Towards Building a Self-replicating Protocell: Nonenzymatic RNA Copying Driven by Potentially Prebiotic Activation and Higher-order Behaviors in Prebiotic Vesicles

HARVARD UNIVERSITY
Graduate School of Arts and Sciences

DISSERTATION ACCEPTANCE CERTIFICATE

The undersigned, appointed by the

Department of Chemistry & Chemical Biology

have examined a dissertation entitled:

Towards Building a Self-replicating Protocell: Nonenzymatic RNA Copying Driven by Potentially Prebiotic Activation and Higher-order Behaviors in Prebiotic Vesicles

presented by: Stephanie J. Zhang

candidate for the degree of Doctor of Philosophy and hereby certify that it is worthy of acceptance.

Signature _____
Typed name: Professor Jack W. Szostak

Signature _____
Typed name: Professor Stuart L. Schreiber

Signature _____
Typed name: Professor Dimitar Sasselov

Date: 8 Aug 2022

Towards Building a Self-replicating Protocell:
Nonenzymatic RNA Copying Driven by Potentially Prebiotic Activation and Higher-order Behaviors in Prebiotic Vesicles

A dissertation presented

by

Stephanie J. Zhang

to

The Graduate School of Arts and Sciences

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Chemistry

Harvard University

Cambridge, Massachusetts

August 2022

Dissertation Advisor: Dr. Jack W. Szostak Stephanie J. Zhang

Towards Building a Self-replicating Protocell: Nonenzymatic RNA Copying Driven by Potentially Prebiotic Activation and Higher-order Behaviors in Prebiotic Vesicles

Abstract

The emergence of a protocell capable of Darwinian evolution would require replicating compartments and genetic materials. Compartmentalization keeps informational polymers localized so that the functions they encode lead to an advantage in terms of replication or survival. Meanwhile, a mechanism for the inheritance of useful functions must exist. RNA and its close relatives are attractive candidate prebiotic molecules because they harbor the capacity for both inheritance and function in a single class of molecules. In this RNA world model, RNA polymers assemble from activated nucleotides, and when RNA oligomers of adequate lengths lead to ribozymes, RNA can begin to facilitate its own reproduction (Gilbert 1986). The propagation of functional RNAs within reproducing compartments leads to replicating protocells, ultimately giving rise to modern cellular life. My dissertation work explores both chemical and physical processes relevant to the transition from chemical to biological evolution on the early earth.

The hypothesized central role of RNA in the origin of life suggests that RNA propagation predated the advent of complex protein enzymes. Such a process in turn requires a source of chemical energy for in situ nucleotide activation. Previously identified activation chemistries lead to damaging side reactions that destroy both templates and substrates (Biron and Pascal 2004, Fahrenbach, Giurgiu et al. 2017). A chemical process that can continuously reactivate hydrolyzed substrates and is compatible with chemical RNA copying is needed. A potentially prebiotic nucleotide activation that might be compatible with RNA copying remained elusive for nearly 60 years. I have demonstrated a potentially prebiotic pathway that chemically activates RNA nucleotides in a manner compatible with RNA copying (Zhang, Duzdevich et al. 2020). Following

that, I have developed and characterized a prebiotically plausible scenario under which previously isolated key steps of nucleotide activation to nonenzymatic RNA copying can happen together under mutually compatible conditions (Zhang, Duzdevich et al. 2022). This pathway enables more realistic models of RNA propagation.

Early cell membranes are thought to have been composed of fatty acids and related single-chain amphiphiles. Their probable involvement in the origin of life has been further demonstrated by their capability to grow and divide without complex biochemical machinery and encapsulate RNA templates that are being nonenzymatically copied (Hanczyc and Szostak 2004, Budin, Debnath et al. 2012, Adamala and Szostak 2013). However, without complex protein machinery, protocells would have had to rely on passive diffusion for internalizing nutrients. I have shown how primitive cells composed of fatty acids endocytose via a purely physicochemical process upon encountering nutrients in conjunction with extra membrane material. I then found that such inward vesiculation events could lead to internalization of nutrient solutes including mononucleotides and oligonucleotides, drawing further parallels to endocytosis. Such processes could have helped primitive cells capture nutrients that are otherwise impermeable.

Taken together, these results demonstrate prebiotically plausible pathways from nucleotides to RNA copying as well as a scenario in which protocells internalize useful nutrients bypassing permeability limits, bringing us closer to developing primitive cells capable of Darwinian evolution.

Acknowledgements

I have always been fascinated by living creatures, even since childhood. And yet, I have never pondered the question, "How did life begin?", which seems so far beyond my reach. Encountering the Miller-Urey experiment amazed me as complex building blocks of life could have arisen through simple chemical reactions. It made me think that simple chemistry experiments might be able to unravel life's origins. However, it was not until joining the Szostak lab that I started to think seriously about life's primordial beginnings.

I would like to thank my advisor, Jack. His constant support throughout my PhD, allowed me to pursue my own interest and develop into an independent thinker. I am thankful for my thesis committee members, Dr. Stuart Schreiber, and Dr. Dimitar Sasselov, for their time and expertise. I am also thankful to all the members of the Szostak lab, especially my collaborators, Dr. Saurja DasGupta, Dr. Anna Wang, Dr. Daniel Duzdevich, Dian Ding, and Xiwen Jia.

I thank my parents for encouraging me to pursue my goal. I am grateful for everything they have done for me. My grandmother, although she is no longer with us, but she always taught me how to be fearless to challenges. Finally, I want to thank Zhen Jiang, from meeting in Cupertino High to Boston. His constant support and friendship have always helped make challenges a bit easier and life a lot more fun.

Chapter 1

Introduction: *Looking backwards to when life began*

"…probably all the organic beings which have ever lived on this earth may be descended from some one primordial form." — Charles R. Darwin

The origin of life remains one of the grand mysteries of science. Research on the origin of life approached the problems either from the top down, by tracing evolution backward from extant life to most primordial ancestors, or from the bottom up, by constructing a minimal system capable of Darwinian evolution (Walker, Packard et al. 2017) (Figure 1.1). Eventually, these two approaches converge with life emerging somewhere along the way, i.e., either descending down the evolutionary tree (top-down) to the existence of a last universal common ancestor, LUCA (Woese and Fox 1977), or continuing after the emergence of living entities in an inanimate environment (bottom-up) to the simplest living cell. However, the lack of direct evidence of the evolutionary epoch poses challenges in working backward from contemporary life to LUCA. From the bottom up, understanding the origin of life requires the discovery of plausible pathways for the transition from complex chemistry to the emergence of chemical assemblies capable of Darwinian evolution. Such endeavors include but not limited to prebiotic chemistry (Orgel 2004), the replication and self-assembly of membrane vesicles (Hanczyc and Szostak 2004, Budin, Debnath et al. 2012), and the nonenzymatic template-directed copying of nucleic acid sequences (Joyce and Orgel 2006, Li, Prywes et al. 2017). The resulting chemical constraints can be used to ascertain the likely geochemical environments of interest. Further, such chemical assemblies, though artificial, may facilitate the reconstruction of systems that are more realistic analogs of probable early earth systems.

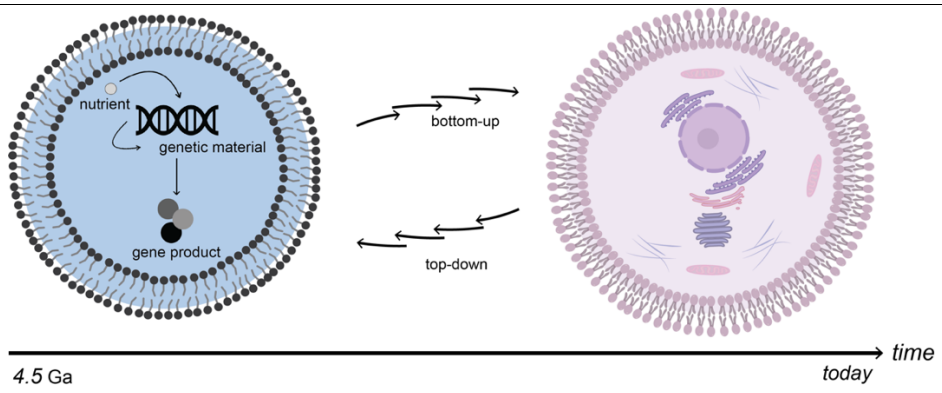


Figure 1.1. General approaches to studying the origin of life include "bottom-up" and "top-down" approaches. The "bottom-up" approach traces molecular evolution from simple to complex, whereas the "top-down" approach seeks to yield a minimal cell by removing complexity and identify the conditions in which it arose.

1.1 Scenarios for the origin of life

RNA world hypothesis

In modern biology, nucleic acids and proteins carry out distinct but vital roles: DNA carries out genetic information and proteins catalyze diverse metabolism reactions. This raises a profound puzzle at life's origins, whether one function precedes the other. A potential solution, the RNA world hypothesis, which postulated that primitive lifeforms might have relied solely on RNA to store genetic information and catalyze reactions, was proposed in the late 1960s (Crick 1968, Orgel 1968). The discovery of a class of RNAs that can catalyze a significant number of diverse chemical reactions, ribozymes, supports the RNA world hypothesis (Kruger, Grabowski et al. 1982, Guerriertakada, Gardiner et al. 1983). For instance, the ribosome, which assembles proteins, is itself a ribozyme (Steitz and Moore 2003, Fox 2010).

In some prebiotic locale, where surface water is present with an atmosphere containing gases such as CH_4 (though currently deemed unrealistic due to its rapid destruction upon UV light), N_2,

and CO_2, produces highly complex mixtures of organic molecules, a very small fraction of which could be potentially life-relevant molecules. Such molecules are steadily accumulated and grow in complexity, potentially driven by energy sources as available, including solar radiation, cosmic impacts, and lightning discharge. The products of this phase are presumed to include amino acids, lipids, sugars, nucleic acid bases, and, ultimately, the first nucleotide. These components could have assembled into activated nucleotides that polymerized to form polynucleotides of random sequences (Ferris, Hill et al. 1996, Kanavarioti, Monnard et al. 2001). Through either monomer addition or ligation, random strands could be copied into RNA oligonucleotides that are long enough to fold into functional structures. In this way, a library of RNA strands with diverse sequences, including the ones capable of hereditary storage and autocatalysis, formed spontaneously on the primitive Earth, and RNA can then facilitate its own reproduction in prebiotic environment (Tjhung, Shokhirev et al. 2020, Walton, DasGupta et al. 2021).

