wiedemann, H. Okamura, C. Schneider, K. Iwan, A. Crisp, M. Rossa, T. Amatov, T. Carell, *Science* **2019**, *366*, 76 – 82.
[2] M. Yadav, R. Kumar, R. Krishnamurthy, *Chem. Rev.* **2020**, *120*, 4766 – 4805.
[3] D. M. Fialho, T. P. Roche, N. V. Hud, *Chem. Rev.* **2020**, *120*, 4806 – 4830.
[4] W. Huang, J. P. Ferris, *J. Am. Chem. Soc.* **2006**, *128*, 8914 – 8919.
[5] J. P. Ferris, A. R. Hill, R. Liu, L. E. Orgel, *Nature* **1996**, *381*, 59 – 61.
[6] P. A. Monnard, A. Kanavarioti, D. W. Deamer, *J. Am. Chem. Soc.* **2003**, *125*, 13734 – 13740.
[7] C. Briones, M. Stich, S. C. Manrubia, *RNA* **2009**, *15*, 743 – 749.
[8] H. Mutschler, A. Wochner, P. Holliger, *Nat. Chem.* **2015**, *7*, 502 – 508.
[9] A. R. Ferré-D'Amaré, *Biopolymers* **2004**, *73*, 71 – 78.
[10] R. Shippy, R. Lockner, M. Farnsworth, A. Hampel, *Appl. Biochem. Biotechnol.* **1999**, *12*, 117 – 129.
[11] S. J. Chen, K. A. Dill, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 646 – 651.
[12] A. D. Pressman, Z. Liu, E. Janzen, C. Blanco, U. F. Müller, G. F. Joyce, R. Pascal, I. A. Chen, *J. Am. Chem. Soc.* **2019**, *141*, 6213 – 6223.
[13] E. Koculi, C. Hyeon, D. Thirumalai, S. A. Woodson, *J. Am. Chem. Soc.* **2007**, *129*, 2676 – 2682.
[14] S. J. Chen, *Annu. Rev. Biophys.* **2008**, *37*, 197 – 214.
[15] T. Pan, T. R. Sosnick, X. W. Fang, K. Littrell, B. L. Golden, V. Shelton, P. Thiyagarajan, *Proc. Natl. Acad. Sci. USA* **2002**, *98*, 4355 – 4360.
[16] A. V. Vlassov, B. H. Johnston, L. F. Landweber, S. A. Kazakov, *Nucleic Acids Res.* **2004**, *32*, 2966 – 2974.
[17] A. V. Vlassov, S. A. Kazakov, B. H. Johnston, L. F. Landweber, *J. Mol. Evol.* **2005**, *61*, 264 – 273.
[18] S. A. Kazakov, S. V. Balatskaya, B. H. Johnston, *RNA* **2006**, *12*, 446 – 456.
[19] L. A. Hegg, M. J. Fedor, *Biochemistry* **1995**, *34*, 15813 – 15828.
[20] H. Mutschler, A. I. Taylor, B. T. Porebski, A. Lightowlers, G. Houlihan, M. Abramov, P. Herdewijn, P. Holliger, *eLife* **2018**, *7*, e43022.
[21] R. Hieronymus, S. P. Godehard, D. Balke, S. Müller, *Chem. Commun.* **2016**, *52*, 4365 – 4368.
[22] R. Hieronymus, S. Müller, *ChemSystemsChem* **2021**, *3*, e2100003.
[23] S. Gwiazda, K. Salomon, B. Appel, S. Müller, *Biochimie* **2012**, *94*, 1457 – 1463.
[24] K. Adamala, J. W. Szostak, *Science* **2013**, *342*, 1098 – 1100.
[25] M. Traïkia, D. E. Warschawski, M. Recouvreur, J. Cartaud, P. F. Devaux, *Eur. Biophys. J.* **2000**, *29*, 184 – 195.
[26] R. C. MacDonald, F. D. Jones, R. Qui, *Biochim. Biophys. Acta Biomembr.* **1994**, *1191*, 362 – 370.
[27] T. Litschel, K. A. Ganzinger, T. Movinkel, M. Heymann, T. Robinson, H. Mutschler, P. Schwille, *New J. Phys.* **2018**, *20*, 055008.
[28] A. A. Hyman, C. A. Weber, F. Jülicher, *Annu. Rev. Cell Dev. Biol.* **2014**, *30*, 39 – 58.
[29] C. P. Brangwynne, C. R. Eckmann, D. S. Courson, A. Rybarska, C. Hoege, J. Gharakhani, F. Jülicher, A. A. Hyman, *Science* **2009**, *324*, 1729 – 1732.
[30] T. Ukmar-Godec, S. Hutten, M. P. Grieshop, N. Rezaei-Ghaleh, M. S. Cima-Omori, J. Biernat, E. Mandelkow, J. Söding, D. Dormann, M. Zweckstetter, *Nat. Commun.* **2019**, *10*, 2909.
[31] J. Crosby, T. Treadwell, M. Hammerton, K. Vasilakis, M. P. Crump, D. S. Williams, S. Mann, *Chem. Commun.* **2012**, *48*, 11832 – 11834.
[32] T.-Y. D. Tang, D. Van Swaay, A. DeMello, J. L. Ross Anderson, S. Mann, *Chem. Commun.* **2015**, *51*, 11429 – 11432.
[33] V. Alva, J. Söding, A. N. Lupas, *eLife* **2015**, *4*, e09410.
[34] C. Das, A. D. Frankel, *Biopolymers* **2003**, *70*, 80 – 85.
[35] A. I. Oparin, *The Origin of Life and the Origin of Enzymes*, Wiley, Hoboken, **1965**.
[36] J. Greenwald, W. Kwiatkowski, R. Riek, *J. Mol. Biol.* **2018**, *430*, 3735 – 3750.
[37] S. Braun, C. Humphreys, E. Fraser, A. Brancale, M. Bochtler, T. C. Dale, *PLoS ONE* **2011**, *6*, e19125.
[38] D. Yang, S. Peng, M. R. Hartman, T. Gupton-Campolongo, E. J. Rice, A. K. Chang, Z. Gu, G. Q. Lu, D. Luo, *Sci. Rep.* **2013**, *3*, 1 – 6.
[39] P. Adamski, M. Eleveld, A. Sood, Á. Kun, A. Szilágyi, T. Czárán, E. Szathmáry, S. Otto, *Nat. Rev. Chem.* **2020**, *4*, 386 – 403.
[40] E. A. Frankel, P. C. Bevilacqua, C. D. Keating, *Langmuir* **2016**, *32*, 2041 – 2049.
[41] B. M. Chowrira, A. Berzal-Herranz, J. M. Burke, *Biochemistry* **1993**, *32*, 1088 – 1095.
[42] S. Tagami, J. Attwater, P. Holliger, *Nat. Chem.* **2017**, *9*, 325 – 332.
[43] B. Drobot, J. M. Iglesias-Artola, K. Le Vay, V. Mayr, M. Kar, M. Kreysing, H. Mutschler, T.-Y. D. Tang, *Nat. Commun.* **2018**, *9*, 3643.
[44] R. R. Poudyal, R. M. Guth-Metzler, A. J. Veenis, E. A. Frankel, C. D. Keating, P. C. Bevilacqua, *Nat. Commun.* **2019**, *10*, 490.
[45] C. Blanco, M. Bayas, F. Yan, I. A. Chen, *Curr. Biol.* **2018**, *28*, 526 – 537.
[46] R. Hayatsu, M. H. Studier, E. Anders, *Geochim. Cosmochim. Acta* **1971**, *35*, 939 – 951.
[47] D. Yoshino, K. Hayatsu, E. Anders, *Geochim. Cosmochim. Acta* **1971**, *35*, 927 – 938.
[48] K. Plankensteiner, H. Reiner, B. M. Rode, *Mol. Diversity* **2006**, *10*, 3 – 7.
[49] B. H. Patel, C. Percivalle, D. J. Ritson, C. D. Duffy, J. D. Sutherland, *Nat. Chem.* **2015**, *7*, 301 – 307.
[50] P. Canavelli, S. Islam, M. W. Powner, *Nature* **2019**, *571*, 546 – 549.
[51] H. Hartman, T. Smith, *Life* **2014**, *4*, 227 – 249.
[52] A. P. Johnson, H. J. Cleaves, J. P. Dworkin, D. P. Glavin, A. Lazcano, J. L. Bada, *Science* **2008**, *322*, 404.
[53] D. A. M. Zaia, C. T. B. V. Zaia, H. De Santana, *Origins Life Evol. Biospheres* **2008**, *38*, 469 – 488.
[54] J. G. Forsythe, S. Yu, I. Mamajanov, M. A. Grover, R. Krishnamurthy, F. M. Fernández, N. V. Hud, *Angew. Chem. Int. Ed.* **2015**, *54*, 9871 – 9875; *Angew. Chem.* **2015**, *127*, 10009 – 10013.
[55] M. Frenkel-Pinter, J. W. Haynes, M. C, A. S. Petrov, B. T. Burcar, R. Krishnamurthy, N. V. Hud, L. J. Leman, L. D. Williams, *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 16338 – 16346.

 © 2021 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH

[56] K. Le Vay, H. Mutschler, *Emerging Top. Life Sci.* **2019**, *3*, 469 – 475.
[57] S. M. Nesbitt, H. A. Erlacher, M. J. Fedor, *J. Mol. Biol.* **1999**, *286*, 1009 – 1024.
[58] T. J. Wilson, D. M. J. Lilley, *RNA* **2002**, *8*, 587 – 600.
[59] Z. Y. Zhao, T. J. Wilson, K. Maxwell, D. M. J. Lilley, *RNA* **2000**, *6*, 1833 – 1846.
[60] W. K. Johnston, P. J. Unrau, M. S. Lawrence, M. E. Glasner, D. P. Bartel, *Science* **2001**, *292*, 1319 – 1325.
[61] J. Attwater, A. Wochner, P. Holliger, *Nat. Chem.* **2013**, *5*, 1011 – 1018.
[62] D. P. Horning, G. F. Joyce, *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 9786 – 9791.
[63] A. V. Vlassov, B. H. Johnston, S. A. Kazakov, *Oligonucleotides* **2005**, *15*, 303 – 309.
[64] L. Wu, N. Shimada, A. Kano, A. Maruyama, *Soft Matter* **2008**, *4*, 744 – 747.
[65] M. Frenkel-Pinter, J. W. Haynes, A. M. Mohyeldin, C. Martin, A. B. Sargon, A. S. Petrov, R. Krishnamurthy, N. V. Hud, L. D. Williams, L. J. Leman, *Nat. Commun.* **2020**, *11*, 1 – 14.
[66] M. Franchi, J. P. Ferris, E. Gallori, *Orig Life Evol Biosph* **2003**, *33*, 1 – 16.
[67] P. A. Chong, R. M. Vernon, J. D. Forman-Kay, *J. Mol. Biol.* **2018**, *430*, 4650 – 4665.
[68] J. R. Vieregg, M. Lueckheide, A. B. Marciel, L. Leon, A. J. Bologna, J. R. Rivera, M. V. Tirrell, *J. Am. Chem. Soc.* **2018**, *140*, 1632 – 1638.
[69] R. R. Poudyal, C. D. Keating, P. C. Bevilacqua, *ACS Chem. Biol.* **2019**, *14*, 1243 – 1248.
[70] E. Spruijt, A. H. Westphal, J. W. Borst, M. A. Cohen Stuart, J. Van Der Gucht, *Macromolecules* **2010**, *43*, 6476 – 6484.
[71] L. Li, S. Srivastava, M. Andreev, A. B. Marciel, J. J. De Pablo, M. V. Tirrell, *Macromolecules* **2018**, *51*, 2988 – 2995.
[72] J. van der Gucht, E. Spruijt, M. Lemmers, M. A. Cohen Stuart, *J. Colloid Interface Sci.* **2011**, *361*, 407 – 422.
[73] F. P. Cakmak, S. Choi, M. C. O. Meyer, P. C. Bevilacqua, C. D. Keating, *Nat. Commun.* **2020**, *11*, 5949.
[74] T. J. Nott, T. D. Craggs, A. J. Baldwin, *Nat. Chem.* **2016**, *8*, 569 – 575.

Manuscript received: July 12, 2021
Accepted manuscript online: September 27, 2021
Version of record online: November 9, 2021

 © 2021 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH

3. DISCUSSION AND CONCLUSIONS

3.1 DISCUSSION

This thesis aimed towards addressing missing gaps in our understanding of increasing complexity in prebiotic RNA-based systems. Ribonucleic acids present an attractive model for early life scenarios in part due to their ability to store genetic information, akin to the role of DNA in modern biology, and to fold into catalytically active structures, a role that is now dominated by proteins. Essential to the RNA World origin of life hypothesis that places RNA in a central genetic and functional role is a self-replicating molecule capable of Darwinian evolution. Therein, exploration of the sequence space through recombination processes is a likely means by which the realization of new catalytic functions may have emerged. However, it is unclear how primordial replicators may have accessed viable pools of activated substrates and sufficient concentrations to fuel propagative processes towards increasing length, complexity, and – eventually – function. Further, constructive processes such as ligation, recombination, and polymerization must have been supported in conjunction with the surroundings to overcome entropic barriers towards forming lifelike out-of-equilibrium polymers.

I first investigated whether diamidophosphate, a prebiotically feasible phosphate source[98,127], is capable of stably generating activated RNA substrates to drive *in situ* ribozyme-catalyzed RNA ligation. While it has been shown that frozen aqueous solutions may have provided a prebiotically plausible environment in which RNA stability and ligative activity are promoted[56,128–130], the absence of a viable activated substrate feedstock hinders momentum towards increasingly complex and functional molecules. Using a simple but robust model ribozyme system under frozen aqueous conditions, I demonstrated that the weakly activating but (pre-)biologically relevant 2',3'-cyclic phosphate (>P) nucleoside species was continually formed and fueled subsequent ligation.

In this publication, I reconciled for the first time stable RNA activation under frozen conditions with constructive, ribozyme-catalyzed RNA ligation. Remarkably, the concentrating effect of eutectic ice enables *cis*-acting ligation of a hairpin ribozyme variant from low initial solute concentrations, illustrating the capability of DAP to maintain pools of activated RNA substrate to fuel both enzymatic and non-enzymatic ligations. In contrast, conventional highly reactive laboratory approaches to generate >Ps such as carbodiimides hydrolyze in water irreversibly and rapidly. Further studies exploring routes

towards accessing the necessary energy for resultant product-ribozyme complex separation would be of interest to realize subsequent propagation of increasingly complex and larger RNA molecules towards the emergence of the first replicator. Moreover, continuous RNA evolution via persistent activation of RNA substrates could be investigated towards maintenance of multiple-ribozyme systems exhibiting cross-catalysis[57].

In the second chapter, I demonstrated experimentally that catalytic RNA-peptide interactions augment robust constructive ribozyme activity. Complex coacervation droplets are known to form with low chemical specificity for oppositely charged polymer species[131] and the resulting droplets are thought to have facilitated spatial compartmentalization of prebiotic processes[11]. However, the path from prebiotic chemistry to lifelike systems necessitates an increasing trend towards polymer length and complexity, whereas only nucleolytic reactions have been demonstrated thus far for RNA therein. Using a modified system of the ribozyme assay described in the first chapter and simple cationic peptides, I examined whether constructive RNA-catalyzed processes can be hosted within this important protocellular model.

Remarkably, I found for the first time that assembly of long RNA molecules from shorter fragments via recombination is enabled within phase separated droplets. In the absence of poly-*L*-lysine and sufficient Mg^{2+} concentrations, the hairpin ribozyme is inactive in bulk solution. Droplet formation driven by RNA-peptide interactions provides an apparent concentration effect therein to enhance both cleavage and ligation reactions under otherwise identical conditions, shifting the reaction equilibrium from cleavage towards ligation. As a result, this spatial confinement allows for robust assembly of complex ~200 nt products from <30 nt starting fragments, highlighting the potentially wide range of recombination-driven path towards increasing complexity in primordial systems. Indeed, the observed rates of ligation are comparable to those in other established concentrating environments such as in eutectic ice[132] and drying[58]. This novel finding is likely due also in part to increased folding of catalytic RNA components and hybridization of substrates to the substrate-binding regions of the ribozyme in addition to concentration of reactants.

With currently available modern technologies to probe RNA sequence characteristics within individual coacervate droplets[133], of particular interest for future experimental considerations is to reconcile complex coacervation of polyribonucleotides and peptides

with stable *in situ* RNA activation towards ribozyme networks. Here, investigating promiscuous catalytic sequences may also prove beneficial for evolving a suite of chemical reaction catalysts[134–136]. I observed through the work described here that phase-separated condensate formation and phenotypes are sensitive to proportions of opposing charges; the zwitterionic state of DAP may facilitate buffering of phase-separated droplet components to stabilize the concentrated compartments.

The concept of an RNA-centric emergence of life is widely assumed. The shortest reported synthetic ribozyme, at just 5 nt, catalyzes aminoacylation of an even shorter 4 nt RNA substrate[137], highlighting the versatility of RNA-based catalytic functions even from minimal sequences. In contrast to its protein-based cousins, the tolerated environmental range of catalytic RNA is also robust, unaffected by otherwise devastating aggregation and even benefiting from extreme conditions[67]. Molecular fossils such as the ribosome, RNA-based cofactors, and viroids, virus-like pathogenic circular RNAs capable of Darwinian evolution[138,139] also suggest continuity from early replicators to modern biology. Therefore, the origins of feedstock molecules and building blocks have been of keen interest in the field[140]. To this end, prebiotically plausible synthesis pathways of conventional RNA polymers and building blocks have been demonstrated[75,141–143], and pyrimidine and purine nucleobases alike have been detected in extraterrestrial sources such as carbonaceous meteorites[144,145].

However, starting materials were likely limited on early Earth and currently observed nucleobases, ribose, and phosphodiester linkages were likely not the original building blocks[146]. Indeed, the advent of proteins dominating the functional role in modern biology is regarded by some as a solution to the problematic instability of the polyanionic RNA backbone in the context of primordial phosphotriester formation[147]. Therefore, while RNA provides an accessible, tractable, robust, and relevant model for early life and enables streamlining experimental design, alternative model prebiotic systems also lend valuable insight.

For example, the origin of the genetic code and translation system are critical and difficult problems in the investigations on the origin of life[148,149]. Indeed, towards the pinnacle of nascent complexity in an RNA-based world, ancestral replicators may have been nucleopeptidic entities partly due to the rather straightforward path to modern biology[150]. Building on the RNA-peptide interactions that drive constructive processes re-

ported here, experimentally realizing a self-replicating system that consists of both nucleic acid and peptide where each component can catalyze replication of the other would lend further support for this notion.

Additionally, a potentially countless number of alternative primitive genetic materials have been suggested beyond conventional nucleic acids and peptides. Cyclic xeno nucleic acids, or XNAs, such as hexitol NAs are capable of duplex formation and cross-pairing with D/RNA[151], and XNA-based functional enzymes have been selected *in vitro*[152–154]. Recently, non-enzymatic phosphodiester backbone ligation of L-threoninol XNAs containing acyclic sugar moieties was demonstrated[155]. It is therefore imaginable that prebiotic soups comprised heterogeneous pools of peptides, conventional nucleic acids, and/or pre-(deoxy)ribonucleic acids, among other constituents, from which the emergence of functions not viable from homogenous chemical species was possible[156]. Further investigating molecular evolution with particular consideration for heterogeneous feedstocks, such as nonconventional genetic material, and environments, such as simulated day/night or wet/dry cycles, is critical to further our understanding of prebiotic systems[67].