Alternative scenarios

RNA world hypothesis, however, does not rule out the possibility that peptides may have been involved in the origin of life. Recently, non-canonical RNA bases (Decatur and Fournier 2002, Carell, Brandmayr et al. 2012), which are found today in transfer and ribosomal RNAs, enable RNA-based peptide synthesis (Muller, Escobar et al. 2022). In addition, recent studies of prebiotic nucleotide synthesis suggest alternative genetic systems during the origin of life, including nucleotides with altered sugar moieties, arabinonucleic acid (ANA), threose nucleic acid (TNA) (Kim, Zhou et al. 2020, Zhang, Kim et al. 2021), as well as a number of noncanonical nucleobases, such as 2-thiopyrimidines and inosine (Fialho, Roche et al. 2020, Yadav, Kumar et al. 2020). These substrates have shown decreased incorporation into primer compared to RNA. However, they can be copied over when present in the template (Kim, Zhou et al. 2020). These findings suggest that

the non-standard nucleotides will be efficiently filtered out and primitive genomes will exhibit selection bias towards RNA.

1.2 Experimental model: template-directed nonenzymatic RNA copying

Prior to the hypothesized emergence of the more elaborate ribozyme-catalyzed replication machinery, the emergence of life in the RNA world scenario would have required nonenzymatic (i.e., chemical) RNA replication to propagate genetic information. Efforts to demonstrate the nonenzymatic replication of RNA oligonucleotides date back almost 60 years (Weimann, Lohrmann et al. 1968, Inoue and Orgel 1982, Acevedo and Orgel 1987). Nonenzymatic template-directed RNA polymerization, or primer extension, in which nucleotides are added to a primer when guided by a template sequence, has been intensively studied as a model of RNA copying. The success of this polymerization relies on the chemical activation of the mononucleotide phosphate groups. A wide range of phosphate activation groups has been explored to drive these polymerization reactions, including imidazoles, methyl-imidazoles, and oxyazabenzotriazoles (Weimann, Lohrmann et al. 1968, Vogel, Deck et al. 2005, Blain and Szostak 2014). Recently, our laboratory has demonstrated efficient copying of various short RNA templates using 5′-phosphoro-2-aminoimidazolide-(2AI) activated mononucleotides (2AImpN), and shown polymerization occurs through the *in-situ* generation of 5′-5′-imidazolium-bridged dinucleotides (Li, Prywes et al. 2017, Walton, Zhang et al. 2019) (Figure 1.2 showing the structures of 2AImpN and 2AImpN*2AIm). The bridged dinucleotide is the key intermediate for efficient template-directed chemical RNA polymerization, unlike the activated mononucleotide itself, which does not significantly contribute to primer extension (Duzdevich, Carr et al. 2021).

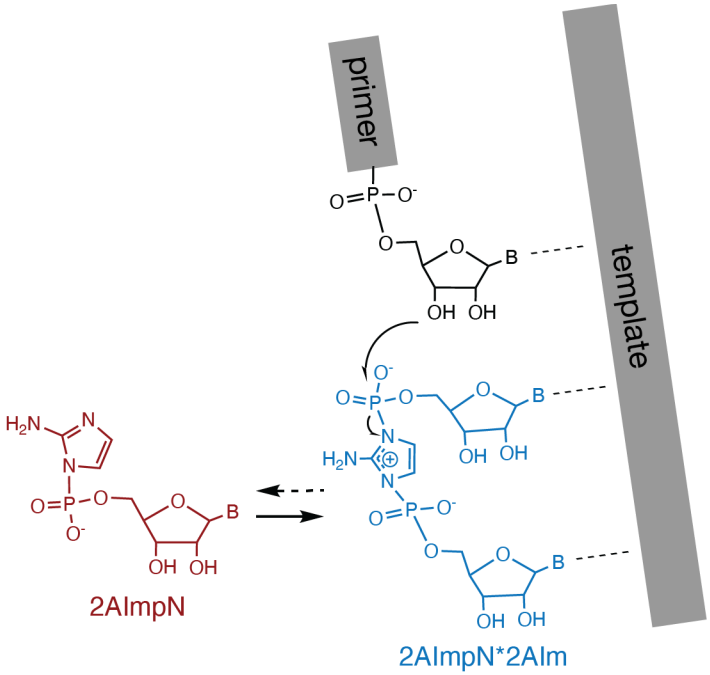


Figure 1.2. Schematic of the proposed mechanism of primer extension by 5'-5'-imidazolium-bridged dinucleotides. Two activated nucleotides (2AImpN) react to form a 5'-5'-imidazolium bridged dinucleotide intermediate (2AImpN*2AIm). Following the formation of the covalent intermediate, this bridged dinucleotide would bind the template and react with the primer, releasing an activated monomer as the leaving group.

RNA copying can be driven by the *in-situ* activation of both monomers and oligonucleotides. A ligation approach can overcome problems, such as copying through adenosine (A) and uridine (U) rich sequences or secondary structures (Joyce, Schwartz et al. 1987). Further, it remains attractive because it requires fewer reaction steps to copy templates and its enhanced fidelity (James and Ellington 1997). However, major challenges remain, including low yield and slow rate of template-directed RNA ligation (Rohatgi, Bartel et al. 1996, Prywes, Blain et al. 2016). There remains

possibility of achieving effective RNA copying by exploiting both ligation and primer extension with monomer addition (Szostak 2011).

1.3 Prebiotic nucleotide activation and reactivation

Previous work has relied on synthetically generated activated nucleotides, which are prone to irreversible hydrolysis and so have short lifetimes in solutions (Whitaker and Powner 2018). However, the implication has always been that on the early earth there would have been some mechanism by which the pool of activated nucleotides could have been actively replenished or maintained. A prebiotically plausible aqueous-phase chemical process that can activate the monomers *in situ* under conditions that are compatible with nonenzymatic RNA copying is needed. In search of such a chemical process, Sutherland and co-workers have recently reported a prebiotically plausible route to selective phosphate activation, without modification to the nucleobases (Mariani, Russell et al. 2018). The activation is achieved with methyl isocyanide—released upon the UV irradiation of transition metal complexes in an aqueous solution in the presence of primordial atmospheric gasses—aldehyde, and imidazole (Figure 1.3). Briefly, the reaction of isocyanide and aldehyde generates a nitrilium ion, which transiently activates the phosphate group. The resulting on-pathway intermediate, imidoyl phosphate, undergoes attack by imidazole or its derivatives to yield phosphorimidazolide. The potential prebiotic syntheses of the other precursors, including aldehydes (Ritson and Sutherland 2013, Patel, Percivalle et al. 2015), and imidazoles (Ferris, Sanchez et al. 1968, Oro, Basile et al. 1984, Vazquez-Salazar, Tan et al. 2017) have been previously described. Though their study is primarily focused on the characterization of the phosphorimidazolide with imidazole, they also show nucleotide activation with 2AI, which may be prebiotically accessible (Fahrenbach, Giurgiu et al. 2017).

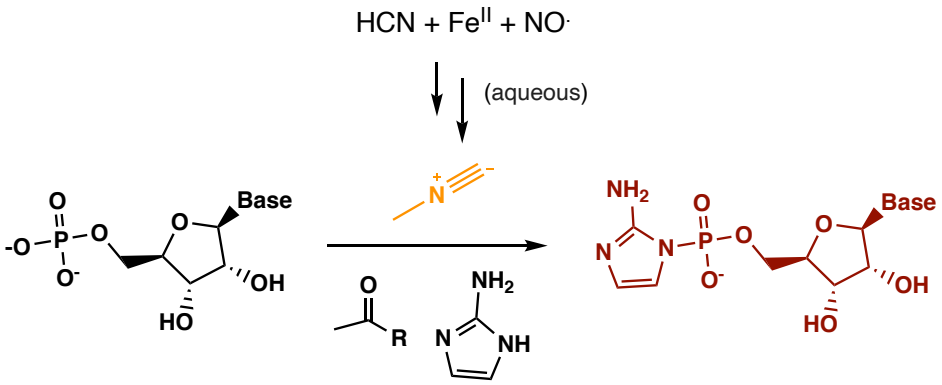


Figure 1.3. The prebiotically relevant nucleotide activation enables selective phosphate activation with various imidazole derivatives. The key ingredient, methyl isocyanide (MeNC, orange), could have been produced in a ferrocyanide and nitroprusside containing environment upon UV irradiation. Briefly, aldehydes react with isocyanide to produce a highly reactive nitrilium ion. It will then react with mononucleotides and be displaced by the activating group, 2-aminoimidazole, 2AI, forming an activated nucleotide that could be fed into primer extension reactions and hence copy RNA.

This isocyanide nucleotide activation chemistry prompts us to apply it to chemical RNA copying. However, a major hurdle to the compatibility of activation chemistry and RNA copying is that excess 2AI is required to drive nucleotide activation, but excess 2AI specifically inhibits nonenzymatic RNA polymerization by attacking the imidazolium-bridged dinucleotide intermediate **3** (Figure 1.2). In Chapter 2, we demonstrate the compatibility between prebiotically plausible nucleotide activation chemistry and nonenzymatic RNA polymerization. We report a novel pathway that activates RNA nucleotides in a manner compatible with template-directed nonenzymatic copying. We show that this pathway, which we refer to as bridge-forming activation, selectively

yields the reactive imidazolium-bridged dinucleotide intermediate required for copying. Ultimately, we anticipate it will lead to an experimental demonstration of continuous rounds of nucleotide activation in the context of primer extension, enabling more realistic simulations of RNA propagation. In Chapter 3, we report a prebiotically plausible pathway starting from chemical nucleotide activation, to bridged dinucleotide formation, and ultimately primer extension under mutually compatible conditions. Future studies shall explore the possibility of bridged formation between monomers and oligonucleotides, which might effectively enhance primer extension and ligations.