3.2 CONCLUDING REMARKS

Through the work described in this thesis, a small portion of the myriad hurdles that the first RNA-based systems may have faced towards increasing complexity and function has been addressed. A prebiotically plausible scenario involving phase-separated droplets formed from short RNA and simple peptides providing compartmentalization and enhanced constructive activity as well as a viable feedstock of activated RNA substrates to feed such reactions have been described. While further propagation of complexity and subsequent coevolution of RNA and peptide towards a proto-genetic code system have yet to be demonstrated, modern technologies and insights in the field perhaps enable a not-so-distant demonstration. Further investigations into out-of-equilibrium plausible prebiotic environments allowing for efficient emergent complexity would also be of keen interest to follow.

REFERENCES

Citation style: Angewandte Chemie International Edition

[1] G. F. Joyce, D. Deamer, G. Fleischaker, in *Orig. Life Cent. Concepts*, Jones And Bartlett, Boston, **1994**.

[2] D. Prieur, G. Erauso, C. Jeanthon, *Planet. Space Sci.* **1995**, *43*, 115–122.

[3] D. B. Johnson, N. Okibe, F. F. Roberto, *Arch. Microbiol.* **2003**, *180*, 60–68.

[4] V. S. Safronov, *Isr. Progr. Sci. Transl. Jerusalem* **1972**, 11.

[5] G. W. Wetherill, G. R. Stewart, *Icarus* **1989**, *77*, 330–357.

[6] M. S. Dodd, D. Papineau, T. Grenne, J. F. Slack, M. Rittner, F. Pirajno, J. O'Neil, C. T. S. Little, *Nature* **2017**, *543*, 60–64.

[7] S. L. Miller, *Science* **1953**, *117*, 528–529.

[8] F. Tian, O. B. Toon, A. A. Pavlov, H. De Sterck, *Science* **2005**, *308*, 1014–1017.

[9] B. Cavalazzi, R. Barbieri, in *Cnidaria, Past, Present Future. World Medusa Her Sisters* (Eds.: S. Goffredo, Z. Dubinsky), Springer, **2016**, pp. 3–13.

[10] F. Wöhler, *Ann. Phys.* **1828**, *88*, 253–256.

[11] A. I. Oparin, in *Orig. Life* (Ed.: J.D. Bernal), Weidenfeld And Nicholson, **1967**, pp. 199–234.

[12] J. B. S. Haldane, *Ration. Annu.* **1929**, *148*, 3–10.

[13] J. R. Cronin, S. Pizzarello, *Science* **1997**, *275*, 951–955.

[14] G. Wächtershäuser, *Prog. Biophys. Mol. Biol.* **1992**, *58*, 85–201.

[15] J. D. Bernal, *Proc. Phys. Soc. Sect. A* **1949**, *62*, 537.

[16] A. Rich, M. Kasha, B. Pullman, in *Horizons Biochem.* (Eds.: M. Kasha, B. Pullman), Academic Press, New York, **1962**.

[17] C. R. Woese, D. H. Dugre, S. A. Dugre, M. Kondo, W. C. Saxinger, *Cold Spring Harb. Symp. Quant. Biol.* **1966**, *31*, 723–736.

[18] F. H. C. Crick, *J. Mol. Biol.* **1968**, *38*, 367–379.

[19] L. E. Orgel, *J. Mol. Biol.* **1968**, *38*, 381–393.

[20] Walter Gilbert, *Nature* **1986**, *319*, 618–618.

[21] G. F. Joyce, J. W. Szostak, *Cold Spring Harb. Perspect. Biol.* **2018**, *10:a034801*

[22] A. M. Weiner, N. Maizels, *Proc. Natl. Acad. Sci. U. S. A.* **1987**, *84*, 7383.

[23] N. Maizels, A. M. Weiner, *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 6729–6734.

[24] A. D. Goldman, B. Kacar, *J. Mol. Evol.* **2021**, *89*, 127–133.

[25] T. R. Cech, A. J. Zaug, P. J. Grabowski, *Cell* **1981**, *27*, 487–496.

[26] C. Guerrier-Takada, K. Gardiner, T. Marsh, N. Pace, S. Altman, *Cell* **1983**, *35*, 849–857.

[27] N. Ban, P. Nissen, J. Hansen, P. B. Moore, T. A. Steitz, *Science* **2000**, *289*, 905–920.

[28] B. T. Wimberly, D. E. Brodersen, W. M. Clemons, R. J. Morgan-Warren, A. P. Carter, C. Vonrheln, T. Hartsch, V. Ramakrishnan, *Nature* **2000**, *407*, 327–339.

[29] M. M. Yusupov, G. Z. Yusupova, A. Baucom, K. Lieberman, T. N. Earnest, J. H. D. Cate, H. F. Noller, *Science* **2001**, *292*, 883–896.

[30] K. Plankensteiner, H. Reiner, B. Rode, *Curr. Org. Chem.* **2005**, *9*, 1107–1114.

[31] D. Segré, D. Ben-Eli, D. W. Deamer, D. Lancet, *Orig. Life Evol. Biosph.* **2001**, *31*, 119–145.

[32] M. P. Robertson, G. F. Joyce, *Cold Spring Harb. Perspect. Biol.* **2012**, 4:a003608

[33] H. Duim, S. Otto, *Beilstein J. Org. Chem.* **2017**, *13*, 1189.

[34] K. Le Vay, H. Mutschler, *Emerg. Top. Life Sci.* **2019**, *3*, 469–475.

[35] C. J. Hutchins, P. D. Rathjen, A. C. Forster, R. H. Symons, *Nucleic Acids Res.* **1986**, *14*, 3627.

[36] M. J. Fedor, *J. Mol. Biol.* **2000**, *297*, 269–291.

[37] R. R. Breaker, *Chem. Rev.* **1997**, *97*, 371–390.

[38] H. Ulrich, C. Trujillo, A. Nery, J. Alves, P. Majumder, R. Resende, A. Martins, *Comb. Chem. High Throughput Screen.* **2006**, *9*, 619–632.

[39] Y. Nomura, Y. Yokobayashi, *Nucleic Acids Res.* **2019**, *47*, 8950.

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

[41] C. A. Chay, X. Guan, G. Bruening, *Virology* **1997**, *239*, 413–425.

[42] J. M. Carothers, S. C. Oestreich, J. H. Davis, J. W. Szostak, *J. Am. Chem. Soc.* **2004**, *126*, 5130–5137.

[43] T. R. Cech, *Gene* **1993**, *135*, 33–36.

[44] I. Hirao, A. D. Ellington, *Curr. Biol.* **1995**, *5*, 1017–1022.

[45] A. R. Ferré-D'Amaré, *Biopolymers* **2004**, *73*, 71–78.

[46] J. Haseloff, W. L. Gerlach, *RNA Catal. Splicing, Evol.* **1989**, 43–52.

[47] A. Hampel, R. Tritz, *Biochemistry* **2002**, *28*, 4929–4933.

[48] Y. Komatsu, I. Kanzaki, M. Koizumi, E. Ohtsuka, *J. Mol. Biol.* **1995**, *252*, 296–304.

[49] A. Hampel, J. A. Cowan, *Chem. Biol.* **1997**, *4*, 513–517.

[50] S. Nesbitt, L. A. Hegg, M. J. Fedor, *Chem. Biol.* **1997**, *4*, 619–630.

[51] K. J. Young, F. Gill, J. A. Grasby, *Nucleic Acids Res.* **1997**, *25*, 3760–3766.

[52] A. Hampel, R. Tritz, *Biochemistry* **1989**, *28*, 4929–4933.

[53] S. M. Nesbitt, H. A. Erlacher, M. J. Fedor, *J. Mol. Biol.* **1999**, *286*, 1009–1024.

[54] L. A. Hegg, M. J. Fedor, *Biochemistry* **2002**, *34*, 15813–15828.

[55] I. Gözen, E. S. Köksal, I. Põldsalu, L. Xue, K. Spustova, E. Pedrueza-Villalmanzo, R. Ryskulov, F. Meng, A. Jesorka, I. Gözen, E. S. Köksal, I. Põldsalu, L. Xue, K. Spustova, E. Pedrueza-Villalmanzo, R. Ryskulov, F. Meng, A. Jesorka, *Small* **2022**, 2106624.

[56] P. A. Monnard, A. Kanavarioti, D. W. Deamer, *J. Am. Chem. Soc.* **2003**, *125*,

13734–13740.

[57] H. Mutschler, A. Wochner, P. Holliger, *Nat. Chem.* **2015**, *7*, 502–508.

[58] S. A. Kazakov, S. V. Balatskaya, B. H. Johnston, *RNA* **2006**, *12*, 446–456.

[59] A. Lazcano, S. L. Miller, *Cell* **1996**, *85*, 793–798.

[60] M. Levy, S. L. Miller, J. Oró, *J. Mol. Evol.* **1999**, *49*, 165–168.

[61] J. L. Bada, A. Lazcano, *Science* **2002**, *296*, 1982–1983.

[62] J. W. Szostak, *J. Syst. Chem.* **2012**, *3*, 1–14.

[63] H. S. Bernhardt, *Biol. Direct* **2012**, *7*, 1–10.

[64] D. Segré, D. Ben-Eli, D. W. Deamer, D. Lancet, *Orig. life Evol. Biosph.* **2001**, *31*, 119–145.

[65] A. Pressman, C. Blanco, I. A. Chen, *Curr. Biol.* **2015**, *25*, R953–R963.

[66] G. F. Joyce, Jones And Bartlett Publishers, Boston, **1994**.

[67] K. Le Vay, E. Salibi, E. Y. Song, H. Mutschler, *Chem. – An Asian J.* **2020**, *15*, 214–230.

[68] N. R. Pace, *Cell* **1991**, *65*, 531–533.

[69] M. Levy, S. L. Miller, *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 7933–7938.

[70] F. H. Westheimer, *Science* **1987**, *235*, 1173–1178.

[71] M. J. Russell, A. J. Hall, *J. Geol. Soc. London.* **1997**, *154*, 377–402.

[72] W. Martin, M. J. Russell, *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **2007**, *362*, 1887–1925.

[73] W. Martin, J. Baross, D. Kelley, M. J. Russell, *Nat. Rev. Microbiol.* **2008**, *6*, 805–814.

[74] H. S. Bernhardt, W. P. Tate, *Biol. Direct* **2012**, *7*, DOI 10.1186/1745-6150-7-4.

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

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

[77] B. J. Weimann, R. Lohrmann, L. E. Orgel, H. Schneider-Bernloehr, J. E. Sulston, *Science* **1968**, *161*, 387.

[78] G. F. Joyce, T. Inoue, L. E. Orgel, *J. Mol. Biol.* **1984**, *176*, 279–306.

[79] S. R. Vogel, C. Deck, C. Richert, *Chem. Commun. (Camb).* **2005**, 4922–4924.

[80] W. G. Scott, A. Szöke, J. Blaustein, S. M. O'Rourke, M. P. Robertson, *Life Open Access J.* **2014**, *4*, 131.

[81] M. W. Powner, B. Gerland, J. D. Sutherland, *Nature* **2009**, *459*, 239–242.

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

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

[84] M. Crowe, J. S.- ChemBioChem, undefined 2006, *Wiley Online Libr.* **2006**, *7*, 951–956.

[85] M. Shigematsu, T. Kawamura, Y. Kirino, *Front. Genet.* **2018**, *9*, 562.

[86] S. Mohr, R. E. Thach, *J. Biol. Chem.* **1969**, *244*, 6566–6576.

[87] M. Renz, R. Lohrmann, *BBA - N. A.* **1971**, *240*, 463-471.

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

[89] H. Mutschler, P. Holliger, *J. Am. Chem. Soc.* **2014**, *136*, 5193–5196.

[90] F. H. Westheimer, *Acc. Chem. Res.* **1968**, *1*, 70–78.

[91] J. A. Rojas Stütz, E. Kervio, C. Deck, C. Richert, *Chem. Biodivers.* **2007**, *4*, 784–802.

[92] M. A. Gilles, A. Q. Hudson, C. L. Borders, *Anal. Biochem.* **1990**, *184*, 244–248.

[93] W. R. Algar, in *Chemoselective Bioorthogonal Ligation React. Concepts Appl.* (Eds.: W.R. Algar, P.E. Dawson, I.L. Medintz), Wiley-VCH Verlag GmbH & Co. KGaA, **2017**, pp. 3–36.

[94] A. Osumah, R. Krishnamurthy, *ChemBioChem* **2021**, *22*, 3001–3009.

[95] C. Gibard, I. B. Gorrell, E. I. Jiménez, T. P. Kee, M. A. Pasek, R. Krishnamurthy, *Angew. Chemie - Int. Ed.* **2019**, *58*, 8151–8155.

[96] M. A. Pasek, T. P. Kee, in *Orig. Life Primal Self-Organization* (Eds.: R. Egel, D.-H. Lankenau, Y. A. Mulkidjanian), Springer Berlin Heidelberg, Berlin, **2011**, pp. 57–84.

[97] M. Gull, *Challenges 2014, Vol. 5, Pages 193-212* **2014**, *5*, 193–212.

[98] C. Gibard, S. Bhowmik, M. Karki, E. K. Kim, R. Krishnamurthy, *Nat. Chem.* **2018**, *10*, 212–217.

[99] J. C. Blain, J. W. Szostak, *Ann. Rev. Biochem.* **2014**, *83*, 615–640.

[100] Y. Bansho, T. Furubayashi, N. Ichihashi, T. Yomo, *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 4045–4050.

[101] S. Matsumura, Á. Kun, M. Ryckelynck, F. Coldren, A. Szilágyi, F. Jossinet, C. Rick, P. Nghe, E. Szathmáry, A. D. Griffiths, *Science* **2016**, *354*, 1293–1296.

[102] K. Adamala, J. W. Szostak, *Science* **2013**, *342*, 1098–1100.

[103] N. Lane, W. F. Martin, *Cell* **2012**, *151*, 1406–1416.

[104] P. Baaske, F. M. Weinert, S. Duhr, K. H. Lemke, M. J. Russell, D. Braun, *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 9346–9351.

[105] C. B. Mast, S. Schink, U. Gerland, D. Braun, *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 8030–8035.

[106] M. Kreysing, L. Keil, S. Lanzmich, D. Braun, *Nat. Chem.* **2015**, *7*, 203–208.

[107] J. P. Ferris, *Philos. Trans. R. Soc. B Biol. Sci.* **2006**, *361*, 1777–1786.

[108] N. Ichihashi, K. Usui, Y. Kazuta, T. Sunami, T. Matsuura, T. Yomo, *Nat. Commun.* **2013**, *4*:2494

[109] F. Hong, F. Zhang, Y. Liu, H. Yan, *Chem. Rev.* **2017**, *117*, 12584–12640.

[110] W. M. Shih, J. D. Quispe, G. F. Joyce, *Nature* **2004**, *427*, 618–621.

[111] R. Schulman, B. Yurke, E. Winfree, *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 6405–6410.

[112] K. Adamala, J. W. Szostak, *Science* **2013**, *342*, 1098–1100.

[113] A. Veis, *Adv. Colloid Interface Sci.* **2011**, *167*, 2–11.

[114] S. L. Perry, Y. Li, D. Priftis, L. Leon, M. Tirrell, *Polymers (Basel).* **2014**, *6*, 1756–1772.

[115] D. Priftis, M. Tirrell, *Soft Matter* **2012**, *8*, 9396–9405.

[116] S. Koga, D. S. Williams, A. W. Perriman, S. Mann, *Nat. Chem.* **2011**, *3*, 720–724.

[117] R. Chollakup, W. Smitthipong, C. D. Eisenbach, M. Tirrell, *Macromolecules* **2010**, *43*, 2518–2528.

[118] T. Y. Dora Tang, C. Rohaida Che Hak, A. J. Thompson, M. K. Kuimova, D. S. Williams, A. W. Perriman, S. Mann, *Nat. Chem.* **2014**, *6*, 527–533.

[119] W. M. Aumiller, F. P. Cakmak, B. W. Davis, C. D. Keating, *Langmuir* **2016**, *32*, 10042-10053.

[120] Y. G. Zhao, H. Zhang, *Dev. Cell* **2020**, *55*, 30–44.

[121] S. Boeynaems, S. Alberti, N. L. Fawzi, T. Mittag, M. Polymenidou, F. Rousseau, J. Schymkowitz, J. Shorter, B. Wolozin, L. Van Den Bosch, P. Tompa, M. Fuxreiter, *Trends Cell Biol.* **2018**, *28*, 420–435.

[122] P. Cramer, D. A. Bushnell, R. D. Kornberg, *Science* **2001**, *292*, 1863–1876.

[123] G. Zhang, E. A. Campbell, L. Minakhin, C. Richter, K. Severinov, S. A. Darst, *Cell* **1999**, *98*, 811–824.

[124] I. Suárez-Marina, Y. M. Abul-Haija, R. Turk-MacLeod, P. S. Gromski, G. J. T. Cooper, A. O. Olivé, S. Colón-Santos, L. Cronin, *Commun. Chem.* **2019**, *2*, 1–8.

[125] B. Drobot, J. M. Iglesias-Artola, K. Le Vay, V. Mayr, M. Kar, M. Kreysing, H. Mutschler, T. Y. D. Tang, *Nat. Commun.* **2018**, *9*, 1–9.

[126] R. R. Poudyal, R. M. Guth-Metzler, A. J. Veenis, E. A. Frankel, C. D. Keating, P. C. Bevilacqua, *Nat. Commun.* **2019**, *10*, 1–13.

[127] M. Karki, C. Gibard, S. Bhowmik, R. Krishnamurthy, *Life* **2017**, *7*, 32.

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

[129] P. A. Monnard, J. W. Szostak, *J. Inorg. Biochem.* **2008**, *102*, 1104–1111.

[130] P. M. G. Löffler, J. Groen, M. Dörr, P. A. Monnard, *PLoS One* **2013**, *8*:e75617.

[131] K. K. Nakashima, M. A. Vibhute, E. Spruijt, *Front. Mol. Biosci.* **2019**, *6*:21.

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

[133] D. Wollny, B. Vernot, J. Wang, M. Hondele, A. Hyman, K. Weis, J. G. Camp, T.-Y. D. Tang, B. Treutlein, *bioRxiv* **2022,** *13*:2626.

[134] E. A. Schultes, D. P. Bartel, *Science* **2000**, *289*, 448–452.

[135] O. Khersonsky, D. S. Tawfik, *Annu. Rev. Biochem.* **2010**, *79*, 471–505.

[136] E. Janzen, E. Janzen, C. Blanco, H. Peng, J. Kenchel, J. Kenchel, I. A. Chen, I. A. Chen, I. A. Chen, *Chem. Rev.* **2020**, *120*, 4879–4897.

[137] R. M. Turk, N. V. Chumachenko, M. Yarus, *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 4585–4589.

[138] R. Flores, S. Gago-Zachert, P. Serra, R. Sanjuán, S. F. Elena, *Ann. Rev. Microbiol.* **2014**, *68*, 395–414.

[139] K. Moelling, F. Broecker, *Int. J. Mol. Sci.* **2021**, *22*, 3476.

[140] Z. Martins, *Life* **2018**, *8*, 28.

[141] B. H. Patel, C. Percivalle, D. J. Ritson, C. D. Duffy, J. D. Sutherland, *Nat. Chem.* **2015**, *7*, 301–307.

[142] N. Prywes, J. C. Blain, F. Del Frate, J. W. Szostak, *Elife* **2016**, *5*:e17756.

[143] A. Biscans, *Life (Basel, Switzerland)* **2018**, *8*, 57.

[144] R. Hayatsu, *Science* **1964**, *146*, 1291–1293.

[145] Y. Oba, Y. Takano, Y. Furukawa, T. Koga, D. P. Glavin, J. P. Dworkin, H. Naraoka, *Nat. Commun.* **2022**, *13*, 1–10.

[146] N. V. Hud, B. J. Cafferty, R. Krishnamurthy, L. D. Williams, *Chem. Biol.* **2013**, *20*, 466–474.

[147] A. C. Forster, *Biochemistry* **2022**, *61*, 749-751.

[148] F. H. C. Crick, S. Brenner, A. Klug, G. Pieczenik, *Orig. Life* **1976**, *7*, 389–397.

[149] H. Grosjean, E. Westhof, *Nucleic Acids Res.* **2016**, *44*, 8020–8040.

[150] B. M. A. G. Piette, J. G. Heddle, *Trends Ecol. Evol.* **2020**, *35*, 397–406.

[151] C. Hendrix, H. Rosemeyer, I. Verheggen, F. Seela, A. Van Aerschot, P. Herdewijn, *Chem. – A Eur. J.* **1997**, *3*, 110–120.

[152] V. B. Pinheiro, A. I. Taylor, C. Cozens, M. Abramov, M. Renders, S. Zhang, J. C. Chaput, J. Wengel, S. Y. Peak-Chew, S. H. McLaughlin, P. Herdewijn, P. Holliger, *Science* **2012**, *336*, 341–344.

[153] A. I. Taylor, V. B. Pinheiro, M. J. Smola, A. S. Morgunov, S. Peak-Chew, C. Cozens, K. M. Weeks, P. Herdewijn, P. Holliger, *Nature* **2015**, *518*, 427–430.

[154] H. Mei, J. Y. Liao, R. M. Jimenez, Y. Wang, S. Bala, C. McCloskey, C. Switzer, J. C. Chaput, *J. Am. Chem. Soc.* **2018**, *140*, 5706–5713.

[155] K. Murayama, H. Okita, T. Kuriki, H. Asanuma, *Nat. Commun. 2021 121* **2021**, *12*, 1–9.

[156] S. Bhowmik, R. Krishnamurthy, *Nat. Chem.* **2019**, *11*, 1009–1018.

APPENDIX A: PUBLICATION III

4.1 PUBLICATION III

Nucleic Acid Catalysis under Potential Prebiotic Conditions

AUTHORS: Dr. Kristian Le Vay*, Elia Salibi[+], **Emilie Yeonwha Song**[+], and Prof. Dr. Hannes Mutschler*

[+] *authors contributed equally.* * *corresponding authors.*

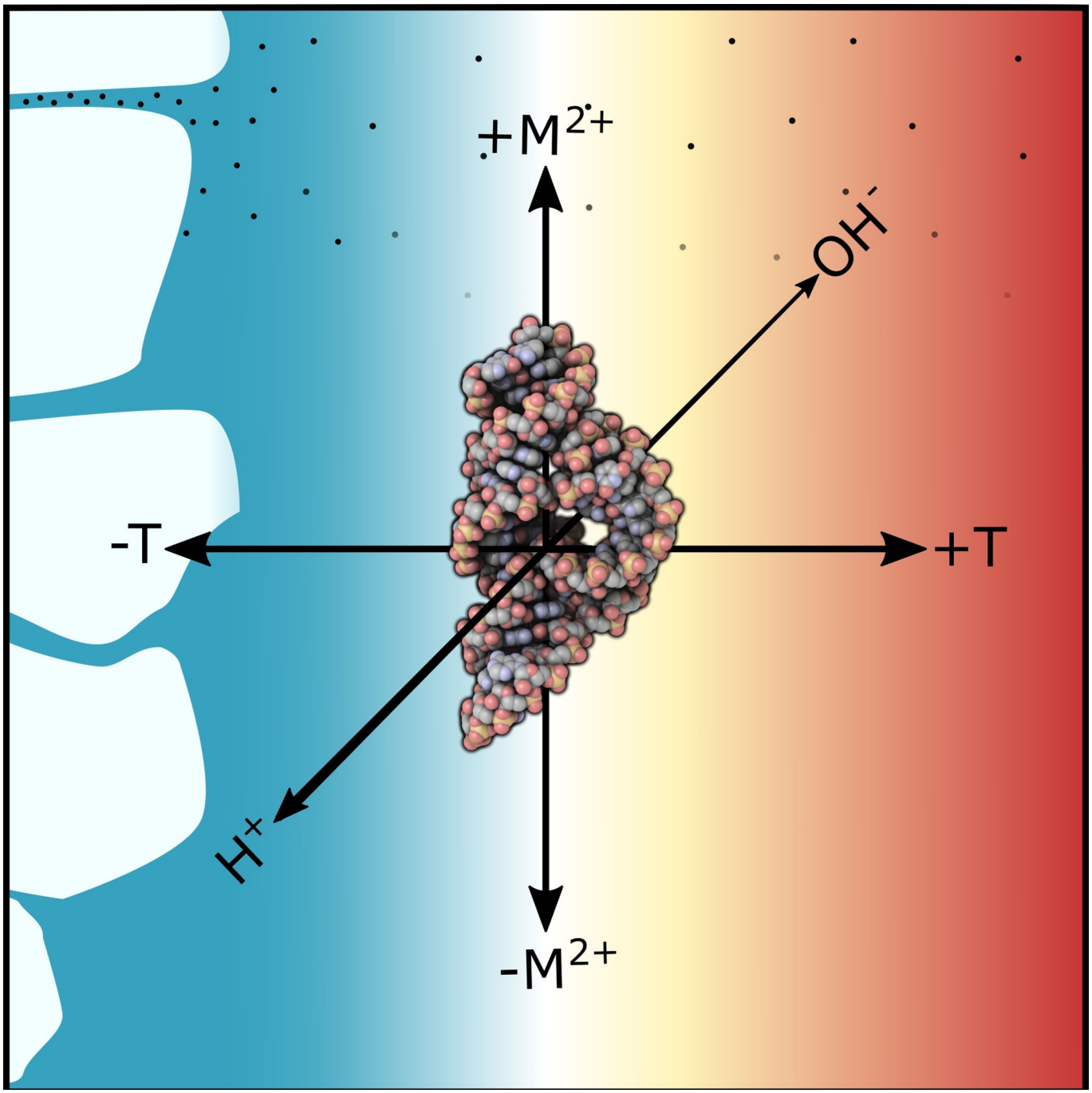
JOURNAL, YEAR, VOLUME, AND PAGE NO.:

Chem. Asian. J., **2020**, *15*, 214-230

ACES Asian Chemical Editorial Society
DOI: 10.1002/asia.201901205
CHEMISTRY AN ASIAN JOURNAL Minireview

Nucleic Acid Catalysis under Potential Prebiotic Conditions

Kristian Le Vay,* Elia Salibi[+], Emilie Y. Song[+], and Hannes Mutschler*[a]



© 2019 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

Abstract: Catalysis by nucleic acids is indispensable for extant cellular life, and it is widely accepted that nucleic acid enzymes were crucial for the emergence of primitive life 3.5-4 billion years ago. However, geochemical conditions on early Earth must have differed greatly from the constant internal milieus of today's cells. In order to explore plausible scenarios for early molecular evolution, it is therefore essential to understand how different physicochemical parameters, such as temperature, pH, and ionic composition, influence nucleic acid catalysis and to explore to what extent nucleic acid enzymes can adapt to non-physiological conditions. In this article, we give an overview of the research on catalysis of nucleic acids, in particular catalytic RNAs (ribozymes) and DNAs (deoxyribozymes), under extreme and/or unusual conditions that may relate to prebiotic environments.

1. Introduction

The discovery of the catalytic properties of nucleic acids by Cech and Altman in 1982-83 both redefined biological catalysis and provided compelling support for origin of life hypotheses centered around nucleic acid-based information storage and catalysis, in particular the "RNA world" hypothesis first suggested by Alexander Rich, in which self-replicating RNA emerged prior to the evolution of DNA and proteins.[1–3] Despite the prevalence of the RNA World hypothesis and related conjectures, such as different "pre-RNA" worlds[4] and mixed chimeric systems including, for example, both RNA and DNA,[5] a key unanswered question is: under which environmental conditions did functional nucleic acids emerge and sustain themselves? Constraining the parameter space of a habitable early Earth is crucial to understanding the emergence of life. One way of achieving this is to consider the sensitivity of nucleic acids to environmental conditions: in what conditions can nucleic acids survive, and do conditions exist which can potentiate nucleic acid catalysis? Exploring conditions more exotic than dilute buffered solutions may yield answers to intractable problems in origin of life and synthetic biology research.[6,7]

A wide range of catalytic nucleic acids are known today. For RNA (ribozymes), the most iconic example is the ribosome,[8] whose central role in peptide bond formation and thus protein synthesis designates it the most important ribozyme in modern biochemistry, and the most obvious "smoking gun" of an early RNA world predating modern biochemistry. Another ubiquitous ribozyme that is essential in all free-living organisms is RNAseP, which processes the 5′-ends of precursor-tRNAs.[9,10] Other prominent examples for ribozymes are small RNA-cleaving ribozymes such as the hammerhead (HH) ribozyme[11,12] (Figure 1 A) and the hairpin (HP) ribozyme[13] (Figure 1 B), which catalyze reversible self-cleavage to process the concatemeric products of rolling circle RNA replication into linear and circular RNA molecules.[14] A related function is carried out by self-splicing introns,[15,16] which catalyze their own excision from messenger, transfer, or ribosomal RNA via two sequential transesterification reactions of the phosphodiester backbone. In addition, in vitro selection experiments have revealed that the palette of RNA catalysis is far broader than these reactions and encompasses RNA ligation,[17,18] aminoacyl transfer, porphyrin metalation[19] and C–C bond formation including the Diels–Alder reaction,[20] Michael addition,[21] aldol condensations[22] and others,[23] suggesting that an early metabolism might have been sustained by ribozymes.

While the main function of DNA in biology is the storage of genetic information, a large number of artificial DNA catalysts have also been isolated by in vitro selection. These deoxyribozymes, or DNAzymes, catalyze a range of bond forming reactions, including the Diels–Alder reaction,[24] Friedel–Crafts reactions,[25] RNA ligation (2′-5′ and 3′-5′),[26,27] DNA ligation,[28] 5′-phosphorylation,[29] adenylation,[30] RNA-nucleopeptide linkage[31] and porphyrin metalation.[32] The full range of DNA catalysis is reviewed in detail by Hollenstein, and an example of a RNA cleaving DNAzyme is shown in Figure 1 D.[33]

Finally, synthetic nucleic acids are also capable of catalysis. In particular, Taylor et al. selected artificial endonuclease and ligase enzymes from random pools of arabino nucleic acid (ANA), 2′-fluoroarabino nucleic acid (FANA), hexitol nucleic acid (HNA) and cyclohexene nucleic acid (CeNA).[34]

While these studies convincingly demonstrate the broad catalytic potential of polynucleotides, they leave open the question of whether some of these reactions could have contributed to early biocatalysis, and whether they are compatible with the environmental conditions on early Earth.

Since the beginning of the Origin of Life field, great efforts have been made to determine, or at least constrain, the conditions under which life originated. Definitive answers have been elusive, due to the extreme timescales under consideration and the combined uncertainties of when, where and how the first primitive forms of life emerged. The lack of fossil evidence of early life, the large number of possible geochemical environments and the difficulty in determining conditions on early Earth make this an almost intractable problem for origin of life researchers, amongst whom there is little consensus on these

[a] *Dr. K. Le Vay, E. Salibi,[+] E. Y. Song,[+] Dr. H. Mutschler*
Biomimetic Systems
Max Planck Institute of Biochemistry
Am Klopferspitz 18, 82152 Martinsried (Germany)
E-mail: levay@biochem.mpg.de
mutschler@biochem.mpg.de

[+] *These authors contributed equally to this work.*

The ORCID identification number(s) for the author(s) of this article can be found under:
https://doi.org/10.1002/asia.201901205.

© 2019 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

 © 2019 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

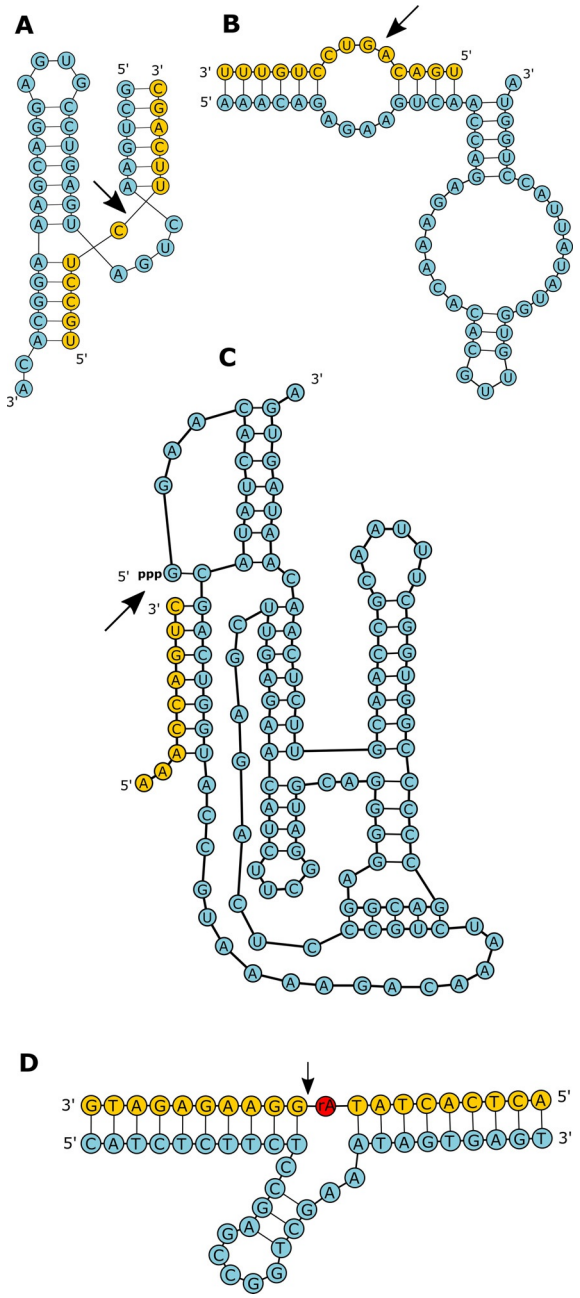


Figure 1. Secondary structures of various nucleic acid enzymes, including the hammerhead ribozyme, hairpin ribozyme, the class I ligase and 8–17 DNAzyme. The hammerhead (A) and hairpin (B) ribozymes catalyze the reversible cleavage of the RNA substrate strand shown in yellow (black arrow indicates cleavage site).[42] The class I ligase (C) binds a substrate strand (yellow) and catalyzes 3′ OH nucleophilic attack on its own 5′ triphosphate, leading to phosphodiester bond formation and release of inorganic pyrophosphate.[43] The 8–17 DNAzyme (D) is a metalloenzyme catalyzing RNA transesterification in the presence of divalent metal ions.[44] The substrate strand is shown in yellow, with the ribonucleotide cleavage site marked in red.

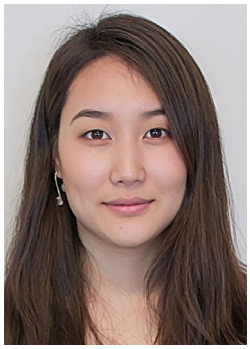
questions.[35,36] In light of this, we and others have previously argued for a flexible approach to the problem, by performing experiments under relaxed but plausible boundary conditions and using the results to inform about possible plausible prebiotic environments.[37–41]

The many studies that aim to constrain the global climate and conditions on early Earth allow some experimental boundaries to be set: As today, divalent magnesium and calcium were abundant in the oceans of early Earth. Historical ocean solute composition is dependent on both pH and reducing potential. Assuming an acidic ocean pH around 4 Ga, hydrogen

Hannes Mutschler *received his M.Sc. in biophysics from the Humboldt University of Berlin. His Ph.D. topic was on bacterial toxin-antitoxin systems at the Max Planck Institute of Medical Research/Heidelberg University, Germany. For his postdoc on origins of life research, he joined Phil Holliger's lab at the Protein and Nucleic Acid Chemistry Division lab at the MRC Laboratory of Molecular Biology in Cambridge, UK. In 2016, he became a Max-SynBio independent research group leader at Max Planck Institute of Biochemistry in Martinsried, Germany. His current research interests include bottom-up synthetic biology, cell-free protein synthesis as well as RNA catalysis under prebiotic conditions.*

Kristian Le Vay *received his M.Sc. in chemistry from the University of Bristol, UK, and completed his Ph.D. at the Bristol Centre for Functional Nanomaterials. During this time, he designed and developed artificial enzyme systems based on photoactive protein-nanoparticle conjugates. He is currently a postdoctoral researcher at the Max Planck Institute for Biochemistry. Here, his research encompasses both synthetic biology and origin of life studies, focusing on prebiotic RNA catalysis in unusual environments and the development of dynamic model protocell systems.*

Elia Salibi *completed his B.Sc. in Biology at the University of Beirut, Lebanon. He then received his M.Sc. in Molecular Biology and Genetics from the University of Pavia, Italy. His thesis focused on developing and characterizing a synthetic modular CRISPR/dCas9-based gene silencing toolbox for Escherichia coli. Currently he is a Ph.D. student at the Max Planck Institute of Biochemistry investigating the role of RNA protocells in the context of the emergence of life.*



Emilie Y. Song *received her B.Sc. in Biochemistry and Chemistry from the University of British Columbia in Vancouver, Canada. She then completed her M.Sc. in Molecular Genetics at the University of Toronto screening for novel Streptomyces bacterial natural products with potential bioactivity against parasitic worms. She is currently a PhD candidate in the CRC235: Emergence of Life graduate program investigating RNA stability and catalysis in plausible prebiotic environments.*

© 2019 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

sulfide present in seawater would have created a reducing environment rich in Fe^{2+}, but low in concentrations of free transition metal and group 12–16 ions due to the formation of insoluble sulfide compounds.[45,46] Early nucleic acid catalysis may have relied on Fe^{2+} as a cofactor, until the advent of aerobic conditions caused the oxidation of Fe^{2+} to Fe^{3+}, necessitating its replacement by Mg^{2+} or other metal ions.[47] Oceanic pH, which is driven by atmospheric CO_2 concentrations, likely rose monotonically from pH 6.6 in the Hadean era to pH 7.9 by the Cambrian era.[48] However, other studies posit that oceanic pH in the late Hadean/ early Archean was as low as 3.5–5.4.[49,50] Further uncertainty is introduced if we consider that life may have emerged in the vicinity of a hydrothermal vents, where local pH may be either very low (pH 2–3) or very high (pH 9–11), depending on type, rather than in the bulk ocean.[51]

Estimates of temperature are more variable, spanning climates ranging from frozen to near boiling. Oxygen, iron and silicon isotope studies suggest temperatures of 70 °C up until as late as 3.3 Ga, a theory additionally supported by evidence of a low viscosity Archean ocean.[52–55] However, evidence of a temperate climate is provided by geological carbon cycle models and isotope evidence from cherts and sediments.[56–58] Studies of Archean glacial deposits suggest the presence of ice caps or cold periods during this time,[58] and some researchers argue that in the absence of extreme levels of greenhouse gases, a glacial Hadean Earth is likely, albeit with intermittent periods of "fire and brimstone" following major impacts.[59,60]

Although these studies provide some useful constraints on the conditions at the Origin of Life, a broad range of conditions remain feasible. The exact microenvironment in which the first replicators emerged was likely more significant than the global conditions at the time. For example, 'warm little ponds' on land would be subject to temperature, composition and concentration fluctuations due to evaporation and condensation driven by day–night cycles,[61] eutectic phases in frozen environments lead to strong solute up-concentration and significant pH shifts,[62] and hydrothermal vents provide extreme temperature and pH gradients.[51] Any of these environments might provide shelter from adverse conditions such as UV radiation, the surface intensity of which was several orders of magnitude higher than today.[63]

In this focus review, we will explore the range of conditions under which nucleic acid catalysis is possible, highlighting how nucleic acids can adapt to extreme conditions, and how these conditions can both support and potentiate function. In order to understand the emergence of life, we must understand the environmental factors that would have acted upon the first functional nucleic acids, for example, in an RNA, proto-RNA or mixed nucleic acid world scenario. In addition, many nucleic acid enzymes catalyze industrially relevant processes and, as such, challenging conditions may be required to increase reaction rates, shift reaction equilibria or improve substrate or product solubility. In both cases, reaction conditions may deviate strongly from in vivo or typical in vitro environments.