1.4 Ribozyme-catalyzed RNA copying

Under the RNA world hypothesis, one must postulate that a library of RNA strands with various sequences formed on primitive earth and this family of sequences can support its own replication. Further, the emergence of RNA catalysts will greatly improve the yield and the fidelity of RNA polymerization relative to nonenzymatic sequence copying, providing selective advantages for early cellular life. Evolved ribozymes that catalyze template-directed ligation and polymerization reactions with nucleoside triphosphates have been successfully identified (Johnston, Unrau et al. 2001). To connect the purely nonenzymatic to ribozyme-catalyzed RNA template copying during the origin of life, an alternative approach to RNA-templated RNA copying that combines ribozyme catalysis with intrinsically reactive RNA substrates, such as imidazolides, has also been demonstrated (Walton, DasGupta et al. 2021). However, typical chemical activation of RNA substrates is incompatible with ribozyme catalysis, it remains unclear how prebiotic systems generated, and sustained pools of activated building blocks for increasingly complex RNA. *In-situ* activation of RNA substrates under potentially prebiotic conditions amenable to RNA catalysis is of great interest. Diamidophosphate (DAP) and imidazole have been demonstrated to drive the formation of

2′,3′-cyclic phosphate RNA mono- and oligonucleotides in frozen water-ice, enabling iterative enzymatic assembly of long RNAs over the course of three months (Song, Jimenez et al. 2021). Although the results are not discussed in detail here, we have demonstrated *in-situ* activation of RNA substrates by isocyanide-mediated chemistry enabling efficient RNA catalyzed ligation within a few hours.

1.5 Primitive cell membranes

Membranes provide the major mechanism for compartmentalization and mediate the chemical fluxes between the cell and its surroundings in modern biology. Without the highly evolved membrane proteins, primitive life must have possessed essential features, such as nutrient transport and basic cellular behaviors like growth and division. Modern membranes composed of diacyl amphiphilic molecules are unsuitable as primitive membranes due to their inertness in the absence of enzymes. Single-chained amphiphiles are considered to constitute early membranes. First, fatty acids and their derivatives can be synthesized abiotically under simulated prebiotic conditions via Fischer-Tropsch-type reactions (McCollom, Ritter et al. 1999) and have been detected in carbonaceous chondrite meteorites (Lawless and Yuen 1979, Yuen, Lawless et al. 1981). Secondly, fatty acids can form bilayer vesicles that are similar to modern cell membranes composed of diacyl phospholipids when the solution pH is near the apparent pKa of the corresponding fatty acid. Intriguingly, clay surfaces promote the vesicle formation of fatty acids as well as catalyzing RNA polymerization, which suggests a potential mechanism for improved RNA encapsulation due to the colocalization of vesicles and RNA (Ferris, Hill et al. 1996, Ferris 2002).

Single-chained amphiphiles, fatty acids and their related molecules, have shown the capability to grow and divide without complex biochemical machinery. Simple osmotic pressures across vesicle membranes could lead to vesicle growth (Chen, Roberts et al. 2004, Chen and Szostak 2004). In

addition, spontaneous membrane growth and division have been demonstrated under alternation of environmental conditions, evaporation, and rain (Budin, Debnath et al. 2012, Budin and Szostak 2013). Further, selective permeable membranes also endow vesicles with properties of competition. For example, preferred growth within RNA-containing vesicles at the expense of empty ones was observed due to the differences in osmotic pressures exerted by encapsulated RNA (Chen, Roberts et al. 2004).

However, the low solute permeability of primitive membranes, on one hand, is desired to prevent the loss of the encapsulated contents. On the other hand, it stands in the way of efficiently acquiring nutrients under nutrient-rich circumstances. A protocell will have to reside within a pool of nutrients for minutes to hours to acquire them. Endocytosis has been previously demonstrated in the phospholipid membranes via various pathways (Lipowsky 1999, Peterlin, Arrigler et al. 2009, Zong, Ma et al. 2017). However, the ability for primitive membranes to endocytose is yet to be definitively shown. In Chapter 4, we demonstrate how primitive cells composed of fatty acids endocytose via a purely physicochemical process. A simultaneous reduction in volume and increase in surface area allows the larger molecules to be imported into protocells bypassing permeability limits, mimicking the process of endocytosis in complex eukaryotes. The process takes several seconds. The protocells, if removed from the nutrient source, continue to retain the internalized solute, thereby providing time to then absorb the nutrients. Further, we construct a numerical model that captures such out-of-equilibrium topological changes. Taken together, we believe this marks an important step forward toward the understanding and development of primitive cells.

Their probable involvement in the origin of life has been demonstrated by nonenzymatic primer extensions and RNA catalysis within vesicles composed of single-chained amphiphiles (Chen, Salehi-Ashtiani et al. 2005, Adamala and Szostak 2013). The nonenzymatic template-directed generation of primer extension products long enough to encode active ribozymes inside model

protocells. In this scenario, the propagation of functional RNAs within reproducing compartments leads to replicating protocells, ultimately giving rise to modern cellular life.

Alternative primitive membranes

Protocells composed of fatty acids and its related molecules can grow, divide, and acquire nutrients from the environments, supporting the life cycles of a protocell. However, it lacks stability towards a broad range of pHs and metal ion concentrations. Alternative prebiotically plausible membrane compositions that can withstand a wide variety of chemical conditions, including the concentrations of metal ions necessary for the folding and activity of biomolecules, would be advantageous. Recently, prebiotically plausible monoacyl cyclophospholipids have shown a remarkable ability to self-assemble in a broad range of pH and high concentrations of metals (Toparlak, Karki et al. 2019). Interesting, isocyanide-driven activation chemistry also converts reactive monoacylglycerol phosphates into inert cyclophospholipids (Bonfio, Caumes et al. 2019), highlighting the compatibility of such membrane constituent with activation chemistries that enables nonenzymatic RNA polymerization as well as supporting their potential role as major constituents of protocellular compartments.

Chapter 2

Prebiotically Plausible Activation Chemistry Compatible with Nonenzymatic RNA Copying

Reproduced and modified with permission from:

Zhang, S. J., Duzdevich, D., and Szostak, J. W. (2020) Potentially Prebiotic Activation Chemistry Compatible with Nonenzymatic RNA Copying. *Journal of the American Chemical Society* 142 (35), 14810-14813. S.J.Z., D. Duzdevich, and J.W.S. designed research; S.J.Z. performed all of the experiments and analyzed data; and S.J.Z., D. Duzdevich, and J.W.S. wrote the paper.

Abstract

The nonenzymatic replication of ribonucleic acid (RNA) oligonucleotides may have enabled the propagation of genetic information during the origin of life. RNA copying can be initiated in the laboratory with chemically activated nucleotides, but continued copying requires a source of chemical energy for in situ nucleotide activation. Recent work has illuminated a potentially prebiotic cyanosulfidic chemistry that activates nucleotides, but its application to nonenzymatic RNA copying remains a challenge. Here we report a novel pathway that enables the activation of RNA nucleotides in a manner that is compatible with template-directed nonenzymatic polymerization. We show that this pathway selectively yields the reactive imidazolium-bridged dinucleotide intermediate required for nonenzymatic template-directed RNA copying. Our results will enable more realistic prebiotic chemical simulations of RNA copying based on continuous in situ nucleotide activation.

Main text

RNA is a leading candidate for the primordial genetic polymer because of its potential to act as both a hereditary and functional biomolecule (Szostak, Bartel et al. 2001, Lilley and Eckstein 2007, Wachowius, Attwater et al. 2017). The emergence of life in the RNA World would have required nonenzymatic RNA replication to propagate genetic information prior to the emergence of more elaborate ribozyme-catalyzed replication machinery (Szostak 2017, Joyce and Szostak 2018, Horning, Bala et al. 2019). Primer extension is a model of RNA copying in which nucleotides **1** are added to a primer when guided by a template sequence (Figure 2.1) (Sulston, Lohrmann et al. 1968, Orgel 1992, Orgel 2004). Nonenzymatic primer extension relies on activation of the

mononucleotide phosphate groups (Weimann, Lohrmann et al. 1968, Westheimer 1987, Vogel, Deck et al. 2005, Prywes, Blain et al. 2016, Li, Prywes et al. 2017).

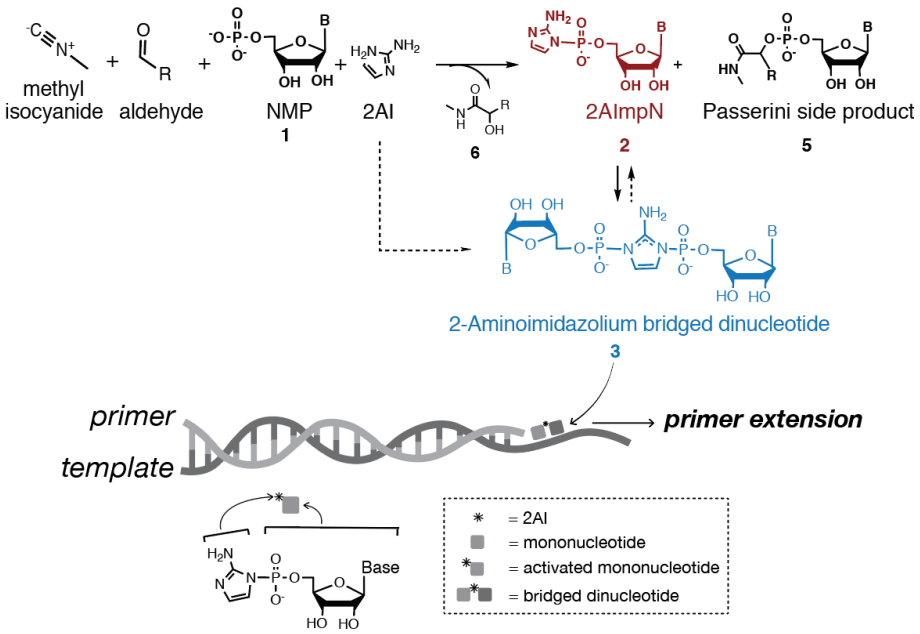


Figure 2.1. The components of nonenzymatic RNA primer extension. Nonenzymatic template-directed RNA polymerization at the 3′-end of a primer proceeds via **3**, which forms spontaneously in a pool of chemically activated nucleotides **2**. Isocyanide nucleotide activation chemistry is incompatible with primer extension due to the required excess 2AI, which inhibits accumulation of **3**.