2. The role of metal ions in nucleic acid folding and catalysis

2.1. Folding of nucleic acids

The range of conditions in which catalytic nucleic acids are functional is largely determined by the mechanism by which nucleic acids can fold into catalytically active three-dimensional structures. Nucleic acid folding differs to that of proteins, which in many cases tend to fold via rapid, cooperative two-state thermodynamic transitions, with no detectable intermediate structures.[64] Nucleic acid chain compaction is driven by ion-mediated electrostatic interaction, conformational entropy, base pairing, base stacking, and noncanonical interactions.[65,66] Compared to proteins, the folding energy landscape of nucleic acids is convoluted due to the high number of competing, energetically similar folding states, and nucleic acid molecules tend to adopt a range of conformations in solution.[67,68] The highly charged polyanionic backbone of nucleic acids usually prevents the irreversible aggregation of misfolded molecules. This means that, whilst activity may be lowered by adverse environmental conditions due to the presence of inactive or poorly active conformers, catalysis can occur under a broad range of environmental conditions. Consequently, conditions that promote folding and the formation of active conformations are of particular interest, as they may directly improve the catalytic activity of nucleic acid enzymes.

2.1. Modes of metal ion—nucleic acid interaction

A key variable determining nucleic acid folding and activity is the presence of counterions, which help to overcome the charge repulsion from the polyphosphate backbone during compaction. For RNA, the most relevant cations under in vivo conditions are Mg^{2+} and K^+, both of which interact with RNA predominantly through electrostatic forces.[69] In particular, Mg^{2+} ions enable the formation of complex folds that allow nucleic acids to stabilize specific structures, recognize binding partners and mediate catalytic processes.[70–73] Generally, interacting Mg^{2+} can be divided into two populations (Figure 2): diffusive ions, which surround the RNA as an ensemble of hydrated ions that are non-specifically attracted to the negative charge of the RNA, and a much smaller group of partially desolvated ions, which bind to specific electronegative sites on the RNA itself.[74] Whilst these specific metal ion-RNA interactions mostly contribute to the conformational specificity of an RNA structure (and thus in many cases to the active conformation of nucleic acid enzymes), diffusive ion-RNA interactions contribute most to the thermodynamic stabilization of the overall RNA fold.[75]

2.2. Impact of metal ions on nucleic acid catalysis

Given that magnesium is the seventh most abundant element in the Earth's crust, and that the Mg^{2+} ion is the second most abundant cation (55 mM) in sea water after Na^+, it is conceivable that similar Mg^{2+} concentrations were present in an Arche-

© 2019 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

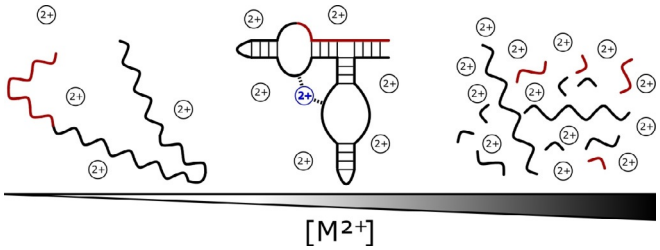


Figure 2. Schematic depicting dependence of RNA folding and hydrolysis on divalent metal ion concentration. Under aqueous conditions, divalent metal ions (in particular Mg^{2+} and Mn^{2+}) can enhance RNA folding by both diffuse binding and site-specific binding (highlighted in blue). In diffuse binding, hydrated Mg^{2+} ions interact nonspecifically with the nucleic acid via long-range electrostatic interactions. In site binding, dehydrated or partially dehydrated Mg^{2+} ions (highlighted in blue) interact specifically with anionic binding sites, which are formed by the RNA fold to act as coordinating ligands for the metal ion. At high M^{2+} concentrations, metal ion catalysis leads to increased RNA hydrolysis.

an ocean,[76] or at varying levels in potential RNA world freshwater environments. However, many other mono-, di- and polyvalent ions can also drive the folding of RNA (and other nucleic acids), including Mn^{2+}, Ca^{2+}, Fe^{2+}, Sr^{2+}, Ba^{2+}, Na^{+} and polyamines.[66,77,78] The ion concentrations required to achieve RNA folding vary between the different ion types, as their charge density and excluded volume largely determine the strength of the coulombic RNA-ion interaction and thus the overall compactness of the folded nucleic acid.[78] For example, the *Tetrahymena* group I ribozyme, which was derived from a self-splicing *Tetrahymena* preribosomal RNA and catalyzes a reaction mimicking the first step of splicing,[79] requires micromolar concentrations of trivalent cations, millimolar concentrations of divalent ions but near-molar concentrations of monovalent ions for folding.[75] However, although the *Tetrahymena* group I ribozyme folds into a native-like state in the presence of various counterions, folding of the catalytically active state requires site-specific binding of Mg^{2+} or Mn^{2+}.[75]

All of the larger natural RNA enzymes, such as RNAseP[9,10] and the various self-splicing introns,[15,16] depend on site-specific metal ion cofactors for chemical reactivity. Likewise, the various artificial RNA ligase and polymerase ribozymes, which rely on nucleoside triphosphate activation chemistry, are strict metalloenzymes with only poor tolerance towards metal ions other than Mg^{2+}.[80] In view of this, it is quite surprising that modern intracellular conditions are somewhat challenging for nucleic acid folding and activity due to low free Mg^{2+} concentrations of approximately 1 mM.[81] The need for higher levels of free Mg^{2+} in vivo is alleviated by the presence of RNA chaperone proteins, which promote RNA folding and annealing.[69] The dependence on intracellular protein co-factors is well illustrated by RNAse P: at low ionic strength, the protein component of this complex is essential for activity in vivo and in vitro.[82,83] However, the RNA itself is active in vitro in the presence of 60 mM $MgCl_2$.[2] The high divalent ion concentration required for RNA-only catalysis in vitro emphasizes that charge screening by either salt or the protein component is essential for folding and activity. Nevertheless, optimal conditions are highly dependent on the catalytic system in question. For example, the family of group II introns has a broad tolerance for Mg^{2+} concentrations and near-optimal activity occurs between 0.1 to 100 mM in vitro.[84]

Like ribozymes, DNAzymes use diffuse electrostatic and specific metal ion interactions for activity and folding. Notably, the high stability, cost-effective production, and easy chemical modification of DNA has enabled the systematic selection of a large number of DNAzymes and aptamers capable of selective metal ion detection. These DNAs can bind to and distinguish between an impressive range of species, including alkali metal ions, alkaline earth metal ions, transition metals, noble metals, post-transition metal ions and lanthanide and actinide ions for catalysis.[85]

It should be mentioned that non-metallic ions can also support folding of nucleic acids into active conformations. For example, polyamines can aid RNA folding; the required $MgCl_2$ concentration for RNAseP RNA folding and activity is reduced from 60 mM to 10 mM in the presence of 1 mM spermidine.[2] However, enhancements in folding are dependent on the characteristics of the polyamine counterion. Longer polyamines destabilize folded structures due to excluded volume effects, which can prevent a complete folding transition to the native state even under usually favorable folding conditions.[77]

Lanthanides (Ln^{3+}) are also of interest, as their interactions with nucleic acids are very different from typical divalent metal ions due to their unusual coordination chemistry. In particular, the absence of a strong ligand field allows for a high degree of structural diversity in lanthanide complexes, as ligands alone dictate the symmetry and coordination of complexes.[86] As a result, lanthanides not only show a high affinity to the phosphate backbone of nucleic acids due to their high charge density (typically only μM concentrations are required for binding), but they can also directly interact with the nucleobase moieties.[87] Because of these unusual properties, the impact of lanthanides on nucleic acid catalysis is rather diverse: Ln^{3+} ions can accelerate a small Pb^{2+}-dependent ribozyme called the leadzyme,[88] yet they inhibit the hammerhead[89] and hairpin[90] ribozymes, and the RNA-cleaving 8–17 DNAzyme.[91] In addition, several strictly Ln^{3+}-dependent RNA-cleaving DNAzymes were discovered by in vitro selection experiments,[92–95] suggesting that nucleic acid enzymes can directly harness the Lewis acid character of lanthanides for catalysis (Figure 3). To the best of our knowledge, Ln^{3+}-specific ribozymes have not yet been described in literature, and at a first glance rare earth metals have little relevance for origin of life scenarios due to their low aqueous solubility. However, low concentrations of lanthanides are available, for example, under hot acidic conditions in volcanic mudpots, and Ln^{3+} ions are essential under these conditions for some acidophilic microbes that use methane as an energy source.[96] This raises the possibility that prebiotic systems relying on nucleic acid catalysis may have been able to harness lanthanides for certain reactions.

2.2. Metal ion induced hydrolysis

While metal ions assist nucleic acid folding and catalysis in many cases, they are often also a threat to the chemical integ-

© 2019 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

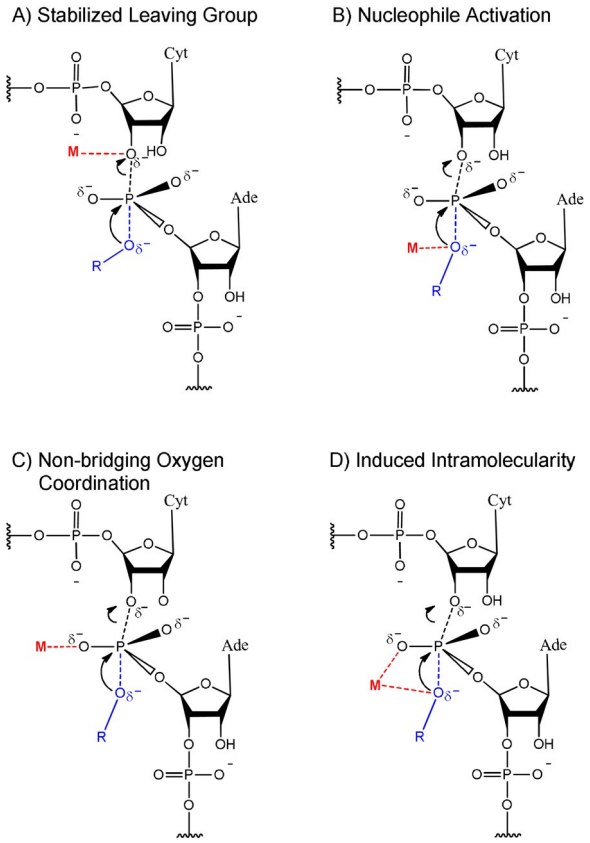


Figure 3. Various modes of interaction between metal ions and RNA during RNA cleavage. The reaction proceeds via a trigonal bipyramidal transition state. The rate of reaction can be accelerated by Lewis acid stabilization of the leaving 3′ oxygen (A), facilitating the deprotonation of the attacking oxygen nucleophile (B), coordination of non-bridging oxygens (C) or coordination of a non-bridging oxygen in addition to the nucleophile (D), which promotes a favorable in-line geometry for nucleophilic attack. The stabilizing metal ion and attacking base are shown in red and blue, respectively. Adapted from Forconi et al. and Frederiksen et al.[104,108]

rity of RNA (Figure 2);[97] heavy metal ions such as Eu^{3+}, La^{3+} and Tb^{3+}, Pb^{2+}, and Zn^{2+} catalyze rapid RNA cleavage in aqueous solutions.[97,98] Zn^{2+} is only about 4% as active as Pb^{2+}, and other metal ions such as Cd^{2+}, Mn^{2+}, Cu^{2+} or Mg^{2+} catalyze degradation one to two orders of magnitude slower than Zn^{2+}.[99] However, at elevated temperatures and/or high ion concentrations, these seemingly weak catalysts (including Mg^{2+}) can reduce RNA half-lives down to minutes.[100] This means that environments with a high concentration of Mg^{2+} and high temperatures, such as hydrothermal vents, are unsuitable settings for RNA-based scenarios of molecular evolution. Likewise, free Ln^{3+} ions are highly nucleolytic under basic conditions, as their ions form multinuclear complexes and cleave RNA nonspecifically at low mM concentrations with a rate acceleration as large as 10^8–10^{12}-fold.[101] DNA is much more resistant towards metal ion-induced scission, and requires additional DNA-binding delivery agents for efficient cleavage under mild aqueous conditions.[102] A notable exception is the ability of Ce^{IV} to accelerate DNA hydrolysis up to 10^{11}-fold under neutral conditions, reducing the half-life of the phosphodiester linkage in DNA from millions of years down to a few hours.[101]

Possible modes of metal ion-catalyzed nucleic acid hydrolysis include Lewis acid catalysis, Brønsted base catalysis, nucleophilic catalysis by metal-bound hydroxides and simple electrostatic stabilization of transition states by positively charged metal ions (Figure 3). The individual mechanisms of each metal ion class are still the subject of some debate and go beyond the focus of this review, but are discussed in excellent detail elsewhere.[101,103–104]

Facing the threat of degradation by metal ions, in particular in the case of RNA, it is interesting from a prebiotic perspective that a number of nucleic acids are capable of efficient catalysis without divalent metal ions. In particular, several families of small nucleolytic ribozymes reversibly catalyze metal-independent and site-specific cleavage/ ligation of the RNA backbone, and can accelerate this reaction by approximately a million-fold using general acid base catalysis.[105] Similarly, purely Na^+-dependent DNAzymes were isolated by targeted in vitro selection.[106,107] Some of these (deoxy-)ribozymes will be discussed later in more detail, as they are compatible with a wide range of conditions.

2.3. Prebiotic alternatives to Mg^{2+}

Of the various ions that can replace Mg^{2+} during nucleic acid folding and catalysis, Fe^{2+} is of great prebiotic interest as it was likely to be highly abundant on Earth before the advent of photosynthesis.[31] Fe^{2+} was speculated to be present in micro- to low millimolar quantities during early Archean Earth.[31] Such concentrations are sufficient to replace Mg^{2+} during RNA cleavage catalyzed by several DNAzymes.[109] As discussed in section 3, Fe^{2+} was used during pH-dependent selection for RNA-cleaving ribozymes, where it enabled the discovery of novel catalytic motifs that are absent in typical selections using Mg^{2+}.[110] Intriguingly, Hsiao et al. showed that substituting Mg^{2+} with Fe^{2+} in an anoxic environment enabled various natural RNAs, such as tRNA or ribosomal RNA, to catalyze single-electron transfer reactions, which are typically limited to cofactor-dependent protein enzymes.[111] Thus, RNA might have catalyzed different electron transfer reactions, which are a prerequisite for metabolic activity, before the rise of oxygen levels.

Zn^{2+} has also been proposed as a key divalent transition metal ion in prebiotic chemistry.[112] In this "Zinc World" hypothesis, porous and photoactive structures comprised of ZnS provided the substrate upon which CO_2 reduction and biomolecular polymerization occurred, driven by UV light. Indeed, Zn^{2+} can substitute Mg^{2+} as the only divalent metal ion during RNAseP catalysis, but only in the presence of high concentrations of ammonium salts.[113] Zn^{2+} was also shown to be strongly beneficial for DNA-catalyzed DNA cleavage. The artificial deoxyribozyme 10MD5 is a bimetallic metalloenzyme (analogous to many protein DNA endonucleases) that catalyzes the Mn^{2+}/Zn^{2+}-dependent DNA phosphodiester bond hydrolysis with at least a 10^{12}-fold rate enhancement.[114] In a follow-up study, Silverman and co-workers demonstrated that only two base substitutions were necessary to alter 10MD5 from hetero-

© 2019 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

bimetallic to a purely Zn^{2+}-dependent monometallic DNAzyme.[115] Later, even faster and smaller deoxyribozymes which require Zn^{2+} alone for catalysis were identified by in vitro selection.[116]

In summary, the availability of metal ions such as magnesium was most likely not a critical factor for early nucleic acid enzymes (especially ribozymes). However, it is possible that Fe^{2+} ions in particular extended the catalytic properties of ribozymes under the anoxic conditions of the late Hadean and early Archean. Further research in this field could uncover new, unexpected catalytic nucleic acids that increase the plausibility of an early metabolism mediated by nucleic acids.

3. The influence of pH on folding and catalysis

3.1. Potential pH values in prebiotic settings

Another crucial physicochemical parameter for early nucleic acid catalysis and stability is pH. Estimates of environmental pH on early Earth are largely hypothetical (see introduction), but most evidence suggests that oceanic pH was initially acidic (pH 6.6,[48] or lower[49,50]).The theory that early molecular evolution originated at alkaline (pH 9–11) hydrothermal vents, similar to the modern Lost City systems, has a number of proponents, but is difficult to reconcile with an RNA-based origin due to the inherent lability of RNA to alkaline hydrolysis, which occurs above pH 6 and is strongly accelerated by higher temperatures and divalent metal ions (Figure 4).[100,117] RNA is most stable at pH 4–5 with significant acid hydrolysis not occurring until below pH 2. Thus, more acidic vent types such as acidic volcanic lakes or comet ponds are credible early scenarios for RNA formation and catalysis.[51]

DNA is less stable than RNA under acidic conditions due to increased depurination below pH 3,[118,119] but is more resistant to basic conditions as it does not possess the 2′-OH group required for base-catalyzed hydrolysis. A DNA-later scenario could therefore be in agreement with a gradual increase of environmental pH over time. Indeed, high CO_2 levels in the Hadean era may have led to a variety of acidic aqueous environments,[49] and the slow transition from acidic to slightly alkaline oceans could have driven the later emergence of the more stable DNA-based systems.[48,120,121]

3.2. The impact of pH on nucleic acid catalysis.

The direct effect of pH on catalysis is inherently dependent on the type and mechanism of the reaction. Catalysis by nucleic acids can occur via transition state stabilization (e.g. by hydrogen bonding or electrostatic stabilization), general acid and/or base catalysis (i.e. by enhancing the nucleophilicity of attacking groups by deprotonation or by stabilizing leaving groups by protonation), or by facilitating active conformational states such as the formation of an in-line transition state during nucleophilic attack.[122] For example, the reversible RNA cleavage reaction catalyzed by small nucleolytic ribozymes, which is based on the nucleophilic attack of an O2′ on an adjacent phosphorus atom, is in most ribozymes accelerated by general

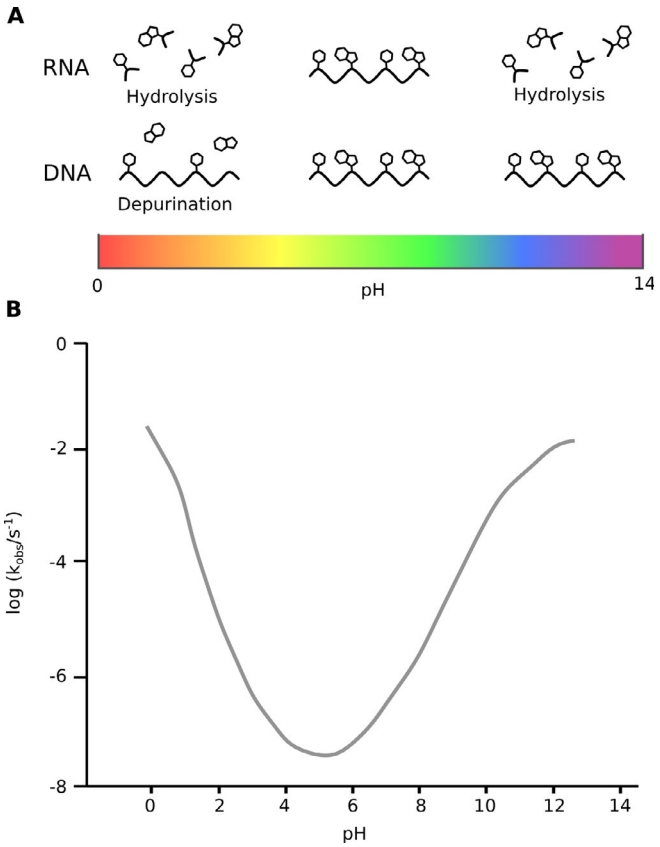


Figure 4. The impact of pH on RNA/ DNA stability. A) Illustration of RNA and DNA stability in different pH ranges. At acidic pH < 2, RNA is prone to hydrolysis, whereas DNA is more susceptible to depurination. At basic pH, the phosphodiester backbone of RNA hydrolyses rapidly, whereas DNA remains stable. B) Relative rate of RNA hydrolysis with respect to pH. Shown is an illustrative pH-rate profile for the cleavage of 3′,5′-UpU at 90 °C based on the data reported by Jarvinen et al.[131]

acid-base catalysis.[122] Here, two ionizable groups stabilize the developing negative and positive charges during the reaction by partial proton transfer in the trigonal bipyramidal phosphorane transition state of the reaction (Figure 5).[122,123] Typically, optimal proton transfer in enzymes requires functional groups with a pK_a in the neutral range.[124] However, the free form of the four canonical nucleobases have pK_a values far from neutrality and are therefore suboptimal for general acid-base catalysis.[125] In some ribozymes, the local molecular environment can cause a considerable shift in the pK_a of both general acid and base towards neutrality, a similar effect to that found in some proteins.[126–128] If both ionizable groups are sufficiently perturbed, the pH dependence of catalytic rates shows a "bell-shaped" pH rate profile, where the rates are maximal around pH 7.[123,129] In other cases, such as for the hairpin (HP) ribozyme, the rates of RNA cleavage (and ligation) increase up to pH 7, but plateau at higher values due to the high pK_a of N1 in the catalytically active guanosine base.[130]

Generally, the acid-base mechanism employed by small ribozymes makes them robust towards changes in pH and enables significant cleavage and ligation activity at pH > 6. However, the rate enhancement is limited by the small fraction of ribo-

© 2019 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

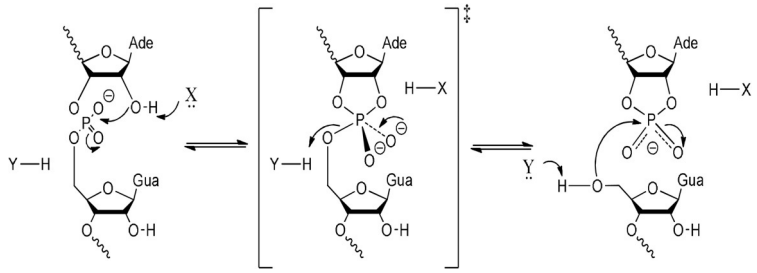


Figure 5. Mechanism of general acid/base-catalyzed RNA phosphodiester cleavage and ligation. General acid-base RNA cleavage and ligation catalyzed by nucleolytic ribozymes. In the cleavage reaction (here, a scissile bond between A and G), the 2′-O attacks the 3′-P in an SN2 process (left). This leads to the formation of a trigonal bipyramidal phosphorane that is probably close to the transition state (middle). Concurrent breaking of the bond to the 5′-O leads to a cyclic 2′,3′ phosphate and 5′-O products. In the ligation reaction, the 5′-O nucleophile attacks the P of the cyclic phosphate. A general base (X) assisting in the removal of the proton from the 2′-OH, and a general acid (Y) protonating the 5′-*O*-oxyanion leaving group catalyze the cleavage reaction. In the reverse ligation reaction, X and Y act as general acid and base, respectively.

zymes that, on average, have the correct ionization state for general acid-base catalyzed cleavage (typically 1 in 10^5 to 10^6 ribozymes for the HP ribozyme at neutral pH[122]). For the reverse ligation reaction the inverse ionization state is more favored, but the resulting rates are offset by a low k_{cat} due to the low reactivity of the neutral base moieties.[123]

The phosphotransfer reactions of large metalloribozymes such as self-splicing introns,[132,133] RNAseP and artificial ligases that make use of triphosphate activation chemistry, show a log-linear relationship between the rate of the chemical step and pH.[134] This is typical for a reaction mechanism involving a pre-equilibrium loss of a proton from a hydroxyl group before in-line nucleophilic attack. Likewise, most RNA-cleaving deoxyribozymes have a log-linear dependence of rate on pH with a slope near unity,[44,135] which is also consistent with the requirement for a single deprotonation event during the reaction.

pH levels also have an important effect on nucleic acid base pairing, as the protonation state of nucleobases dictates their ability to form hydrogen bonds. In particular, at low pH most nucleic acids are denatured (or at least destabilized) due to the protonation of G-C base pairs and resultant Hoogsteen base pair formation.[51] While this mechanism is detrimental for nucleic acid folding, for example, of active ribozymes, environmental pH cycles or gradients[136] may have lowered DNA and RNA duplex melting temperatures, and therefore facilitated non-enzymatic and enzymatic copying reactions.[137] Furthermore, non-canonical A–C and C–C base pairs have been shown to occur under mildly acidic conditions, with A–C base pairs at pH 5 reaching the stability of A–U and G–U base pairs under neutral conditions.[138] Thus, different pH regimes can enable the exploration of structural motifs and thus catalytic sequences that are otherwise inaccessible at neutral pH.

3.3. In vitro selection of nucleic acids catalysts under non-physiological pH conditions

Indeed, in vitro selection experiments have shown that nucleic acids can be readily evolved towards improved catalysis at lower pH where the chemical stability of the RNA backbone is strongly increased. For example, a de novo selection of self-cleaving ribozymes at low pH resulted in a variant that showed pH-dependent kinetics with an optimum of around pH 4.[139] Another study by Popović et al. investigated the effects of both pH and divalent cations on the isolation of self-cleaving RNA in iterative in vitro selection experiments from random libraries.[110] Depending on pH, and whether Mg^{2+} or Fe^{2+} was included as the divalent metal ion during selection, different sequences and secondary structure motifs were isolated. Neutral pH in the presence of Fe^{2+} led to the selection of hammerhead (HH)-like motifs, whilst at pH 5 a variety of previously unknown motifs were discovered and the abundance of HH motifs dropped to less than 0.1 %. Thus, both pH and substitutions between Fe^{2+} and Mg^{2+} strongly influence the relative fitness of different motifs.

Short RNA-cleaving DNAzymes have also been evolved to function in trans at low pH. The reaction proceeds optimally at pH of 4–4.5 in the absence of Mg^{2+}, demonstrating that low pH can facilitate the Mg^{2+}-free cleavage of RNA by a DNAzyme.[140] Moreover, of the 20 clones sampled after selection, 14 did not share extensive sequence similarities, suggesting that the catalysis of the cleavage reaction at low pH has different or relaxed sequence requirements.

Ligation reactions represent an important catalytic function, for example, for nucleic acid self-replication.[141] Consequently, RNA ligases have also been evolved to function at acidic pH. For example, random mutagenesis of a derivative of the triphosphate-dependent class I RNA ligase ribozyme (Figure 1C), followed by four rounds of evolution of the randomized pool under acidic pH, allowed for the selection of clones that function optimally at pH 4 instead of at neutral conditions.[142] Additional mutagenesis of the selected ribozyme further enhanced the rate of ligation by 8000-fold.[143] Kühne and Joyce implemented a continuous in vitro evolution strategy to progressively decrease or increase the optimal pH of the class I ligase ribozyme, beginning with an optimal pH of 8.5.[144] The result was two highly active class I ribozyme variants with only very few mutations that shifted the optimal pH to either pH 5.8 or 9.8.

Early peptide synthesis and even translation may have also occurred over a broad pH range. The peptidyltransfer reaction that takes place at the heart of the ribosome does not involve acid-base catalysis and so is relatively pH-insensitive.[145] A considerable decrease in peptide bond formation is observed only at pH $<$ 6.5 due to inactivation of the attacking amino group of the A-site aa-tRNA by protonation.[146,147] Notably, the activation of amino acids by aminoacetylation, a key step in protein biosynthesis, can also be catalyzed by RNA under acidic condi-

© 2019 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

tions: Kumar et al. reported the selection of a calcium-dependent ribozyme capable of activating amino acids in this manner, with an optimal of pH 4.0–4.5.[148]

4. Heat tolerance of nucleic acid catalysis

Temperature is a further critical parameter in nucleic acid catalysis and stability (Figure 6). As for proteins, reaction rates increase with increasing temperature, until the point at which activity falls due to denaturation. In the absence of magnesium, the duplex melting temperature (T_m) of nucleic acids is sufficiently low to reduce the catalytic potential at even slightly elevated temperatures. In addition, the faster reaction kinetics at elevated temperatures are offset by the increasing rate of phosphodiester hydrolysis, especially in the presence of divalent metal cations such as magnesium as discussed above, which prevents sustained catalysis.

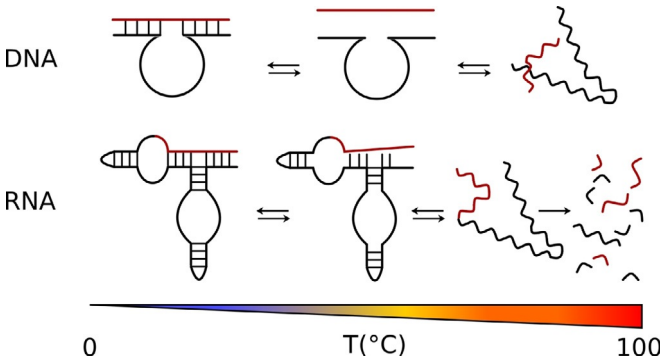


Figure 6. Stability of (deoxy-)ribozymes with increasing temperature. In aqueous environments, low and moderate temperatures support folding of typical secondary and tertiary DNA and RNA structures. Higher temperatures generally support the reversible melting and the resulting formation of unfolded single-stranded nucleic acids. However, the individual melting points and pathways are strongly dependent on the overall number and strength of tertiary and secondary interactions, as well as the concentration of counter-ions. Generally, hybridization of RNA is stronger than that of DNA. High temperatures also increase the rate of spontaneous and irreversible RNA backbone hydrolysis, which is typically not the case for DNA.

4.1. Prebiotic temperatures and thermophilic RNAs

Temperature estimations of the early Earth are a matter of debate. Several lines of evidence exist that support a hot climate during the Archean eon, 4 to 2.5 billion years ago, by which point the Earth's crust is thought to have cooled sufficiently to allow for the dawn of life. Based on oxygen and silicon isotope analyses in sedimentary rocks,[52,54,149] turbidity current deposits that suggest a possible low viscosity ancient ocean,[55] and the progressively decreasing thermostabilities of resurrected ancestral proteins,[150] Archean surface seawater temperatures have been interpreted to range between 60 °C and 80 °C. In contrast, temperatures below 40 °C at the surface have also been proposed based on evidence including deuterium and phosphate isotope analyses,[56,57] and Archean glacial deposits suggest the presence of ice caps.[58] Indeed, more recent 3D climate-carbon models by Charnay et al. predict global mean temperatures between around 8 °C (281 K) and 30 °C (303 K) 3.8 billion years ago, suggesting that cold and even frozen environments may have been present on early Earth.[151] Hydrothermal vent temperatures are highly variable, with gradients from the hot interior (>350 °C) to much colder seawater (or surrounding surface freshwater).[51] This precludes the occurrence of biochemical processes on or near to the surface of the vent, particularly given that the function of typical mesophilic nucleic acid enzymes is lost above ≈70 °C, but conditions in the immediate surroundings may have been rather more amenable.

Despite the temperature sensitivity of RNA, living systems have adapted to survive at extreme temperatures. Comparison of homologous ribozymes in mesophilic and thermophilic organisms reveals how sequence adaptations can lead to higher temperature stability. A study on RNase P homologs in mesophilic and thermophilic bacteria by Pan et al. observed that folding was more cooperative for thermophilic RNA, and the folding pathway proceeded via a different set of intermediate structures despite the high similarity of the final states.[152] Further work revealed that the thermophilic homolog possesses several mutations that increase its stability by increasing GC content and eliminating non-canonical base pairs.[153] In addition, insertions in diverse motifs throughout the thermophilic homolog structure increase tertiary interactions and folding cooperativity while creating a more densely packed core.

4.2. In vitro selection of thermophilic nucleic acid enzymes

Several reports focusing on heat adaptation of nucleic acid enzymes to higher temperatures have been published. Guo et al. used directed evolution to select for thermally stable variants of the *Tetrahymena* ribozyme.[154] A family of temperature stable variants were identified, which were slower than the original ribozyme but had 10.5 °C higher melting temperatures. Whilst the consensus sequence of this family contained nine point mutations, only one served to strengthen the helical secondary structure. The remaining 8 mutations increased tertiary interactions between adjacent motifs, thus improving the packing of the ribozyme structure and presumably favoring active conformations.

Saksmerprome et al. discovered highly thermostable variants of the HH ribozyme.[155] Through in vitro selection, two groups of minimal HH ribozymes were isolated that exhibited trans catalytic activity at elevated temperatures due to strong tertiary interactions between terminal loops and internal bulges that strengthen ribozyme folding and ribozyme-substrate binding. High thermal stability may also be achieved without dedicated selection experiments: Vazquez-Tello et al. discovered that the SMα1 HH ribozyme found in the human parasite *Schistosoma mansoni* HH ribozyme is most active at ≈70 °C in vitro without additional sequence optimizations.[156] Moreover, the same ribozyme can also be successfully cloned and expressed in the thermophile *Thermus thermophilus* where it catalyzes efficient cis- and trans- cleavage of mRNA in vivo at temperatures up to 80 °C. In this case, temperature modulates the rate limiting steps of the reaction: at 37 °C, catalysis is lim-

© 2019 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

ited by substrate dissociation, whereas at high temperature RNA degradation, ribozyme-substrate association, and secondary structure denaturation limit activity.

DNAzymes capable of high temperature catalysis have also been obtained by in vitro selection. Nelson et al. selected a range of Zn^{2+}-dependent RNA-cleaving DNAzymes with activity at 90 °C.[157] The selected sequences share little sequence similarity with other metal dependent DNAzymes, and only slightly enhance cleavage above background levels. Interestingly, no secondary structural features are predicted in the selected sequences at 90 °C, implying that the DNAzyme is capable of binding Zn^{2+} and maintaining catalysis with minimal secondary structure.

These studies demonstrate that the catalysis of nucleic acids can be retained at elevated temperatures. Temperature adaptation in ribozymes is generally achieved through additional RNA-RNA interactions stabilizing both the catalytically active conformation and RNA-substrate interactions, allowing activity to be sustained up to 80 °C. These adaptive mechanisms may generally also decrease the M^{2+} dependency of nucleic acid folding and catalysis, which, in the case of RNA, helps to reduce degradation. More work investigating the stabilization of more primitive, short ribozyme systems is required to examine the range of temperatures that permit the emergence or even self-replication of functional RNAs at increased temperatures. DNA is more resistant to degradation than RNA, so selected DNAzymes can operate at up to 90 °C by relying on metal cofactor binding rather than the maintenance of a well-folded active site. It is as yet unknown whether such systems are limited to simple reactions such as substrate cleavage.

5. Pressure as a modulator of nucleic acid catalysis

In addition to temperature and pH, hydrostatic pressure is also a potentially important environmental factor when considering oceanic or subterranean origins of life. High-pressure conditions are typically defined as 10 MPa or greater, corresponding to a water depth of 1000 m or more. 88% of the volume of modern oceans may be considered high pressure, with an average pressure of 38 MPa and a maximum on the abyssal plane of 110 MPa.[158] Thus, any model of abiogenesis that includes deep-sea vents must account for hydrostatic pressure, which often has profound effects on biological systems by changing the balance of intermolecular interactions. Long-range interactions such as Van der Waals forces and salt bridges become weaker under compression, and shorter interactions such as hydrogen bonds are favored. Under pressure, systems shift towards low volume states in accordance with Le Chatelier's principle. In proteins, dissociation and unfolding is associated with a large negative volume change (−30 to −110 mL mol^{-1}), whilst the DNA double helix dissociation has a positive ΔV of 1–5 mL mol^{-1}.[159–161] This leads to dissociation and unfolding of protein systems as hydrophobic surfaces become hydrated, but nucleic acid structures that are dependent on hydrogen bonding are stabilized. The double helical forms of DNA and RNA are typically stabilized by pressure, with a concomitant increase in melting temperature and no major structural changes other than slight structural distortion due to compression of hydrogen bonding interactions.[162,163] The stabilizing effect is dependent on solution ionic strength and T_m, with duplexes that melt below 50 °C being destabilized by pressure and those melt above 50 °C being stabilized.[159] Certain non-canonical nucleic acid structures, such as the DNA G quadruplex, exhibit negative ΔVs and melt under pressure.[161]

RNA structures are also remarkably stable under high hydrostatic pressure: few structural changes are observed in $tRNA^{Phe}$ up to 1 GPa.[164] Some RNA structures, such duplexes consisting of A-U base pairs, are slightly destabilized by pressure, and more critically the formation of tertiary interactions and docked conformations required for ribozyme catalysis may be disfavored due to positive activation volumes.[165,166] Indeed, the observed rate of cleavage (k_{obs}) and overall equilibrium constant of HP ribozyme self-cleavage decreases with increasing pressure.[166,167] However, despite the overall retardation of the reaction, the actual self-cleavage step is accelerated by hydrostatic pressure and the decrease in rate is attributed to the positive activation volume of docking between catalytic loops.[168] The overall yields of RNA strand cleavage by certain hairpin (HH) ribozymes are improved by high hydrostatic pressure, which can even potentiate catalysis in the absence of the Mg^{2+} typically required for cleavage under ambient pressure.[169,170] Whilst the hammerhead (HH) ribozyme also has a positive activation volume associated with a transition to an active conformation (although significantly smaller than for HP ribozyme), no observable ΔV is associated with the cleavage reaction itself.[171] Molecular dynamics simulations have demonstrated that enhanced hydrogen bonding interactions in the core of the HP and HH ribozymes are responsible for an enhancement in the rate of cleavage under hydrostatic pressure.[172] The effect of hydrostatic pressure appears to extend to deoxyribozyme catalysis: the 10–23 DNAzyme was shown to be active under pressure in the absence of magnesium, albeit with reduced overall yield.[169]

These studies demonstrate that hydrostatic pressure can promote nucleic folding and compensate for a lack of magnesium in certain nucleic acid catalysts. The increase in melting temperature associated with pressurization could permit increased reaction temperatures for weakly folding systems, and be used to avoid Mg^{2+}-catalyzed degradation of RNA. When considering undersea environments, the resistance of nucleic acid to pressure-induced denaturation lends support to a nucleic acid-based origin of life, especially when considering the drastic effect of such conditions on protein folding.