Our laboratory has demonstrated efficient copying of various short RNA templates using 5′-phosphoro-2-aminoimidazolide-(2AI) activated ribonucleotides (2AImpN **2**) (Li, Prywes et al. 2017), and shown that polymerization proceeds through the *in situ* generation of 5'-5'-imidazolium-bridged dinucleotides **3** (Walton and Szostak 2016) (Figure 2.1). The superiority of 2AI as a phosphate activating group over other imidazole derivatives is due at least in part to the higher accumulation and greater stability of 5'-5'-imidazolium-bridged dinucleotides (Walton and Szostak 2017). The

bridged dinucleotide forms spontaneously, accumulating to a low level in a pool of chemically activated nucleotides. Activated mononucleotides hydrolyze to generate free 2AI, which in turn attacks the bridged dinucleotide to yield two 2AImpNs **2** (Walton and Szostak 2017, Zhang, Walton et al. 2018). Bridged dinucleotides also decay through hydrolysis—yielding one 2AImpN **2** and one nucleoside monophosphate (NMP **1**).

A prebiotically relevant model requires *in situ* activation that is also compatible with primer extension (Prywes, Blain et al. 2016, Li, Prywes et al. 2017, Whitaker and Powner 2018). Sutherland and co-workers have recently reported a prebiotically plausible route to phosphate activation with methyl isocyanide, aldehyde, and imidazole (Mariani, Russell et al. 2018) in a pH regime that is potentially compatible with primer extension (Walton and Szostak 2017, Mariani, Russell et al. 2018). This prompted us to seek conditions under which prebiotically plausible nucleotide activation chemistry can be applied to template-directed nonenzymatic RNA polymerization (Figure 2.1).

A major hurdle to the compatibility of activation chemistry and RNA copying is that excess 2AI is required to drive nucleotide activation, but excess 2AI specifically inhibits nonenzymatic RNA polymerization by attacking the imidazolium-bridged dinucleotide intermediate **3** (Figure 2.1). We report a new pathway that both circumvents this issue and yields significantly higher concentrations of bridged dinucleotides than the spontaneous self-reaction of activated mononucleotides **2** (Walton and Szostak 2017). Our findings point to new experiments with continuous activation of nucleotides during RNA copying.

As a first step to combining isocyanide activation with primer extension, we sought reaction conditions compatible with both. Optimal primer extension requires Mg^{2+} and mildly basic buffer (pH ~8) (Szostak 2012, Walton and Szostak 2017). We examined the effects of Mg^{2+} concentration and pH on the activation of NMPs **1** to 2AImpN **2** using isocyanide (Figure S2.1) and acetaldehyde. All four canonical ribonucleotides were activated under primer extension conditions (Figure S2.2a).

However, an undesirable full Passerini reaction product **5** (Sutherland, Mullen et al. 2008), which depletes the starting NMP pool (supplementary text, Scheme S2.1), also formed in addition to the desired activated nucleotides (Figures S2.3, Tables S2.1, S2.3). In a screen of longer chain aldehydes and ketones in place of acetaldehyde, 2-methylbutyraldehyde (2MBA) decreased the formation of **5** from 12% to 3%, while increasing the yield of 2AImpN **2** from 31% to 81% (Figure S2.4, Table S2.2 with 30 mM Mg^{2+}). The higher yield of **2** may stem from reduced hydrolysis of the imidoyl intermediate **4** without affecting the 2AI attack on the phosphate group.

Although the above optimizations define reaction conditions that could enable compatibility with primer extension, there remained a significant obstacle. The high concentration of 2AI required for NMP activation (Table 2.1) was expected to prohibit accumulation of the imidazolium bridged dinucleotides **3** necessary for RNA copying by driving the equilibrium toward 2AImpN **2** (Figure 2.1) (Walton and Szostak 2017).

Table 2.1. Yields of 2AImpA 2 and bridged dinucleotide 3 at different 2AI concentrations measured by NMR spectroscopy.

[2AI] mM	Activation (**2**) yield %	Bridged dinucleotide (**3**) yield %
10	6±2	1.8±0.5
20	13.4±0.6	2.8±0.9
50	28.1±0.2	2.1±0.2
100	33.0±0.5	0.8±0.4
200	48±2	0
400	34±1	0
800	19±1	0

All reactions were carried out using AMP (10 mM), acetaldehyde (200 mM), methyl isocyanide (200 mM), 2AI (varied), Mg^{2+} ($MgCl_2$, 10 mM), and HEPES (200 mM) at pH 8.0 and t = 6 hours. Errors are standard deviations of the mean, n = 3 replicates.

Confirming this required a primer extension assay compatible with isocyanide activation chemistry. Because the isocyanide chemistry modified the fluorophores used for primer labelling (Figure S2.5), we developed a post-labeling strategy for visualization of primer extension (supplementary text, Figures S2.6, S2.7). Using a standard primer extension reaction in which the template sequence is 5′-CCG-3′, we found that the excess 2AI (200 mM) required for efficient activation severely inhibits primer extension in the presence or absence of activation chemistry (Figure S2.8). Thus, the requirement for excess 2AI is a fundamental incompatibility between primer extension and *in situ* activation with isocyanide.

Reflecting on the overall primer extension pathway—from nucleotides **1**, via activated nucleotides **2**, to the bridged dinucleotides **3** that actually promote the elongation of the primer—we asked whether the isocyanide activation chemistry might be relevant to the formation of the bridged dinucleotide as well as activated monomers. We postulated that the 2AI of pre-activated 2AImpN **2** could act as a nucleophile to displace the imidoyl leaving group of intermediate **4**, directly forming the bridged imidazolium dinucleotide **3** necessary for primer extension (Figure 2.1). We therefore introduced 2AImpN **2** to a mixture of isocyanide, aldehyde, and AMP **1** without any free 2AI. We found that not only did the bridged dinucleotide species **3** form, but it accumulated to a significantly higher yield than is typical through the self-reaction of activated monomers in the absence of activation chemistry. For an equimolar mix of AMP and 2AImpA in the presence of isocyanide activation, ^{31}P NMR spectra show 16% bridged dinucleotide **3** at t = 229 min. (the timepoint at which the concentration of bridged dinucleotide peaks), compared with only 2% in its absence (Figure 2). To differentiate this scenario from the one in which excess 2AI drives NMP **1** activation, we call it bridge-forming activation (supplementary text).

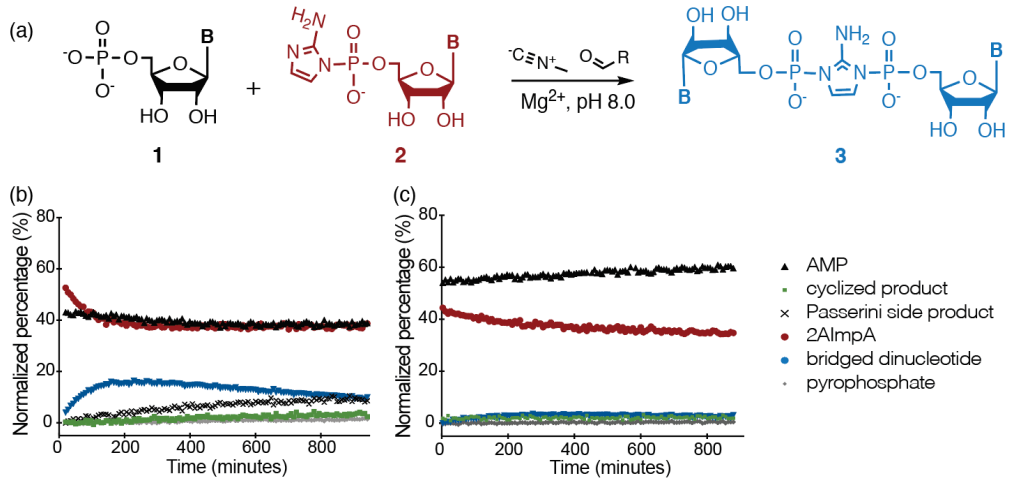


Figure 2.2. Bridge-forming activation. (a) Prebiotically plausible chemistry drives bridged dinucleotide formation. Analyses of the reaction over the course of 15 hours by ^{31}P-NMR (b) with and (c) without bridge-forming activation. The relative percentage of each species is calculated based on the corresponding peak integration normalized to the number of phosphorus atoms. Reaction conditions: AMP (5 mM), 2AImpA (5 mM), Mg^{2+} (30 mM, $MgCl_2$), HEPES (200 mM) at pH 8 with or without methyl isocyanide (200 mM) and 2MBA (200 mM).

We expect that in a prebiotically plausible scenario for RNA copying, the ratio of activated to unactivated nucleotides would vary with time. We find that bridge-forming activation functions across a broad range of ratios, with bridged dinucleotide **3** detected in every case in which activated mononucleotide **2** is present (Figure S2.9a-c). Bridged dinucleotide **3** did not form in the absence of activated monomer **2** (Figure S2.9d).

Interestingly, the treatment of 100% activated mononucleotides **2** with the bridge-forming activation reagents also efficiently yielded bridged dinucleotide **3** with little accompanying hydrolysis (Figure S2.10): from t = 12 min. to t = 135 min., only 1% of the mononucleotides hydrolyzed whereas the bridged dinucleotide **3** yield was 40%. These observations suggest a significant contribution from a novel pathway in which the activation chemistry directly mediates the bridging

of two already-activated nucleotides (rather than only the bridging of pairs of activated and unactivated nucleotides). The mechanism of this pathway is under investigation.

We next considered whether bridge-forming activation shows any preference for 2AI over 2-methylimidazole (2MI), the historically most common activating imidazole, given there is no reason to assume that superiority of 2AI-activated nucleotides **2** for primer extension can be extended in the context of bridge-forming chemistry. Treatment of an equimolar mixture of AMP and 2MImpA **7** with bridge-forming activation did yield bridged dinucleotide **8**, though markedly less than with the use of 2AImpN. Without bridge-forming activation, no detectable bridged dinucleotide **8** formed (Figure S2.11). The significant difference in bridged dinucleotide accumulation between 2AI- and 2MI-activated mononucleotides led us to consider how they would behave together. If bridge-forming activation is promiscuous, then treatment of an equimolar mixture of 2AI- and 2MI-activated nucleotides should yield both imidazole bridged dinucleotides. However, the reaction only yielded one detectable bridged species: 2-aminoimidazolium bridged dinucleotide **3** (Figure S2.12). This demonstrates that bridge-forming activation is efficiently selective towards 2AI over 2MI in a nucleotide concentration regime that is functional in primer extension.