6. Activity enhancement by freezing, evaporation and presence of organic solvents

Apart from the typical physicochemical parameters such as pressure, ionic conditions, pH and temperature described above, more exotic environmental conditions can strongly influence nucleic acid catalysis. A notable example is the extraordinary effect of dehydrating conditions on ribozyme and deox-

© 2019 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

yribozyme catalysis induced by freezing, evaporation, or the presence of organic solvents.

6.1. Freezing and dehydration induced ribozyme catalysis

The discovery that freezing or evaporation can enhance or even trigger ribozyme catalysis was serendipitous. The first reports of (undesired) HH ribozyme activity at sub-zero temperatures came from investigations of the autocatalytic processing of dimeric tobacco ringspot virus satellite RNA (STobRV RNA) by Prody et al.[12] The authors reported difficulties during long-term storage of dimeric STobRV RNA due to self-processing into monomers during one week of storage of the RNA at −20 °C as a precipitate in 67% ethanol. Similar observations of "unwelcome" RNA cleavage in hairpin ribozyme/yeast-mRNA constructs during repeated freezing and thawing were later also reported by Donahue and Fedor.[173] The first systematic investigation of this effect was carried out in 1998 by Kazakov et al., who reported efficient freezing-induced self-ligation of the hairpin (HP) ribozyme even in absence of divalent metal ions such as Mg^{2+}, which are usually indispensable for catalysis in low-salt conditions.[174] Kazakov and his co-workers later expanded their work, and showed that alcohol-induced dehydration and simple evaporation also induced M^{2+}-independent RNA ligation by HP ribozymes in both trans and in cis, while disfavoring the reverse cleavage reaction.[175–177] While divalent metal ions were irrelevant for the freezing-induced ligation, monovalent ions had a strong impact on ligation yields. In particular, sodium salts of acetate-phosphate-borate mixtures, EDTA, and acetate/LiCl led to increased ligation yields.

A first conjecture as to why monovalent salts are important for HP ribozyme catalysis under frozen conditions is provided by previous studies, which have shown that the absence of M^{2+} can be compensated by high concentrations (>1.5 M) of monovalent cations.[178] As already discussed above, several of the small nucleolytic ribozymes such as the HP, HH and VS ribozymes are not obligate metalloenzymes (i.e. metal ions are not involved directly in catalysis) but rely on nucleotide-mediated general acid base catalysis. M^{2+} ions in dilute aqueous solution are still vital for tertiary RNA folding and stabilization of the active conformation.[179,180] The high concentrations of monovalent cations required to substitute for divalent metal ions are readily available in the aqueous phase of water-ice mixtures at temperatures above the eutectic point, in which the crystallization of nearly pure water crystals highly concentrates the remaining aqueous phase (Figure 7).[181]

The activation of the HP ribozyme by the high salt concentration in eutectic brine does not at first seem to explain the alcohol-induced activation of catalysis, since the typical alcohol concentrations used to trigger ribozyme catalysis are not sufficient to co-concentrate or precipitate monovalent counterions.[176,182] However, high concentrations of organic molecules such as primary alcohols or polyethers decrease the dielectric constant of the solvent, thereby strengthening cation-RNA interactions.[183,184] Thus, M^{2+}-independent ribozyme catalysis in presence of primary alcohols or poly(ethlyene glycol) (PEG) might, as in freezing, be at least partially due to the enhanced

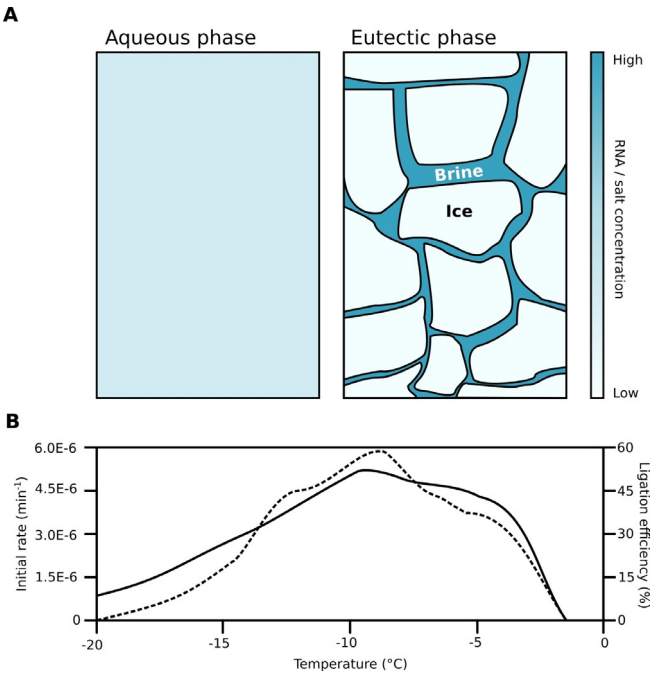


Figure 7. Water ice above the eutectic point is a benign reaction environment for nucleic acids catalysts. A) Schematic showing morphology of eutectic ice phase and relative changes in solute concentration on freezing. The left panel illustrates a dilute aqueous system in an unfrozen state. The right panel shows a partially frozen aqueous solution (e.g. a binary NaCl-water system containing RNA) above the eutectic point. Solutes in the mother liquor (dark blue) are concentrated as a large fraction of almost pure H_2O is sequestered in the ice crystals (light blue). This concentration effect leads to a decreased freezing point of the mother liquor and crystal growth stops when the equilibrium between the ice phase and the liquid phase has been reached. B) Illustrated variation in rate (dashed line) and ligation efficiency (solid line) of the HP ribozyme (excess substrate concentration) in a partially frozen, dilute buffer solution (25 mM NaCl, 1 mM Tris·HCl pH 7.5).[175] Both ligation rate and yields are optimal between −4 °C and −12 °C. At lower temperatures, the low thermal energy available in the system makes it difficult to surmount the activation barrier for the reaction. At temperatures approaching 0 °C, melting of the ice inactivates hairpin ribozyme catalysis in absence of Mg^{2+}.

RNA-Na^+ interactions that can compensate for the missing divalent metal ions.[176] Indeed, even under normal (aqueous) concentrations, ethanol at concentrations above 30% significantly increases the Mg^{2+}-dependent activity of ribozymes and mitigates the effects of destabilizing mutations, although higher levels of ethanol in the presence of Mg^{2+} diminishes this activity, presumably due to RNA aggregation.[185–187] In addition to enhancing ion-ion interactions, dehydration induced by high levels of ethanol or PEG could also support ribozyme activity by promoting the formation of A-form helices (and therefore the catalytic loop structures of ribozymes defined by adjacent helical segments).[186,188]

Kazakov et al. also reported that HP ribozyme-catalyzed ligation during evaporation is considerably improved by the presence of PEG, which had no impact on ligation under aqueous or frozen conditions or ethanol-induced ligation. The authors concluded that PEG might decrease the rate of evaporation, thereby extending the windows of partial dehydration where the water activity is still sufficient to allow hairpin ribozyme

© 2019 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

catalysis.[189] The notion that at least some minimal hydration is required for HP ribozyme catalysis is also in agreement with the reports by Seyhan and Burge, who found that low but non-zero levels of water activity are required for HP and HH ribozyme catalysis in dry RNA films. Intriguingly, hydrated RNA films support cis and trans catalysis over a broad range of temperatures between −70 °C and 37 °C (and probably above), which has potential implications for RNA catalysis under prebiotic conditions.[190]

The formation of active ribozyme conformations in the absence of divalent metal ions can be induced by conditions that promote electrostatic shielding and RNA compaction, such as partial dehydration, up-concentration of monovalent cations, or reduced dielectric constant. Furthermore, the effective increase in RNA concentration during freezing facilitates RNA-RNA association, even from very stable monomeric structures,[191] and has been shown to induce the stretching and alignment of single stranded DNA, which in turn enables its adsorption onto a variety of surfaces.[192]

Freezing favors ligation in reversible transesterification reactions, even from highly fragmented ribozymes.[175,193,194] Freezing can enable highly thermodynamically disfavored reactions, such as ligation of monomeric 2′, 3′-cyclic nucleoside monophosphates to a free 5′ end of RNA.[195] While this reversal of exonucleolytic cleavage has an equilibrium constant of ≈2.2 M^{-1} under aqueous conditions (at 0 °C[196]), it can be decreased ≈20-fold by freezing to −9 °C in the presence of 25 mM NaCl and 10 mM $MgCl_2$, enabling quantitative non-canonical 3′-5′ nucleotidyltransfer of RNA.

Both HP and HH ribozyme ligation yields strongly benefit from repeated freeze-thaw (FT) cycling.[175,193,194] This effect can even be used to enable the in trans assembly of long structured RNAs, such as the ≈200 nt RNA polymerase ribozymes, from fragments between 20–30 nt.[193] The beneficial effects of FT cycles are likely the result of reducing the propensity of small ribozymes to form inactive or poorly active ribozyme-substrate complexes that attenuate bulk catalysis. Repeated freezing and thawing leads to periodic disruption and re-formation of both active and unproductive complexes (in the absence or at low levels of M^{2+}) thereby providing unproductive complexes a "second chance" at catalysis.

Attwater et al. demonstrated the beneficial effects of a frozen environment on strictly M^{2+}-dependent ribozymes such as the R18 RNA polymerase, which catalyzes templated primer extension using nucleoside triphosphates.[197,198] The cold environment considerably extends the lifetime of the polymerase, whilst the concentrating power of freezing above the eutectic temperature enables RNA polymerase activity even at extremely low (unfrozen) starting concentrations of RNA, NTPs and Mg^{2+} salts. The authors also investigated the impacts of different negative counter-ions to Mg^{2+}, and found that they markedly influence activity, presumably due to their influence on the eutectic freezing point, which dictates the concentrating effect of the eutectic brine. The ice microstructure has been shown to provide a quasi-cellular compartmentalization enabling robust phenotype-genotype linkage, which is one of the key requirements for Darwinian Evolution.[198] Indeed, this ice compartmentalization was later used by Attwater et al. to isolate a cold-adapted RNA polymerase ribozyme with considerably increased activity compared to ribozymes selected at ambient temperatures.[199] Recently, Attwater et al. were also able to evolve an ice-adapted RNA trinucleotide polymerase ribozyme that is able to copy its own 170 nt catalytic subunit via the ligation of its almost exclusively triplet-synthesized fragments.[200]

6.2. Freezing and dehydration induced deoxyribozyme catalysis

Zhou et al reported the isolation of the DNAzyme EtNa (Figure 8) from a random DNA library, which is specifically adapted to catalyze RNA cleavage in concentrated organic solvents containing only monovalent Na^+.[201] EtNa shows a rate

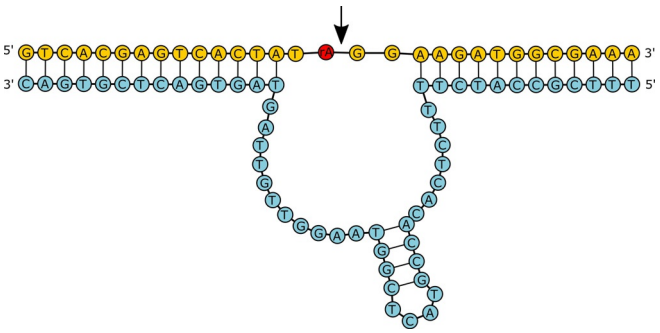


Figure 8. Secondary structure of the EtNa DNAzyme.[201] The substrate strand is shown in yellow, with the ribonucleotide marked in red. The cleavage site is marked by an arrow.

enhancement of up to 1000-fold in 54% ethanol compared to water in presence of 4 mM NaCl, and is completely independent from divalent metal ions. The EtNa RNA cleavage rate can be directly used as a biosensor for the precise measurement of alcohol levels in spirits such as whisky or vodka. Interestingly, EtNa activity drastically decreases at ethanol concentrations beyond 72% (v/v) ethanol, where the B-form helix of DNA is converted into the A-form that (in contrast to ribozymes) seems to be incompatible with the formation of the active DNAzyme conformation. Given that EtNa shows cooperative binding of and activation by Ca^{2+} (in contrast to Mg^{2+})[202] it can also be used as an ultrasensitive biosensor capable of detecting Ca^{2+} levels down to 1.4 μM Ca^{2+}.[203] Eutectic freezing can also activate EtNa, while other DNAzymes that depend on divalent or trivalent metals are inhibited under these conditions.[204] This again highlights the interchangeability of freezing, organic dehydration or evaporation to achieve activation of metal-independent nucleic acid catalysts.

6.3. The potential of wet-dry cycles

The remarkable ability of dehydration to potentiate ribozyme function suggests that such conditions may have been important to the emergence of replicating RNA. Wet-dry cycles, per-

© 2019 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

haps driven by day-night cycles or geothermal activity on early Earth, have been proposed as possible drivers of the emergence of function. Viscous environments formed by water evaporation facilitate non-enzymatic RNA replication cycles slowing reannealing and thereby circumventing strand inhibition.[205] This effect was used by He et al. to form a HH ribozyme by the enzymatic ligation of short fragments, which was functional following dilution in water.[206]

Wet-dry cycles can also be produced by the application of thermal gradients at an air–water interface (Figure 9). The resulting environment up-concentrates a variety of components including RNA precursors and oligonucleotides, enabling a compelling variety of prebiotically important processes including precursor crystallization and phosphorylation.[207] Furthermore, the same environment substantially improves ribozyme catalysis and encapsulation within lipid vesicles. The improved ribozyme catalysis is primarily the result of local high magnesium and RNA concentrations at the air–water interface, but dehydration may also be significant.

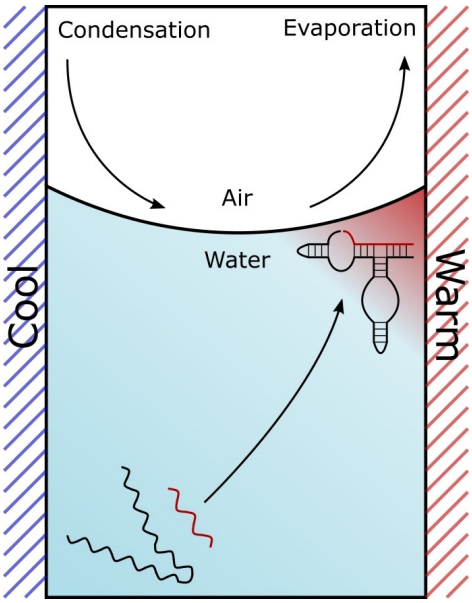


Figure 9. Schematic of a heated rock pore. Thermal gradients at an air–water interface can result in an environment which up-concentrates a variety of components including ribozymes and ions.[207] The improved ribozyme catalysis is most likely the result of local high magnesium and RNA concentrations at the interface. However, direct dehydration of the RNA at the temporally dried interface on the warm side (red) may also contribute to activity. Depending on the geometry of the system, evaporated water condenses at the cold side. The forming water droplets can fall back into the mother solution and wash off the dried components. This can lead to microscopic wet-dry cycles.

7. Ultraviolet light

Exposure to UV radiation presents a challenge to the survival of prebiotic nucleic acids, and is often raised as a major problem in any RNA world scenario due to the elevated levels of surface UV radiation compared to the present day.[208,209] Absorption of ultraviolet photons by nucleobase aromatic rings leads to an excited and highly reactive electronic state, which can give rise to chemical lesions such as adenine cycloaddition to A or T in DNA,[210] as well as the formation of cyclobutane pyrimidine dimers in both DNA and RNA (Figure 9).[211] The effect of UV damage on nucleic acids has been investigated extensively (reviewed by Wurtmann and Wolin),[211] and UV-induced RNA-RNA crosslinking is now an established method for characterizing tertiary or quaternary RNA structure.[212]

Despite its deleterious effect of nucleic acids, ultraviolet radiation has been observed to promote prebiotic chemical reactions that yield ribonucleotides[213–215] and amino acids,[216] and has been proposed as a possible energy source to drive prebiotic chemistry on early Earth.[217] As such, UV radiation could provide an important link between prebiotic chemistry and emergence of an RNA World, but only if radiation levels required to drive such prebiotic reactions can be reconciled with nucleic acid stability under irradiation. Key questions are: To what degree can nucleic acid enzymes sustain photodamage and retain function? Is it possible for nucleic acid enzymes to adapt to strong UV environments?

Despite the well-documented exploration of UV-induced nucleic acid damage, relatively few insights are available regarding the role of UV exposure on functional RNA (or other nucleic acid) enzymes. This may be in part due to a complex interplay between UV radiation and other factors influencing RNA catalysis, such as the presence of metal ions. When exposed to UV radiation, tobacco mosaic virus (TMV) RNA accumulates lesions in the form of uridine hydrates and pyrimidine dimers. However, in the presence of magnesium the rate of accumulation was approximately one-third than that in water, implying that folded RNA is more resistant to UV radiation damage than the unfolded random coil.[218]

The influence of structure and conformation on nucleic acid UV sensitivity was further demonstrated by Kundu et al., who reported an unexpected discrepancy between the UV sensitivities of dTdT dinucleotides in either RNA or DNA hairpins.[219] dTdT dinucleotides embedded in DNA hairpins, which typically adopt a B-form double strand, were susceptible to the formation of photolesions, whilst those in A-form RNA hairpins were protected from damage. The authors also demonstrated that the photosensitivity of the dTdT dinucleotides is modulated by sequence context, with the accumulation of dTdT lesions reduced by neighbouring dA nucleotides, and almost completely inhibited by neighbouring dG nucleotides.[219] It is fascinating that nucleic acids can gain UV resistance simply by adopting a more compact helical conformation, and the sequence dependence of UV photosensitivity suggests that adaptation of nucleic acids to strong UV environments could be possible. Despite this, it must be noted that the effect of UV exposure on functional RNA in vivo typically decreases function.[220–222]

Recently, Saha and Chen monitored the function, folding, and kinetics of RNA aptamers that bind conditionally fluorescent ligands in vitro following UV induced photodamage.[223] One aptamer, Spinach2, retained significant levels of fluorescence after UV exposure compared to the malachite green aptamer. This may be because a large portion of the Spinach aptamer's binding site is comprised of a photostable G-quad-

© 2019 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

ruplex. Single-stranded binding regions were found to be more UV sensitive, confirming that duplex formation is protective against UV radiation,[223,224] and that UV sensitivity significantly depends on folding and conformation.[219]

While UV irradiation has been generally demonstrated to have a detrimental on functional nucleic acids, some examples of UV-dependent nucleic acid catalysts have been reported. Chinnapen and Sen reported the in vitro selection of a DNAzyme with photolyase activity, UV1C, from a pool of random sequences.[225] UV1C is capable of repairing dTdT dimers caused by UV exposure, and requires UV light to function in a manner similar to extant protein photolyase enzymes (Figure 10). The authors later demonstrated that a G-quadruplex near the substrate binding site functions as both an antenna to absorb UV photons and as an electron source for the repair reaction.[226] Intriguingly, a serotonin cofactor dependent photolyase DNAzyme was later selected, which is able to repair both thymine and uracil dimers on ribose and deoxyribose backbones.[227] The discovery that nucleic acids can both harness UV radiation and use this energy to repair photodamage is important, as it provides a mechanism for early replicating systems to survive heavy UV irradiation on Early Earth. In the absence of such a mechanism, early replicators would have to depend on environmental protection from UV radiation, such as the protective effect of montmorillonite clay particles,[228] or shielding by oceanic UV absorbers.[208]

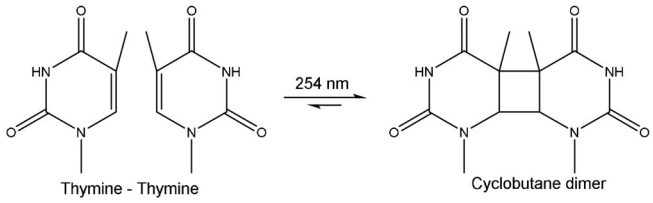


Figure 10. Photodimerization of adjacent thymine nucleotides (dTdT) following exposure to UV light. This photolesion is repaired by the UV1C photolyase DNAzyme, whilst the Sero1C DNAzyme can repair a more diverse range of dimers including thymine, uracil, and several deoxypyrimidine–ribopyrimidine chimeras.[225, 227]

8. Conclusion and perspectives

The activities of both ribozymes and deoxyribozymes are compatible with a broad range of potentially prebiotic conditions. Despite being less versatile and powerful than protein-based catalysis, nucleic acid catalysts are capable of escaping irreversible aggregation, while also tolerating or even benefiting from much harsher conditions such as freezing, drying or dehydration. Moreover, nucleic acid catalysts often require only modest changes in their sequences to adapt to novel challenging conditions such as harsher pH values or higher temperatures, and can often tolerate or adapt to a broad range of different metal ion cofactors. These combined features make them ideal candidates for early biocatalysis, which presumably emerged and remained functional outside the sheltered and constant milieu of the modern cell.