Encouraged by these results, we sought to apply bridge-forming activation to primer extension. To copy the 5′-GCC-3′ template, various concentrations of 2-AI activated C and G mononucleotides were mixed with unactivated C and G and treated or not treated with bridge-forming activation (Figure 2.3). As a control we performed primer extension with 5 mM each 2AImpG and C and observed a distribution of +1, +2, and +3 products to serve as a baseline for other experiments (Figure 2.3a). The addition of 10 mM NMPs inhibited the reaction because unactivated mononucleotides compete for the binding sites of bridged dinucleotides (compare Figure 2.3a to Figure 2.3b). The application of bridge-forming activation increased product yield, with 43% +3 products compared to 21% without the bridge-forming activation (Figure 2.3b,c). This

distribution of products from an equimolar ratio of activated and unactivated mononucleotides plus bridge-forming activation is comparable to that found with the use of 20 mM pure activated mononucleotides (Figure 2.3d). Finally, applying bridge-forming activation to 10 mM pure activated mononucleotides, with no initial unactivated nucleotides, resulted in even more +3 product (53%) (Figure 2.3e). Note that in these experiments the product length is template-limited, because there are no template bases beyond the +3 position. These experiments demonstrate the compatibility of isocyanide-based nucleotide activation with nonenzymatic RNA polymerization.

Bridge-forming activation provides several advantages for primer extension. It requires lower mononucleotide concentrations (1.5-5 mM) to generate appreciable proportions of bridged dinucleotide than required by spontaneous bridging (10-100s mM range) or direct activation with free 2AI (50 mM-400 mM NMP). The lower mononucleotide concentration combined with the higher proportion of bridged dinucleotides also raises the possibility that the Mg^{2+} concentration can be further reduced. High Mg^{2+} concentrations are notoriously problematic for primer extension, causing bridged dinucleotide hydrolysis (Walton, Pazienza et al. 2019), monomer cyclization (Breslow and Huang 1991, Harada and Orgel 1991, Szostak 2012), and template degradation (Szostak 2012, Adamala and Szostak 2013). Finally, the reaction is effective over a range of unactivated to activated mononucleotide ratios (Figure S2.9), a key environmental variable that would be expected to fluctuate in any plausible prebiotic scenario.

Bridge-forming activation provides several advantages for primer extension. It requires lower mononucleotide concentrations (1.5-5 mM) to generate appreciable proportions of bridged dinucleotide than required by spontaneous bridging (10-100s mM range) or direct activation with free 2AI (50 mM-400 mM NMP). The lower mononucleotide concentration combined with the higher proportion of bridged dinucleotides also raises the possibility that the Mg^{2+} concentration can be further reduced. High Mg^{2+} concentrations are notoriously problematic for primer extension,

causing bridged dinucleotide hydrolysis (Walton, Pazienza et al. 2019), monomer cyclization (Breslow and Huang 1991, Harada and Orgel 1991, Szostak 2012), and template degradation (Szostak 2012, Adamala and Szostak 2013). Finally, the reaction is effective over a range of unactivated to activated mononucleotide ratios (Figure S2.9), a key environmental variable that would be expected to fluctuate in any plausible prebiotic scenario.

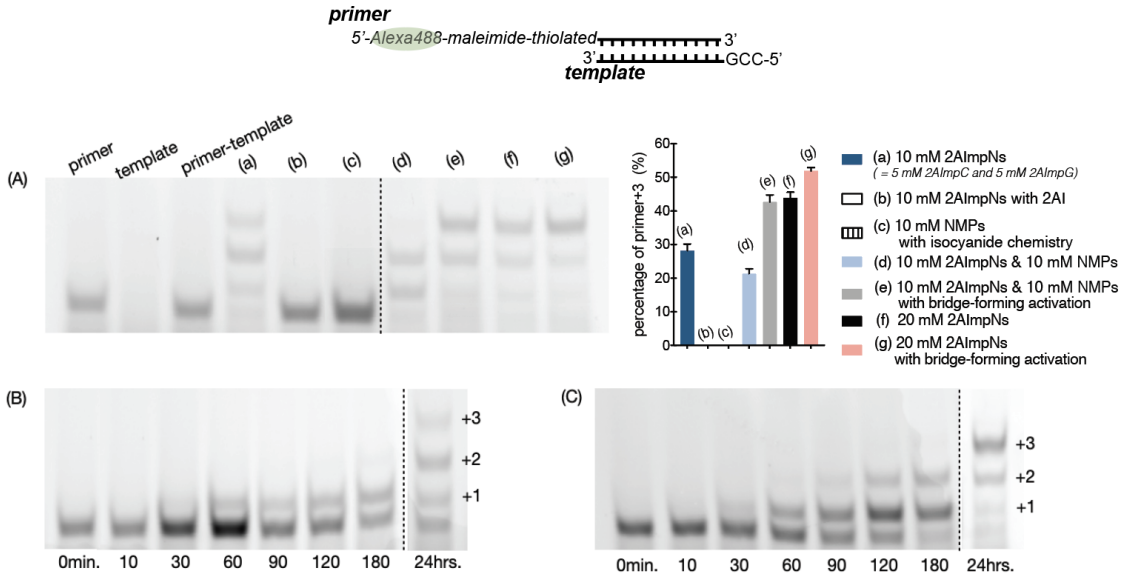


Figure 2.3. Primer extension using bridge-forming activation. (A) Products from primer extension reactions using (a) 10 mM 2AImpN (N=C and G), (b) in the presence of 200 mM 2AI, (c) 10 mM NMPs with isocyanide chemistry, (d) an equimolar mixture of 10 mM 2AImpNs and 10 mM NMPs without bridge-forming activation and (e) with bridge-forming activation, (f) 20mM 2AImpNs without bridge-forming activation and (g) with bridge-forming activation.

Reaction conditions: indicated quantities of 2AImpNs and NMPs were added to primer (1 μM, green text), template (1.5 μM, black text), HEPES (200 mM) pH 8.0, and Mg^{2+} (30 mM, $MgCl_2$) with isocyanide chemistry (grey and coral pink) or without (dark blue, light blue, black). Extension products were assayed by PAGE (Figure S13) at t = 24 hours. Error bars indicate standard deviations of the mean, n = 3 replicates. The time course of primer extension reaction using

equimolar mixture of **1** and **2** (B) with and (C) without bridge-forming activation were assayed. Positions of primer and +1 to +3 products are indicated. Reaction conditions as described in (d) and (e).

Although bridge-forming activation is compatible with primer extension and promotes the formation of the required intermediate, it depends on a source of previously activated mononucleotides. One possibility is that nucleotide activation might occur under partial dry-down conditions where all reactants including 2AI are at very high concentrations, followed by dilution to nucleotide concentrations sufficient for bridge-forming activation and primer extension, but with low enough free 2AI to minimize loss of the bridged dinucleotides. Additional processes that might sequester or degrade 2AI after efficient nucleotide activation should also be investigated. For example, UV radiation, the presumptive energy source for producing isocyanide, photodegrades 2AI on the order of days (Todd, Szabla et al. 2019) although the photodegradation rates of **2** and **3** remain unknown.

Life is not expected to emerge after a single round of template-directed primer extension. A highly desirable feature of bridge-forming activation is the potential re-activation of spent nucleotides for further rounds of polymerization. Previously identified activation chemistries lead to damaging side reactions that destroy both templates and substrates (Ferris and Hagan 1984, Biron and Pascal 2004, Leman, Orgel et al. 2004, Leman, Orgel et al. 2004, Powner, Gerland et al. 2009, Szostak 2012), whereas bridge-forming activation relies on RNA-compatible and specific reagents. Further work is needed to demonstrate nucleotide re-activation in the context of indefinitely continued primer extension.

Chapter 3

Freeze-thaw Cycles Enable a Prebiotically Plausible and Continuous Pathway from Nucleotide Activation to Nonenzymatic RNA Copying

Reproduced and modified with permission from:

Zhang, S. J., Duzdevich, D., Ding, D., & Szostak, J. W. (2022). Freeze-thaw cycles enable a prebiotically plausible and continuous pathway from nucleotide activation to nonenzymatic RNA copying. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(17). S.J.Z., D. Duzdevich, and J.W.S. designed research; S.J.Z. and D. Ding performed research; S.J.Z. analyzed data; and S.J.Z., D. Duzdevich, and J.W.S. wrote the paper.

Abstract

Nonenzymatic template-directed RNA copying using chemically activated nucleotides is thought to have played a key role in the emergence of genetic information on the early Earth. A longstanding question concerns the number and nature of different environments that might have been necessary to enable all of the steps from nucleotide synthesis to RNA copying Here we explore three sequential steps from this overall pathway: nucleotide activation, synthesis of imidazolium-bridged dinucleotides, and template-directed RNA copying. We find that all three steps can take place in one reaction mixture undergoing multiple freeze-thaw cycles. Recent experiments have demonstrated a potentially prebiotic methyl isocyanide-based nucleotide activation chemistry. However, the original version of this approach is incompatible with nonenzymatic RNA copying because the high required concentration of the imidazole activating group prevents the accumulation of the essential imidazolium-bridged dinucleotide. Here we report that ice eutectic phase conditions facilitate not only the methyl isocyanide-based activation of ribonucleotide 5'-monophosphates with stoichiometric 2-aminoimidazole, but also the subsequent conversion of these activated mononucleotides into imidazolium-bridged dinucleotides. Furthermore, this one pot approach is compatible with template-directed RNA copying in the same reaction mixture. Our results suggest that the simple and common environmental fluctuation of freeze-thaw cycles could have played an important role in prebiotic nucleotide activation and nonenzymatic RNA copying.

Introduction

Template-directed nonenzymatic RNA copying requires activated nucleotides. A range of phosphate-activating groups has been explored to chemically activate nucleotides for such polymerization reactions (Weimann, Lohrmann et al. 1968, Vogel, Deck et al. 2005). Our laboratory has demonstrated efficient copying of several short RNA templates using 2-aminoimidazole (2AI) activated ribonucleotides (2AImpN, abbreviated as *pN) (Li, Prywes et al. 2017), and shown that polymerization proceeds predominantly through spontaneously generated 5'-5'-imidazolium-bridged dinucleotides (Walton and Szostak 2016). Template copying is greatly enhanced in the presence of activated downstream helper oligonucleotides, through the formation of monomer-bridged-oligonucleotide species which both bind more tightly to the template and are intrinsically more reactive (Prywes, Blain et al. 2016, Li, Prywes et al. 2017, Ding, Zhou et al. 2022).