Despite the large body of research, further explorations of nucleic acid enzymes under prebiotic conditions may yield yet more unforeseen properties relevant for abiogenesis, and warrant further investigation. For example, selection experiments under prebiotically plausible conditions beyond aqueous solutions in a modern oxygen-rich atmosphere could reveal further unexpected catalytic properties of ribozymes. In addition to the factors discussed in this review, other environmental factors such as mineral surfaces,[228,229,230] pH gradients,[231] high viscosities[206] or combination of various different environments may further enhance the functional repertoire of early nucleic acids. For example, the clay montmorillonite inhibits HP ribozyme catalysis, but surface adsorption to this mineral offers protection against UV degradation.[228] Clay can also enhance recombination ribozymes and favor ligation by preferentially adsorbing longer RNA strands.[230] Furthermore, it is possible that heterogeneous complexes such as RNA/peptide complexes or mixed RNA/DNA (or alternative preRNA/preDNA) systems were important forerunners to modern biochemistry, and allowed the catalysis of biochemical or replicative processes that "pure" RNA or DNA systems are presently incapable of.[5]

Finally, it remains essential to further expand far-from-equilibrium scenarios to explore different stages of molecular evolution (including nucleic acid catalysis) experimentally under heterogeneous conditions, such as the continuous provision of chemical fuel and/or pH, temperature, or salinity cycles.

Acknowledgements

Financial support from the German Research Foundation (DFG) through CRC/SFB 235 Project P01 (E.Y.S.) and P14 (E.S.) is acknowledged. H.M. is supported by the MaxSynBio consortium, which is jointly funded by the Federal Ministry of Education and Research of Germany and the Max Planck Society. H.M. and K.L.V. are supported by the Volkswagen Initiative "Life?—A Fresh Scientific Approach to the Basic Principles of Life".

Conflict of interest

The authors declare no conflict of interest.

Keywords: catalysis · deoxyribozymes · nucleic acids · origin of life · ribozymes

[1] K. Kruger, P. J. Grabowski, A. J. Zaug, J. Sands, D. E. Gottschling, T. R. Cech, *Cell* **1982**, *31*, 147 – 157.
[2] C. Guerrier-Takada, K. Gardiner, T. Marsh, N. Pace, S. Altman, *Cell* **1983**, *35*, 849 – 857.
[3] M. Neveu, H.-J. Kim, S. A. Benner, *Astrobiology* **2013**, *13*, 391 – 403.
[4] N. V. Hud, *Nat. Commun.* **2018**, *9*, 5171.
[5] S. Bhowmik, R. Krishnamurthy, *Nat. Chem.* **2019**, *11*, 1009 – 1018.
[6] J. W. Szostak, *J. Syst. Chem.* **2012**, *3*, 2.
[7] L. E. Orgel, *Nature* **1992**, *358*, 203 – 209.
[8] T. R. Cech, *Science* **2000**, *289*, 878 – 879.
[9] B. Klemm, N. Wu, Y. Chen, X. Liu, K. Kaitany, M. Howard, C. Fierke, *Biomolecules* **2016**, *6*, 27.
[10] O. Esakova, A. S. Krasilnikov, *RNA* **2010**, *16*, 1725 – 1747.

© 2019 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

[11] C. J. Hutchins, P. D. Rathjen, A. C. Forster, R. H. Symons, *Nucleic Acids Res.* **1986**, *14*, 3627 – 3640.
[12] G. A. Prody, J. T. Bakos, J. M. Buzayan, I. R. Schneider, G. Bruening, *Science* **1986**, *231*, 1577 – 1580.
[13] N. G. Walter, J. M. Burke, *Curr. Opin. Chem. Biol.* **1998**, *2*, 24 – 30.
[14] R. Flores, D. Grubb, A. Elleuch, M. Á. Nohales, S. Delgado, S. Gago, *RNA Biol.* **2011**, *8*, 200 – 206.
[15] T. Cech, *Annu. Rev. Biochem.* **1990**, *59*, 543 – 568.
[16] F. Michel, *Annu. Rev. Biochem.* **1995**, *64*, 435 – 461.
[17] D. Bartel, J. Szostak, *Science* **1993**, *261*, 1411 – 1418.
[18] J. Rogers, G. F. Joyce, *RNA* **2001**, *7*, 395 – 404.
[19] M. M. Conn, J. R. Prudent, P. G. Schultz, *J. Am. Chem. Soc.* **1996**, *118*, 7012 – 7013.
[20] B. Seelig, A. Jäschke, *Chem. Biol.* **1999**, *6*, 167 – 176.
[21] G. Sengle, A. Eisenführ, P. S. Arora, J. S. Nowick, M. Famulok, *Chem. Biol.* **2001**, *8*, 459 – 473.
[22] S. Fusz, A. Eisenführ, S. G. Srivatsan, A. Heckel, M. Famulok, *Chem. Biol.* **2005**, *12*, 941 – 950.
[23] F. Wachowius, J. Attwater, P. Holliger, *Q. Rev. Biophys.* **2017**, *50*, e4.
[24] M. Chandra, S. K. Silverman, *J. Am. Chem. Soc.* **2008**, *130*, 2936 – 2937.
[25] U. Mohan, R. Burai, B. R. McNaughton, *Org. Biomol. Chem.* **2013**, *11*, 2241 – 2244.
[26] A. Flynn-Charlebois, Y. Wang, T. K. Prior, I. Rashid, K. A. Hoadley, R. L. Coppins, A. C. Wolf, S. K. Silverman, *J. Am. Chem. Soc.* **2003**, *125*, 2444 – 2454.
[27] W. E. Purtha, R. L. Coppins, M. K. Smalley, S. K. Silverman, *J. Am. Chem. Soc.* **2005**, *127*, 13124 – 13125.
[28] B. Cuenoud, J. W. Szostak, *Nature* **1995**, *375*, 611 – 614.
[29] W. Wang, L. P. Billen, Y. Li, *Chem. Biol.* **2002**, *9*, 507 – 517.
[30] Y. Li, Y. Liu, R. R. Breaker, *Biochemistry* **2000**, *39*, 3106 – 3114.
[31] S. S. Athavale, A. S. Petrov, C. Hsiao, D. Watkins, C. D. Prickett, J. J. Gossett, L. Lie, J. C. Bowman, E. O'Neill, C. R. Bernier, N. V. Hud, R. M. Wartell, S. C. Harvey, L. D. Williams, *PLoS One* **2012**, *7*, e38024.
[32] Y. Li, D. Sen, *Nat. Struct. Mol. Biol.* **1996**, *3*, 743 – 747.
[33] M. Hollenstein, *Molecules* **2015**, *20*, 20777 – 20804.
[34] A. I. Taylor, V. B. Pinheiro, M. J. Smola, A. S. Morgunov, S. Peak-Chew, C. Cozens, K. M. Weeks, P. Herdewijn, P. Holliger, *Nature* **2015**, *518*, 427 – 430.
[35] L. M. Barge, *Nat. Commun.* **2018**, *9*, 5170.
[36] J. Jortner, *Philos. Trans. R. Soc. London Ser. B* **2006**, *361*, 1877 – 1891.
[37] J. W. Szostak, *Angew. Chem. Int. Ed.* **2017**, *56*, 11037 – 11043; *Angew. Chem.* **2017**, *129*, 11182 – 11189.
[38] L.-F. Wu, J. D. Sutherland, *Emerg. Top. Life Sci.* **2019**, ETLS20190011.
[39] N. V. Hud, *Nat. Commun.* **2018**, *9*, 5171.
[40] R. Krishnamurthy, *Nat. Commun.* **2018**, *9*, 5175.
[41] K. Le Vay, H. Mutschler, *Emerg. Top. Life Sci.* **2019**, ETLS20190024.
[42] Y. Takagi, *Nucleic Acids Res.* **2001**, *29*, 1815 – 1834.
[43] N. Bergman, N. Lau, V. Lehnert, E. Westhof, D. Bartel, *RNA* **2004**, *10*, 176 – 184.
[44] M. Bonaccio, A. Credali, A. Peracchi, *Nucleic Acids Res.* **2004**, *32*, 916 – 925.
[45] E. Bellion, *J. Chem. Educ.* **1992**, *69*, A326.
[46] R. J. P. Williams, J. J. R. F. da Silva, *The Chemistry of Evolution*, **2006**.
[47] C. D. Okafor, J. C. Bowman, N. V. Hud, J. B. Glass, L. D. Williams, in *Nucleic Acids and Molecular Biology*, **2018**, pp. 227 – 243.
[48] J. Krissansen-Totton, G. N. Arney, D. C. Catling, *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 4105 – 4110.
[49] H. S. Bernhardt, W. P. Tate, *Biol. Direct* **2012**, *7*, 4.
[50] J. Kua, J. L. Bada, *Orig. Life Evol. Biosph.* **2011**, *41*, 553 – 558.
[51] W. Martin, J. Baross, D. Kelley, M. J. Russell, *Nat. Rev. Microbiol.* **2008**, *6*, 805 – 814.
[52] F. Robert, M. Chaussidon, *Nature* **2006**, *443*, 969 – 972.
[53] M. J. Whitehouse, C. M. Fedo, *Geology* **2007**, *35*, 719.
[54] L. P. Knauth, D. R. Lowe, *Geol. Soc. Am. Bull.* **2003**, *115*, 566 – 580.
[55] P. Fralick, J. E. Carter, *Precambrian Res.* **2011**, *191*, 78 – 84.
[56] R. E. Blake, S. J. Chang, A. Lepland, *Nature* **2010**, *464*, 1029 – 1032.
[57] M. T. Hren, M. M. Tice, C. P. Chamberlain, *Nature* **2009**, *462*, 205 – 208.
[58] M. J. de Wit, H. Furnes, *Sci. Adv.* **2016**, *2*, e1500368.
[59] E. G. Nisbet, N. H. Sleep, *Nature* **2001**, *409*, 1083 – 1091.
[60] N. H. Sleep, K. Zahnle, *J. Geophys. Res. E* **2001**, *106*, 1373 – 1399.
[61] B. K. D. Pearce, R. E. Pudritz, D. A. Semenov, T. K. Henning, *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 11327 – 11332.
[62] V. L. Bronshteyn, A. A. Chernov, *J. Cryst. Growth* **1991**, *112*, 129 – 145.
[63] I. Cnossen, J. Sanz-Forcada, F. Favata, O. Witasse, T. Zegers, N. F. Arnold, *J. Geophys. Res.* **2007**, *112*, E02008.
[64] A. Fersht, *Enzyme Structure and Mechanism*, W. H. Freeman, New York, **1985**, p. 475.
[65] D. Thirumalai, C. Hyeon, *Biochemistry* **2005**, *44*, 4957 – 4970.
[66] S.-J. Chen, *Annu. Rev. Biophys.* **2008**, *37*, 197 – 214.
[67] S.-J. Chen, K. A. Dill, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 646 – 651.
[68] A. D. Pressman, Z. Liu, E. Janzen, C. Blanco, U. F. Müller, G. F. Joyce, R. Pascal, I. A. Chen, *J. Am. Chem. Soc.* **2019**, *141*, 6213 – 6223.
[69] K. A. Leamy, S. M. Assmann, D. H. Mathews, P. C. Bevilacqua, *Q. Rev. Biophys.* **2016**, *49*, e10.
[70] D. E. Draper, *Biophys. J.* **2008**, *95*, 5489 – 5495.
[71] J. C. Bowman, T. K. Lenz, N. V. Hud, L. D. Williams, *Curr. Opin. Struct. Biol.* **2012**, *22*, 262 – 272.
[72] M. Brännvall, L. A. Kirsebom, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 12943 – 12947.
[73] L. Jenner, N. Demeshkina, G. Yusupova, M. Yusupov, *Nat. Struct. Mol. Biol.* **2010**, *17*, 1072 – 1078.
[74] V. K. Misra, D. E. Draper, *Biopolymers* **1998**, *48*, 113 – 135.
[75] T. J. Wilson, Y. Liu, C. Domnick, S. Kath-Schorr, D. M. J. Lilley, *J. Am. Chem. Soc.* **2016**, *138*, 6151 – 6162.
[76] C. E. J. De Ronde, D. M. deR. Channer, K. Faure, C. J. Bray, E. T. C. Spooner, *Geochim. Cosmochim. Acta* **1997**, *61*, 4025 – 4042.
[77] E. Koculi, N. K. Lee, D. Thirumalai, S. A. Woodson, *J. Mol. Biol.* **2004**, *341*, 27 – 36.
[78] E. Koculi, C. Hyeon, D. Thirumalai, S. A. Woodson, *J. Am. Chem. Soc.* **2007**, *129*, 2676 – 2682.
[79] D. Herschlag, T. R. Cech, *Biochemistry* **1990**, *29*, 10159 – 10171.
[80] M. E. Glasner, N. H. Bergman, D. P. Bartel, *Biochemistry* **2002**, *41*, 8103 – 8112.
[81] A. M. P. Romani in *Magnesium in the Central Nervous System* (Eds.: R. Vink, M. Nechifor), University Of Adelaide Press, Adelaide, **2011**, pp. 13 – 58.
[82] C. Reich, G. J. Olsen, B. Pace, N. R. Pace, *Science* **1988**, *239*, 178 – 181.
[83] J. C. Kurz, S. Niranjanakumari, C. A. Fierke, *Biochemistry* **1998**, *37*, 2393 – 2400.
[84] R. K. O. Sigel, *Eur. J. Inorg. Chem.* **2005**, 2281 – 2292.
[85] W. Zhou, R. Saran, J. Liu, *Chem. Rev.* **2017**, *117*, 8272 – 8325.
[86] N. C. Martinez-Gomez, H. N. Vu, E. Skovran, *Inorg. Chem.* **2016**, *55*, 10083 – 10089.
[87] R. Nishiyabu, N. Hashimoto, T. Cho, K. Watanabe, T. Yasunaga, A. Endo, K. Kaneko, T. Niidome, M. Murata, C. Adachi, Y. Katayama, M. Hashizume, N. Kimizuka, *J. Am. Chem. Soc.* **2009**, *131*, 2151 – 2158.
[88] N. Sugimoto, T. Ohmichi, *FEBS Lett.* **1996**, *393*, 97 – 100.
[89] A. L. Feig, *Science* **1998**, *279*, 81 – 84.
[90] N. G. Walter, N. Yang, J. M. Burke, *J. Mol. Biol.* **2000**, *298*, 539 – 555.
[91] H.-K. Kim, J. Li, N. Nagraj, Y. Lu, *Chem. Eur. J.* **2008**, *14*, 8696 – 8703.
[92] V. Dokukin, S. K. Silverman, *Chem. Sci.* **2012**, *3*, 1707 – 1714.
[93] M. Vazin, P. J. J. Huang, Z. Matuszek, J. Liu, *Biochemistry* **2015**, *54*, 6132 – 6138.
[94] P.-J. J. Huang, M. Vazin, J. Liu, *Biochemistry* **2016**, *55*, 2518 – 2525.
[95] P. J. J. Huang, M. Vazin, Z. Matuszek, J. Liu, *Nucleic Acids Res.* **2015**, *43*, 461 – 469.
[96] A. Pol, T. R. M. Barends, A. Dietl, A. F. Khadem, J. Eygensteyn, M. S. M. Jetten, H. J. M. Op den Camp, *Environ. Microbiol.* **2014**, *16*, 255 – 264.
[97] A. Dallas, A. V. Vlassov, S. A. Kazakov, *Principles of Nucleic Acid Cleavage by Metal Ions*, Springer, Berlin, Heidelberg, **2004**, pp. 61 – 88.
[98] *Organic Chemistry of Nucleic Acids* (Eds.: N. K. Kochetkov, E. I. Budovskii), Springer US, Boston, MA, **1972**.
[99] W. R. Farkas, *Biochim. Biophys. Acta Nucleic Acids Protein Synth.* **1968**, *155*, 401 – 409.
[100] Y. Li, R. R. Breaker, *J. Am. Chem. Soc.* **1999**, *121*, 5364 – 5372.
[101] M. Komiyama, N. Takeda, H. Shigekawa, *Chem. Commun.* **1999**, 1443 – 1451.
[102] L. A. Basile, A. L. Raphael, J. K. Barton, *J. Am. Chem. Soc.* **1987**, *109*, 7550 – 7551.
[103] J. K. Bashkin, L. A. Jenkins, *Comments Inorg. Chem.* **1994**, *16*, 77 – 93.
[104] M. Forconi, D. Herschlag in *Methods in Enzymology*, **2009**, pp. 91 – 106.