Sutherland and co-workers recently reported a potentially prebiotic route to selective phosphate activation with methyl isocyanide (MeNC), acetaldehyde, and 2AI (Mariani, Russell et al. 2018). MeNC could have been produced in a ferrocyanide- and nitroprusside-containing environment upon UV irradiation (Mariani, Russell et al. 2018). A major hurdle to the compatibility of this activation chemistry with RNA copying is that activation requires an excess of 2AI to drive the reaction forward; however, excess 2AI prevents nonenzymatic RNA polymerization by shifting the equilibrium between activated monomers and imidazolium-bridged dinucleotides in favor of activated monomers (Walton and Szostak 2016, Walton and Szostak 2017). We have previously identified a variant of isocyanide activation chemistry, named bridge-forming activation, that circumvents this issue by directly generating bridged dinucleotides from activated mononucleotides (Zhang, Duzdevich et al. 2020). The resulting high concentration of bridged dinucleotides increases the yield of primer extension products. Although we were able to demonstrate the compatibility of bridge-forming activation with nonenzymatic RNA polymerization, bridge-forming activation

depends on the presence of already activated mononucleotides. To begin to address the challenge of initial nucleotide activation under conditions that do not inhibit bridged dinucleotide accumulation, we here explore the effects of repeated freeze-thaw cycles on activation chemistry (Monnard, Deamer et al. 2004, Menor-Salvan and Marin-Yaseli 2012).

Under partially frozen conditions, solutes are excluded from growing ice crystals and become concentrated in the interstitial liquid solution. This phenomenon of eutectic phase concentration has been demonstrated to facilitate the synthesis of nucleotide precursors (Bada, Bigham et al. 1994) as well as the untemplated condensation of activated nucleotides into oligomers (Monnard, Kanavarioti et al. 2003). Here we report the *in situ* MeNC-mediated activation of mononucleotides with stoichiometric 2AI by repeatedly subjecting the solution to eutectic phase freezing, which temporarily brings the reactants into close physical proximity and obviates the need for the large excess of 2AI otherwise required to drive activation in relatively dilute liquid phase reactions. We find that all four canonical ribonucleotides (as well as helper oligonucleotides) can be activated with this approach, which also yields high concentrations of bridged species. Consequently, we were also able to show that the reaction is compatible with template-directed nonenzymatic RNA copying. The geochemically plausible conditions that we have identified enable a potentially prebiotic pathway that begins with ribonucleotide 5'-monophosphates, generates 2AI activated nucleotides and then the essential bridged dinucleotides, and finally yields template-directed primer extension products, all in one reaction mixture.

Results

Addressing the incompatibility between template-directed primer extension and the initial *in situ* activation of mononucleotides requires a plausible pathway to nucleotide activation using a concentration of 2AI comparable to the concentration of nucleotides, rather than the large excess currently required (Figure 3.31A). We reasoned that the ice eutectic phase as a medium for activation

could temporarily increase the effective local concentration of 2AI, thereby alleviating the requirement for an absolute excess of 2AI relative to nucleotides (Figure 3.1B). To begin exploring the effect of ice eutectic phase concentration on nucleotide activation, we treated 10 mM adenosine monophosphate (pA) with 10 mM 2AI, 100 mM 2-methylbutyraldehyde (2MBA), and 200 mM MeNC, at pH 8 under both room temperature and ice eutectic phase conditions (Zhang, Duzdevich et al. 2020). With these reactant concentrations, and incubation at room temperature, only 2.0±0.1 mM activated nucleotides (approximately 20% of the 10 mM total available nucleotides) and only 0.11±0.01 mM bridged dinucleotides were formed after 24 hours (Figure 3.1C). In contrast, using the same concentrations of reactants but subjecting the mixture to ice eutectic phase concentration at -13 °C for 24 hours yielded 50% total activation, consisting of 3.0±0.1 mM activated nucleotides and 1.0±0.1 mM bridged dinucleotides. The ice eutectic phase not only increases the yield of activated nucleotides but also acts as a storage mechanism, because activated nucleotides are less susceptible to hydrolysis at low temperatures (Figure S3.1C-D), which further aids in the accumulation of bridged dinucleotides (Figure 3.1C).

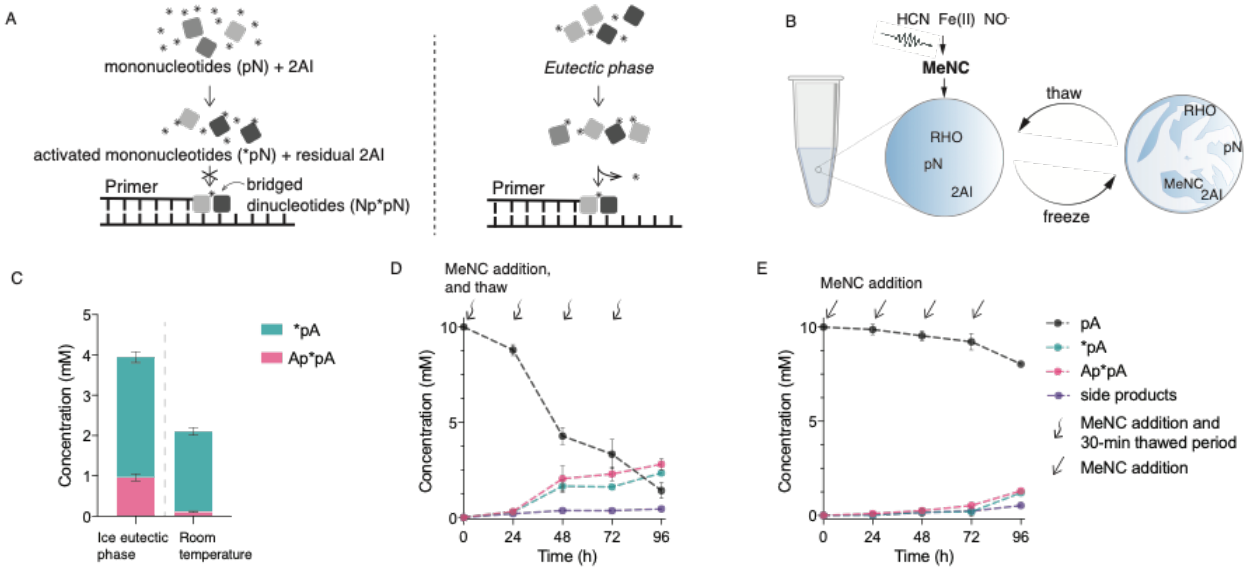


Figure 3.1. Activation in the ice eutectic phase enhances the yield of activated nucleotides (*pN) and promotes formation of bridged dinucleotides (Np*pN). (A) Scheme showing role of ice eutectic concentration. Left: At room temperature, the excess 2AI needed to drive nucleotide activation inhibits bridged dinucleotide accumulation. Right: activation in ice eutectic phase requires only equimolar 2AI and ribonucleotides, allowing bridged dinucleotide accumulation and template copying. (B) Solutes are concentrated in the ice eutectic phase. (C) Nucleotide activation in the ice eutectic phase (left) and at room temperature (right) with one round of MeNC addition and equimolar initial nucleotides and 2AI (see also Figure S3.3A). (D) Cycles of MeNC addition and eutectic freezing drive continued nucleotide activation. Curvy arrows: addition of MeNC after each thawing of the reaction mixture. Note that the reaction mixture is frozen as eutectic ice for the entire reaction course except during the brief thawing steps (~30 minutes at room temperature for the mixture to thaw completely) (see also Figure S3.3B). (E) Activation is inefficient at room temperature in the absence of ice eutectic freezing. The straight arrows indicate the addition of MeNC at room temperature (see also Figure S3.3C).

The possibility that in a prebiotic environment MeNC could have been released periodically in response to day/night cycles of UV irradiation(Sutherland, Mullen et al. 2008) prompted us to test multiple cycles of eutectic phase concentration and MeNC delivery. Instead of adding 200 mM MeNC from the start, 50 mM MeNC was introduced to the solution at the beginning of each of four sequential freeze-thaw cycles. This approach further increased nucleotide activation to 75%, yielding 2.33±0.01 mM activated nucleotides and 2.8±0.3 mM bridged dinucleotides (Figure 3.1D). By comparison, periodic addition of MeNC at room temperature without eutectic freezing yielded only 1.30±0.6 mM activated nucleotides and 0.7± 0.6 mM bridged dinucleotides, or 27% total

activation (Figure 3.1E). A range of pHs was examined to rule out the possibility that freezing-induced changes in pH might account for the increase in activation (Figure S3.2).