 © 2019 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

[105] D. M. J. Lilley, *Philos. Trans. R. Soc. London Ser. B* **2011**, *366*, 2910 – 2917.
[106] S. F. Torabi, P. Wu, C. E. McGhee, L. Chen, K. Hwang, N. Zheng, J. Cheng, Y. Lu, *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 5903 – 5908.
[107] L. Ma, J. Liu, *ChemBioChem* **2019**, *20*, 537 – 542.
[108] J. K. Frederiksen, R. Fong, J. A. Piccirilli in *Nucleic Acid-Metal Ion Interactions*, Royal Society Of Chemistry, Cambridge, **2008**, pp. 260 – 306.
[109] W. J. Moon, J. Liu, *ChemBioChem* **2019**, https://doi.org/10.1002/cbic.201900344.
[110] M. Popović, P. S. Fliss, M. A. Ditzler, *Nucleic Acids Res.* **2015**, *43*, 7070 – 7082.
[111] C. Hsiao, I. C. Chou, C. D. Okafor, J. C. Bowman, E. B. O'neill, S. S. Athavale, A. S. Petrov, N. V. Hud, R. M. Wartell, S. C. Harvey, L. D. Williams, *Nat. Chem.* **2013**, *5*, 525 – 528.
[112] A. Y. Mulkidjanian, M. Y. Galperin, *Biol. Direct* **2009**, *4*, 27.
[113] S. Cuzic, R. K. Hartmann, *Nucleic Acids Res.* **2005**, *33*, 2464 – 2474.
[114] M. Chandra, A. Sachdeva, S. K. Silverman, *Nat. Chem. Biol.* **2009**, *5*, 718 – 720.
[115] Y. Xiao, E. C. Allen, S. K. Silverman, *Chem. Commun.* **2011**, *47*, 1749 – 1751.
[116] H. Gu, K. Furukawa, Z. Weinberg, D. F. Berenson, R. R. Breaker, *J. Am. Chem. Soc.* **2013**, *135*, 9121 – 9129.
[117] W. Martin, M. J. Russell, *Philos. Trans. R. Soc. London Ser. B* **2003**, *358*, 59 – 83.
[118] J. A. Zoltewicz, D. F. Clark, T. W. Sharpless, G. Grahe, *J. Am. Chem. Soc.* **1970**, *92*, 1741 – 1750.
[119] R. An, Y. Jia, B. Wan, Y. Zhang, P. Dong, J. Li, X. Liang, *PLoS One* **2014**, *9*, e115950.
[120] I. Halevy, A. Bachan, *Science* **2017**, *355*, 1069 – 1071.
[121] M. Oivanen, S. Kuusela, H. Lönnberg, *Chem. Rev.* **1998**, *98*, 961 – 990.
[122] D. M. J. Lilley, F. Eckstein, *Ribozymes and RNA Catalysis*, Royal Society Of Chemistry, Cambridge, **2007**.
[123] P. C. Bevilacqua, *Biochemistry* **2003**, *42*, 2259 – 2265.
[124] A. Fersht, *Enzyme Structure and Mechanism*, W. H. Freeman, New York, **1985**.
[125] R. M. Izatt, J. J. Christensen, J. H. Rytting, *Chem. Rev.* **1971**, *71*, 439 – 481.
[126] B. Gong, J.-H. Chen, E. Chase, D. M. Chadalavada, R. Yajima, B. L. Golden, P. C. Bevilacqua, P. R. Carey, *J. Am. Chem. Soc.* **2007**, *129*, 13335 – 13342.
[127] B. Gong, D. J. Klein, A. R. Ferré-D'Amaré, P. R. Carey, *J. Am. Chem. Soc.* **2011**, *133*, 14188 – 14191.
[128] E. A. Frankel, C. A. Strulson, C. D. Keating, P. C. Bevilacqua, *Biochemistry* **2017**, *56*, 2537 – 2548.
[129] T. J. Wilson, A. C. McLeod, D. M. J. Lilley, *EMBO J.* **2007**, *26*, 2489 – 2500.
[130] S. Kath-Schorr, T. J. Wilson, N. S. Li, J. Lu, J. A. Piccirilli, D. M. J. Lilley, *J. Am. Chem. Soc.* **2012**, *134*, 16717 – 16724.
[131] P. Jarvinen, M. Oivanen, H. Lonnberg, *J. Org. Chem.* **1991**, *56*, 5396 – 5401.
[132] D. Herschlag, M. Khosla, *Biochemistry* **1994**, *33*, 5291 – 5297.
[133] J. L. Hougland, A. V. Kravchuk, D. Herschlag, J. A. Piccirilli, *PLoS Biol.* **2005**, *3*, e277.
[134] J. Hsieh, C. A. Fierke, *RNA* **2009**, *15*, 1565 – 1577.
[135] J. Li, *Nucleic Acids Res.* **2000**, *28*, 481 – 488.
[136] L. M. R. Keil, F. M. Möller, M. Kieß, P. W. Kudella, C. B. Mast, *Nat. Commun.* **2017**, *8*, 1897.
[137] A. Mariani, C. Bonfio, C. M. Johnson, J. D. Sutherland, *Biochemistry* **2018**, *57*, 6382 – 6386.
[138] M. Meroueh, C. S. Chow, *Nucleic Acids Res.* **1999**, *27*, 1118 – 1125.
[139] V. K. Jayasena, L. Gold, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 10612 – 10617.
[140] A. Kasprowicz, K. Stokowa-Sołtys, M. Jeżowska-Bojczuk, J. Wrzesiński, J. Ciesiołka, *ChemistryOpen* **2017**, *6*, 46 – 56.
[141] N. Paul, G. F. Joyce, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 12733 – 12740.
[142] Y. Miyamoto, N. Teramoto, Y. Imanishi, Y. Ito, *Biotechnol. Bioeng.* **2001**, *75*, 590 – 596.
[143] Y. Miyamoto, N. Teramoto, Y. Imanishi, Y. Ito, *Biotechnol. Bioeng.* **2005**, *90*, 36 – 45.
[144] H. Kühne, G. F. Joyce, *J. Mol. Evol.* **2003**, *57*, 292 – 298.
[145] P. Bieling, M. Beringer, S. Adio, M. V. Rodnina, *Nat. Struct. Mol. Biol.* **2006**, *13*, 423 – 428.
[146] M. Johansson, K.-W. Ieong, S. Trobro, P. Strazewski, J. Aqvist, M. Y. Pavlov, M. Ehrenberg, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 79 – 84.
[147] M. V. Rodnina, *Curr. Opin. Struct. Biol.* **2013**, *23*, 595 – 602.
[148] R. K. Kumar, M. Yarus, *Biochemistry* **2001**, *40*, 6998 – 7004.
[149] L. P. Knauth, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **2005**, *219*, 53 – 69.
[150] E. A. Gaucher, S. Govindarajan, O. K. Ganesh, *Nature* **2008**, *451*, 704 – 707.
[151] B. Charnay, G. Le Hir, F. Fluteau, F. Forget, D. C. Catling, *Earth Planet. Sci. Lett.* **2017**, *474*, 97 – 109.
[152] T. Pan, T. R. Sosnick, X. W. Fang, K. Littrell, B. L. Golden, V. Shelton, P. Thiyagarajan, *Proc. Natl. Acad. Sci. USA* **2002**, *98*, 4355 – 4360.
[153] N. J. Baird, *RNA* **2006**, *12*, 598 – 606.
[154] F. Guo, T. R. Cech, *Nat. Struct. Biol.* **2002**, *9*, 855 – 861.
[155] V. Saksmerprome, M. Roychowdhury-Saha, S. Jayasena, A. Khvorova, D. H. Burke, *RNA* **2004**, *10*, 1916 – 1924.
[156] A. Vazquez-Tello, P. Castán, R. Moreno, J. M. Smith, J. Berenguer, R. Cedergren, *Nucleic Acids Res.* **2002**, *30*, 1606 – 1612.
[157] K. E. Nelson, P. J. Bruesehoff, Y. Lu, *J. Mol. Evol.* **2005**, *61*, 216 – 225.
[158] I. Daniel, P. Oger, R. Winter, *Chem. Soc. Rev.* **2006**, *35*, 858 – 875.
[159] D. N. Dubins, A. Lee, R. B. Macgregor, T. V. Chalikian, *J. Am. Chem. Soc.* **2001**, *123*, 9254 – 9259.
[160] J. Roche, J. A. Caro, D. R. Norberto, P. Barthe, C. Roumestand, J. L. Schlessman, A. E. Garcia, B. E. Garcia-Moreno, C. A. Royer, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 6945 – 6950.
[161] S. Takahashi, N. Sugimoto, *Molecules* **2013**, *18*, 13297 – 13319.
[162] R. B. Macgregor, *Biopolymers* **1998**, *48*, 253 – 263.
[163] D. J. Wilton, M. Ghosh, K. V. A. Chary, K. Akasaka, M. P. Williamson, *Nucleic Acids Res.* **2008**, *36*, 4032 – 4037.
[164] C. Schuabb, M. Berghaus, C. Rosin, R. Winter, *ChemPhysChem* **2015**, *16*, 138 – 146.
[165] F. Hughes, R. F. Steiner, *Biopolymers* **1966**, *4*, 1081 – 1090.
[166] G. Hervé, S. Tobé, T. Heams, J. Vergne, M.-C. Maurel, *Biochim. Biophys. Acta Proteins Proteomics* **2006**, *1764*, 573 – 577.
[167] S. Tobé, T. Heams, J. Vergne, G. Hervé, M. C. Maurel, *Nucleic Acids Res.* **2005**, *33*, 2557 – 2564.
[168] C. Schuabb, N. Kumar, S. Pataraia, D. Marx, R. Winter, *Nat. Commun.* **2017**, *8*, 14661.
[169] A. Fedoruk-Wyszomirska, E. Wyszko, M. Giel-Pietraszuk, M. Z. Barciszewska, J. Barciszewski, *Int. J. Biol. Macromol.* **2007**, *41*, 30 – 35.
[170] M. Giel-Pietraszuk, A. Fedoruk-Wyszomirska, J. Barciszewski, *Mol. Biol. Rep.* **2010**, *37*, 3713 – 3719.
[171] H. Kaddour, J. Vergne, G. Hervé, M. C. Maurel, *FEBS J.* **2011**, *278*, 3739 – 3747.
[172] N. Kumar, D. Marx, *Phys. Chem. Chem. Phys.* **2018**, *20*, 20886 – 20898.
[173] C. P. Donahue, M. J. Fedor, *RNA* **1997**, *3*, 961 – 973.
[174] R. H. Sarma, M. H. Sarma, *J. Biomol. Struct. Dyn.* **1998**, *15*, 10 – 14.
[175] A. V. Vlassov, B. H. Johnston, L. F. Landweber, S. A. Kazakov, *Nucleic Acids Res.* **2004**, *32*, 2966 – 2974.
[176] S. A. Kazakov, S. V. Balatskaya, B. H. Johnston, *RNA* **2006**, *12*, 446 – 456.
[177] A. V. Vlassov, B. H. Johnston, S. A. Kazakov, *Oligonucleotides* **2005**, *15*, 303 – 309.
[178] E. A. Curtis, D. P. Bartel, *RNA* **2001**, *7*, 546 – 552.
[179] J. L. O'Rear, S. Wang, A. L. Feig, L. Beigelman, O. C. Uhlenbeck, D. Herschlag, *RNA* **2001**, *7*, 537 – 545.
[180] J. B. Murray, A. A. Seyhan, N. G. Walter, J. M. Burke, W. G. Scott, *Chem. Biol.* **1998**, *5*, 587 – 595.
[181] G. L. Stepakoff, D. Siegelman, R. Johnson, W. Gibson, *Desalination* **1974**, *15*, 25 – 38.
[182] A. L. Fink, M. A. Geeves in *Methods in Enzymology*, **1979**, pp. 336 – 370.
[183] S. Nakano, D. Miyoshi, N. Sugimoto, *Chem. Rev.* **2014**, *114*, 2733 – 2758.
[184] S. Nakano, Y. Kitagawa, H. Yamashita, D. Miyoshi, N. Sugimoto, *ChemBioChem* **2015**, *16*, 1803 – 1810.
[185] K. J. Gardiner, T. L. Marsh, N. R. Pace, *J. Biol. Chem.* **1985**, *260*, 5415 – 5419.
[186] M. Hanna, J. W. Szostak, *Nucleic Acids Res.* **1994**, *22*, 5326 – 5331.
[187] A. L. Feig, G. E. Ammons, O. C. Uhlenbeck, *RNA* **1998**, *4*, 1251 – 1258.
[188] S. Beneventi, G. Onori, *Biophys. Chem.* **1986**, *25*, 181 – 190.
[189] S. J. Prestrelski, T. Arakawa, J. F. Carpenter, *Arch. Biochem. Biophys.* **1993**, *303*, 465 – 473.
[190] A. A. Seyhan, J. M. Burke, *RNA* **2000**, *6*, 189 – 198.
[191] X. Sun, J. M. Li, R. M. Wartell, *RNA* **2007**, *13*, 2277 – 2286.

 © 2019 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

[192] B. Liu, T. Wu, Z. Huang, Y. Liu, J. Liu, *Angew. Chem. Int. Ed.* **2019**, *58*, 2109 – 2113; *Angew. Chem.* **2019**, *131*, 2131 – 2135.
[193] H. Mutschler, A. Wochner, P. Holliger, *Nat. Chem.* **2015**, *7*, 502 – 508.
[194] L. Lie, S. Biliya, F. Vannberg, R. M. Wartell, *J. Mol. Evol.* **2016**, *82*, 81 – 92.
[195] H. Mutschler, P. Holliger, *J. Am. Chem. Soc.* **2014**, *136*, 5193 – 5196.
[196] S. C. Mohr, R. E. Thach, *J. Biol. Chem.* **1969**, *244*, 6566 – 6576.
[197] W. K. Johnston, P. J. Unrau, M. S. Lawrence, M. E. Glasner, D. P. Bartel, *Science* **2001**, *292*, 1319 – 1325.
[198] J. Attwater, A. Wochner, V. B. Pinheiro, A. Coulson, P. Holliger, *Nat. Commun.* **2010**, *1*, 76.
[199] J. Attwater, A. Wochner, P. Holliger, *Nat. Chem.* **2013**, *5*, 1011 – 1018.
[200] J. Attwater, A. Raguram, A. S. Morgunov, E. Gianni, P. Holliger, *eLife* **2018**, *7*, e35255.
[201] W. Zhou, R. Saran, Q. Chen, J. Ding, J. Liu, *ChemBioChem* **2016**, *17*, 159 – 163.
[202] W. Zhou, R. Saran, P. J. J. Huang, J. Ding, J. Liu, *ChemBioChem* **2017**, *18*, 518 – 522.
[203] T. Yu, W. Zhou, J. Liu, *ChemBioChem* **2018**, *19*, 31 – 36.
[204] T. Yu, W. Zhou, J. Liu, *ChemBioChem* **2018**, *19*, 1012 – 1017.
[205] C. He, I. Gállego, B. Laughlin, M. A. Grover, N. V. Hud, *Nat. Chem.* **2017**, *9*, 318 – 324.
[206] C. He, A. Lozoya-Colinas, I. Gállego, M. A. Grover, N. V. Hud, *Nucleic Acids Res.* **2019**, *47*, 6569 – 6577.
[207] M. Morasch, J. Liu, C. F. Dirscherl, A. Ianeselli, A. Kühnlein, K. Le Vay, P. Schwintek, S. Islam, M. K. Corpinot, B. Scheu, D. B. Dingwell, P. Schwille, H. Mutschler, M. W. Powner, C. B. Mast, D. Braun, *Nat. Chem.* **2019**, *11*, 779 – 788.
[208] H. J. Cleaves, S. L. Miller, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 7260 – 7263.
[209] I. Cnossen, J. Sanz-Forcada, F. Favata, O. Witasse, T. Zegers, N. F. Arnold, *J. Geophys. Res. E* **2007**, *112*, E02008.
[210] K. Heil, D. Pearson, T. Carell, *Chem. Soc. Rev.* **2011**, *40*, 4271 – 4278.
[211] E. J. Wurtmann, S. L. Wolin, *Crit. Rev. Biochem. Mol. Biol.* **2009**, *44*, 34 – 49.
[212] C. Zwieb, A. Ross, J. Rinke, M. Meinke, R. Brimacombe, *Nucleic Acids Res.* **1978**, *5*, 2705 – 2720.
[213] B. H. Patel, C. Percivalle, D. J. Ritson, C. D. Duffy, J. D. Sutherland, *Nat. Chem.* **2015**, *7*, 301 – 307.
[214] M. W. Powner, J. D. Sutherland, *ChemBioChem* **2008**, *9*, 2386 – 2387.
[215] J. Xu, M. Tsanakopoulou, C. J. Magnani, R. Szabla, J. E. Šponer, J. Šponer, R. W. Góra, J. D. Sutherland, *Nat. Chem.* **2017**, *9*, 303 – 309.
[216] P. K. Sarker, J. Takahashi, Y. Kawamoto, Y. Obayashi, T. Kaneko, K. Kobayashi, *Int. J. Mol. Sci.* **2012**, *13*, 1006 – 1017.
[217] S. Ranjan, D. D. Sasselov, *Astrobiology* **2016**, *16*, 68 – 88.
[218] B. Singer, *Virology* **1971**, *45*, 101 – 107.
[219] L. M. Kundu, U. Linne, M. Marahiel, T. Carell, *Chem. Eur. J.* **2004**, *10*, 5697 – 5705.
[220] T. V. Ramabhadran, J. Jagger, *Proc. Natl. Acad. Sci. USA* **1976**, *73*, 59 – 63.
[221] S. Boldissar, M. S. De Vries, *Phys. Chem. Chem. Phys.* **2018**, *20*, 9701 – 9716.
[222] M. S. Iordanov, D. Pribnow, J. L. Magun, T.-H. Dinh, J. A. Pearson, B. E. Magun, *J. Biol. Chem.* **1998**, *273*, 15794 – 15803.
[223] R. Saha, I. A. Chen, *ChemBioChem* **2019**, *20*, 2609 – 2617.
[224] M. Pearson, H. E. Johns, *J. Mol. Biol.* **1966**, *20*, 215 – 229.
[225] D. J.-F. Chinnapen, D. Sen, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 65 – 69.
[226] D. J. F. Chinnapen, D. Sen, *J. Mol. Biol.* **2007**, *365*, 1326 – 1336.
[227] R. E. Thorne, D. J. F. Chinnapen, G. S. Sekhon, D. Sen, *J. Mol. Biol.* **2009**, *388*, 21 – 29.
[228] E. Biondi, S. Branciamore, M. C. Maurel, E. Gallori, *BMC Evol. Biol.* **2007**, *7*, S2.
[229] J. D. Stephenson, M. Popović, T. F. Bristow, M. A. Ditzler, *RNA* **2016**, *22*, 1893 – 1901.
[230] R. Mizuuchi, A. Blokhuis, L. Vincent, P. Nghe, N. Lehman, D. Baum, *Chem. Commun.* **2019**, *55*, 2090 – 2093.
[231] L. M. R. Keil, F. M. Möller, M. Kieß, P. W. Kudella, C. B. Mast, *Nat. Commun.* **2017**, *8*, 1897.

Manuscript received: August 27, 2019
Revised manuscript received: November 5, 2019
Accepted manuscript online: November 12, 2019
Version of record online: December 9, 2019

 © 2019 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

CURRICULUM VITAE

Yeonwha (Emilie) Song Wratil

EDUCATION

Ph.D. (02/2019 – 05/2022)	Faculty of Biology, LMU Munich, Max Planck Institute of Biochemistry AG Prof. Dr. Hannes Mutschler *Paths to constructive processes in RNA-based early life*
M.Sc. (09/2016 – 12/2018)	Department of Molecular Genetics, University of Toronto Roy and Nodwell Labs *Towards the discovery of novel stress-sensitizing natural product anthelmintics*
B.Sc. (09/2011 – 05/2016)	Bachelor of Science, Biochemistry and Organic Chemistry University of British Columbia

FURTHER RESEARCH EXPERIENCE

05/2015 – 07/2015	Research Internships in Science and Engineering (RISE) Scholarship recipient Max Planck Institute of Colloids and Interfaces
10/2013 – 08/2014	Synthetic Radiochemist, Department of Chemistry University of British Columbia

SELECT AWARDS AND HONOURS

2021	Junior Scientist Publication Award, Max Planck Institute of Biochemistry
2018	Student Poster Prize, Anthelmintics III Conference
2018	Ontario Graduate Scholarship
2015	Undergraduate Summer Research Award, Natural Sciences and Engineering Research Council of Canada
2011	President's Entrance Scholarship, University of British Columbia