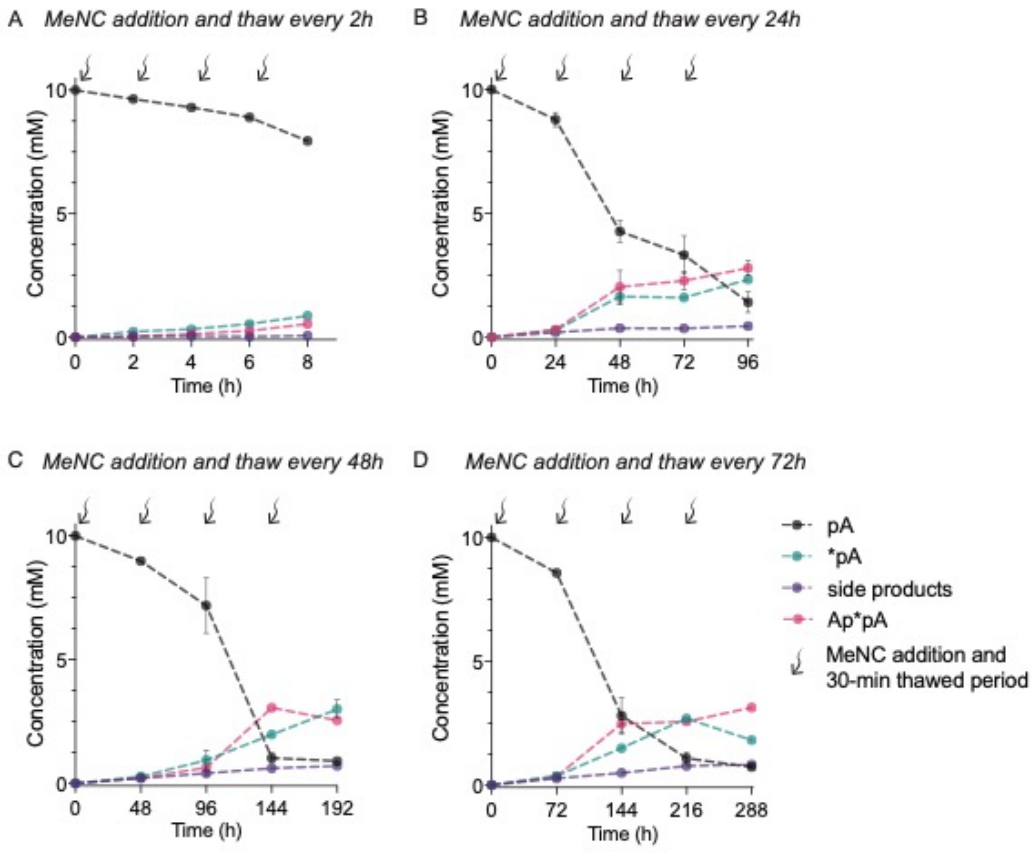


Figure 3.2. The frequency of MeNC delivery at long time intervals does not affect the overall yield of activated products. The yields of nucleotide activation were measured after each of four freeze-thaw cycles with periodic addition of MeNC. (A) Yields of activated species are low with time intervals on the order of hours in between MeNC additions. (B-D) However, lengthening of the time interval beyond 24 hours does not significantly affect the final yield. See also Figure S5. *Reaction conditions*: 10 mM pA, 10 mM 2AI, 100 mM 2MBA, 50 mM MeNC, 30 mM $MgCl_2$, 50 mM Na^+-HEPES pH 8, and subsequent periodic addition of MeNC in three aliquots of 50 mM. After the last

addition of MeNC, the solution was left under eutectic ice phase conditions for the indicated time, then the solution was thawed for room temperature NMR measurement.

We then varied the frequency of MeNC addition to the system and evaluated the effect on activated nucleotide yield. We found that an interval of roughly one day between MeNC additions led to efficient activation (Figure 3.2A-B). Further increases in the interval, up to several days, did not significantly change the activation yield (Figures 3.2 & S3.2 (B) to (D)). This not only indicates the robustness of the system in response to changes in the timing of activation but again demonstrates that freezing is a storage mechanism in our system: the accumulation of bridged dinucleotides is consistent with the expectation that activated mononucleotides and bridged dinucleotides should not readily hydrolyze in the ice eutectic state (Figure S3.1) (Kanavarioti, Monnard et al. 2001, Walton and Szostak 2017). Furthermore, the presence of high concentrations of bridged dinucleotides in these experiments suggests that in addition to initial nucleotide activation, bridge-forming activation is also occurring. To confirm this, we incubated 10 mM pure pre-activated mononucleotides in the ice eutectic phase without additional activation chemistry and observed only 1.05±0.05 mM bridged dinucleotides (Figure S3.4), compared to 2.8±0.3 mM bridged dinucleotides in the presence of activation reagents. Thus, bridge-forming activation proceeds under ice eutectic phase conditions.

Copying mixed sequence RNA templates requires G, C, and U in addition to A. We therefore proceeded to test the activation of all four canonical nucleotides, separately and together, under ice eutectic phase conditions. We began with 10 mM of each individual ribonucleotide and found good yields of activated species with cytidine monophosphate (pC), adenosine monophosphate (pA) and uridine monophosphate (pU). In contrast, guanosine monophosphate (pG) exhibited poor activation (Figure 3.3A). We were able to rescue high efficiency activation by

treating a mixture of all four nucleotides (10 mM total) with activation chemistry under ice eutectic conditions (Figure 3.3B). Dilution experiments showed that the rescue is partially due to the decreased relative concentration of pG (from 10 mM to 2.5 mM, Figure 3C), suggesting that concentration-driven aggregation may interfere with the activation of pG. Our observation is another striking example of the desirability of an optimal degree of heterogeneity in prebiotic processes (Szostak 2011).

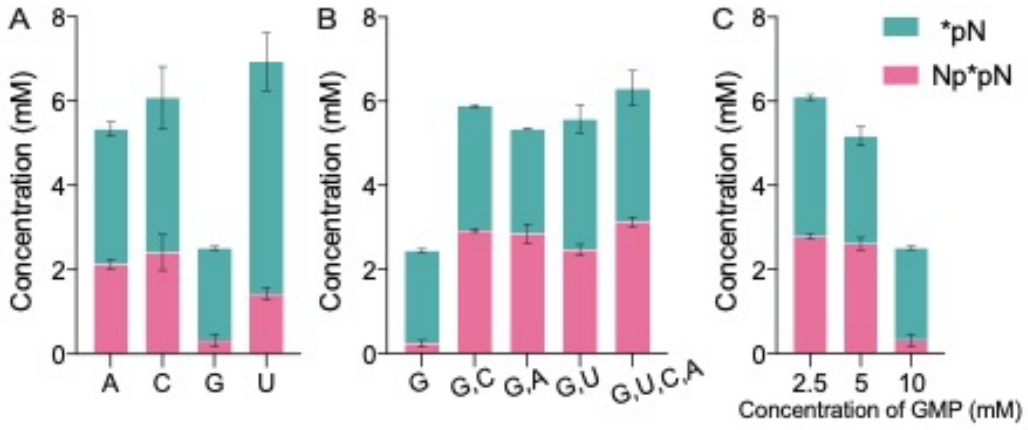


Figure 3.3. Efficient eutectic ice phase activation of all four canonical ribonucleotides. (A) Each of the four canonical ribonucleotides could be activated using the same approach as optimized for pA (Figure 1D). However, pG exhibited lower activation compared to the other three. (B) All four nucleotides can be activated efficiently when incubated simultaneously. Indicated concentrations are sum totals across all the nucleotides. (C) GMP activation yields decrease as GMP concentration is increased, suggesting that the low activation of GMP is partly due to its low solubility in aqueous phase (compare with Figure 3A). Total [pN] = 10 mM, 10 mM 2AI, 100 mM 2MBA, 50 mM MeNC, 30 mM $MgCl_2$, 50 mM Na^+-HEPES pH 8, and periodic addition of MeNC in three aliquots of 50 mM each. See also Figure S3.6.

Encouraged by these results, we sought to apply combined ice eutectic phase nucleotide activation and bridged species formation to template-directed primer extension (Figure 3.4A). First, we aimed to confirm that nucleotides activated using this approach behave the same way as synthetically prepared and purified activated nucleotides. We activated 5 mM each of CMP and GMP under ice eutectic phase conditions (Figure 3.3B) and then added the resulting mixture to a primer-template complex in solution phase. Under these conditions, the primer was extended (Figure 3.4C), just as in the control case with purified, synthetically prepared activated mononucleotides (Figure 3.4B). Thus, activation in the ice eutectic phase generates both activated mononucleotides and bridged dinucleotides which are functional in primer extension.

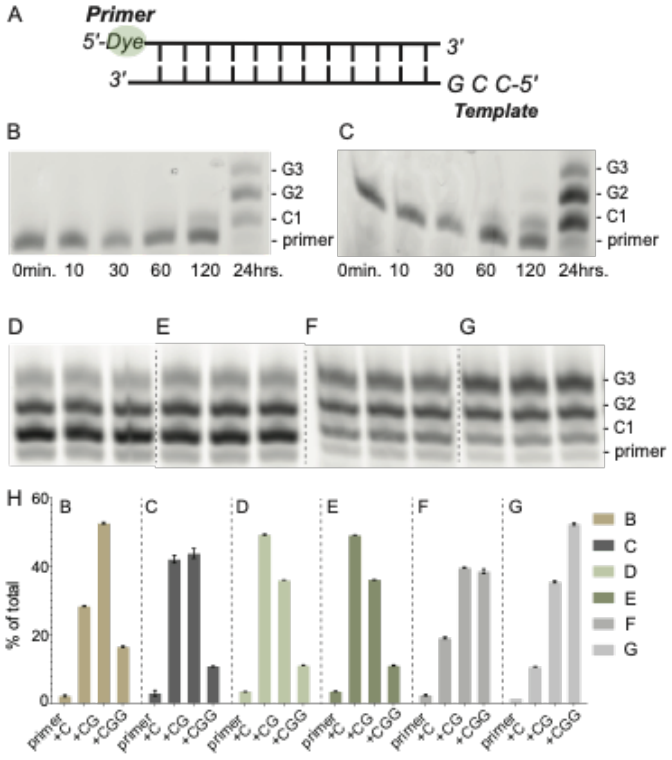


Figure 3.4. Ice eutectic phase activation enables a complete pathway from unactivated mononucleotides through activation to primer extension. (A) Schematic of primer-template complex. (B) Time course of primer extension using pure activated mononucleotides (5 mM *pG

and 5 mM *pC, room temperature). (C) Time course of primer extension reactions using pNs (N = C and G) generated with MeNC activation chemistry under ice eutectic phase. (D) Primer-template was added to already-activated mononucleotides generated by MeNC-based nucleotide activation under ice eutectic phase conditions, yielding extended primer (as in C). (E) Primer-template duplex was directly treated with pNs (N = C and G) and MeNC activation chemistry under ice eutectic phase. The products of the reaction were then incubated for an additional 24 hours at room temperature (F) without and (G) with bridge-forming activation. Each set of three lanes in (D-G) represents independent experimental replicates under the indicated condition. (H) Quantification of primer extension products. Positions of primer and +1 to +3-nucleotide extension products are indicated. Gel image of replicates for (B-C) is included in Figure S7.

Next, we asked whether *in situ* activation and primer extension can occur together in a single reaction mixture during freeze-thaw cycles. In this experiment the primer-template complex was included along with all other reagents from the start and exposed to cycles of ice eutectic phase activation. We found that primer extension was comparable to that observed with pure activated mononucleotides, or with sequential eutectic activation of reactants and subsequent solution phase primer extension (compare Figure 3.4E to Figures 3.4B, 3.4C and 3.4D). We then asked whether this *in situ* pathway could also be combined with our previously developed bridge-forming activation. In this process isocyanide and aldehyde are added to activated mononucleotides to generate enhanced levels of bridged dinucleotides from activated mononucleotides in solution, which thereby promote primer extension. We performed *in situ* nucleotide activation together with primer extension as in Figure 3.4E, but then further treated the mixture with bridge-forming activation at room temperature and incubated the mixture for an additional 24 hours. As expected, primer extension was even more efficient, with 52% of the total primer extended to +3 product (Figures 3.4

G&H) compared with 39% in the absence of additional bridge-forming activation (Figures 3.4 F&H). This experiment demonstrates that ice eutectic phase concentration enables nucleotide activation with bridge-forming activation and primer extension, all in one reaction mixture.

The experiments discussed so far used templates with at most two different template nucleobases. To test the limits of our approach with respect to prebiotic plausibility and heterogeneity, we next asked whether the ice eutectic phase activation of all four nucleotides could be combined with bridge-forming activation and primer extension to copy mixed sequence templates. Mixed-sequence templates containing AU are known to be challenging for nonenzymatic RNA copying (Wu and Orgel 1992), but activated downstream helper oligonucleotides greatly facilitate primer extension on templates containing all four bases (Prywes, Blain et al. 2016, Li, Prywes et al. 2017). We have previously shown that a primer can be extended by up to 7 nucleotides by copying the template sequence 3'-UACUCCGCG-5', if six activated trimers (UGA, GAG, AGG, GGC, GCG, CGC) are included in the reaction mix (Figure 3.5A). We repeated this experiment as a control, using 1.2 μM primer and 1.5 μM template in the presence of 5 mM of each activated mononucleotide (*pN, N=A, U, C, and G) and 0.5 mM of each of the six activated trimers. After 24 hours, 93% of the primer had been extended by one or more nucleotides, with approximately 61% of the primer extended by 7 nucleotides (Figure 3.5B&D), consistent with previous results (Li, Prywes et al. 2017). In order to test monomer and helper oligonucleotide activation, the formation of imidazolium-bridged species, and primer extension all in a one pot reaction, we subjected 1.2 μM primer and 1.5 μM template, as above, along with 5 mM each NMP and 0.5 mM each trimer to four cycles of ice eutectic phase nucleotide activation. This procedure resulted in 97% of the primer being extended by one or more nucleotides, and 75% by 4 or more nucleotides, but only 22% of +7 extended product (Figure 3.5C&E). However, a subsequent 48-hour incubation at room temperature with two additions of bridge-forming activation chemistry further increased the yield

and extent of primer extension. After this extended incubation, 98% of the primer was extended by one or more nucleotides, 87% by 4 or more nucleotides, and 84% by 7 or more nucleotides, an improvement over the established result with synthetically prepared reactants. Overall, this experiment demonstrates that ice eutectic phase concentration drives both monomer and trimer activation, bridged species formation, and consequently, efficient copying of RNA templates with mixed sequences containing all four canonical nucleotides. Further, the inclusion of bridge-forming activation during a subsequent liquid phase incubation consistently improves the yield of primer extension.

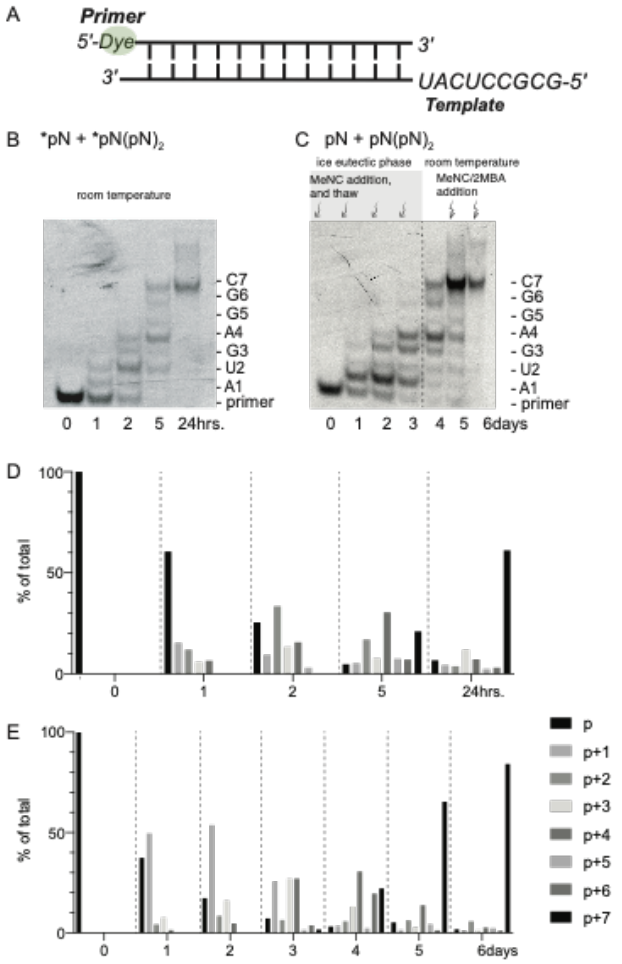


Figure 3.5. Ice eutectic phase activation enables efficient primer extension with all four nucleotides. (A) Schematic of primer-template complex. (B) Time course of primer extension using pure activated mononucleotides and trinucleotides (5 mM each *pN and 0.5 mM each *pN$(pN)_2$, room temperature). (C) Primer-template duplex was directly treated with pNs (N = A, U, C, and G) and $(pN)_3$ and isocyanide activation chemistry under ice eutectic phase. The products of the reaction were then incubated for 24 hours at room temperature, and finally treated with additional (two rounds of) bridge-forming activation for another 48 hours (see details in Materials and Methods 1.6). (D-E) Quantification of primer extension products under conditions (B-C).

Discussion

The plausibility of any scenario for the origin of life on the early Earth is greatly enhanced if different steps or stages in the overall process can occur together under mutually compatible physical and chemical conditions. There are many examples of potentially prebiotic chemical reactions that at first appeared to be incompatible but were subsequently reconciled. For example, it used to be thought that the oxygenous chemistry leading from formaldehyde to sugars such as ribose, and the nitrogenous chemistry leading from cyanide to the nucleobases were strictly incompatible due to the formation of unreactive glycolonitrile and therefore had to occur in separate environments. However, the cyanosulfidic reaction network studied by Sutherland and colleagues has shown that this is not the case, and that it is actually advantageous for these steps to occur together because the facile reduction of glycolonitrile makes it an excellent starting point for many synthetic reactions (Ritson and Sutherland 2013, Patel, Percivalle et al. 2015, Xu, Ritson et al. 2018).

Resolving apparent incompatibilities between distinct aspects of prebiotic chemistry, and especially between sequential chemical reactions, is a significant challenge in origin of life studies, where the goal is to elucidate robust and realistic pathways leading from simple chemical feedstocks

to replicating and evolving protocells. In the present work, we have focused on one such incompatibility, namely the very high concentrations of the activating moiety 2-aminoimidazole (2AI) thought to be required for nucleotide activation, and the fact that high concentrations of 2AI prevent the accumulation of the imidazolium-bridged species required for efficient template-directed RNA copying. We have found that ice eutectic phase incubation enables nucleotide activation with equimolar concentrations of 2AI and nucleotides, thus allowing bridged dinucleotides to accumulate under a prebiotically plausible scenario. Furthermore, cycles of MeNC addition and eutectic freezing drive both efficient activation and bridged dinucleotide formation, directly enabling efficient primer extension. Thus, an environment subject to repeated cycles of freezing and thawing enables the operation of a potentially prebiotic pathway that begins with ribonucleotide 5'-monophosphates, generates 2AI activated nucleotides and then the essential bridged dinucleotides, and finally yields primer extension products, all in one mutually compatible system. Furthermore, in the presence of short helper oligonucleotides, monomer- bridged-oligonucleotide species are formed and enable the copying of mixed sequence templates (Prywes, Blain et al. 2016, Li, Prywes et al. 2017, Ding, Zhou et al. 2022). The greater affinity and higher reactivity of monomer-bridged-oligonucleotide species combine to enable the copying of templates that contain all four canonical nucleotides (Ding, Zhou et al. 2022).

Our observation that the initial activation of mononucleotides can occur under ice eutectic phase conditions suggests a potential role for icy early Earth environments in nonenzymatic RNA copying. The formation and thawing of ice are common phenomenon on Earth's surface, which can take place during day-night temperature cycles, seasonal temperature changes, or random weather-based changes (Gibb, Whittet et al. 2004, Lunine 2006). Ice has previously been implicated in other relevant phenomena, including promoting processes from prebiotic synthesis (Bada, Bigham et al. 1994) to RNA copying, and may even has astrobiological implications (Worth, Sigurdsson et al.

2013, Wordsworth, Knoll et al. 2021). We observed that the efficiency of nucleotide activation is robust with respect to changes in the length of freeze-thaw cycles and the frequency of MeNC addition. MeNC is thought to have been released from prebiotic feedstocks following exposure to UV irradiation from the Sun (Mariani, Russell et al. 2018). The relative independence of activation efficiency from the frequency of MeNC delivery suggests that the required temperature cycling conditions are not stringently constrained, and therefore any corresponding early Earth environment need not be highly contrived. Ice eutectic-based nucleotide activation may not even have been required continuously: once the initial pool of activated mononucleotides had been generated, the application of solution-phase bridge-forming activation could slowly re-activate hydrolyzed nucleotides as well as regenerate bridged dinucleotides in solution. Additional experiments will be needed to determine how long this system can replenish hydrolyzed activated nucleotides without the need for additional freezing after the initial activation. It is possible that although the ice eutectic phase concentration favors activation chemistry, the low temperature may inhibit the actual primer extension reaction. Indeed, primer extension during the freeze-thaw phase of our experiments may be occurring primarily during the relatively short but non-negligible thaw intervals. Nucleotide activation during cycles of partial drying and rehydration may also be relevant (Damer and Deamer 2020).

The key ingredient in the chemical nucleotide activation procedure that we used is MeNC, so it is important to consider whether MeNC could be generated in an environment compatible with our multi-step pathway. For example, the strong UV absorption of nucleotides might leave insufficient UV light for MeNC generation. However, nucleotides absorb UV light most strongly between 240 nm and 280 nm, whereas the generation of MeNC proceeds with longer wavelength irradiation (Cavaluzzi and Borer 2004). Specifically, two steps in the plausible synthesis of MeNC exploit the use of UV light: the production of nitroprusside has been achieved with 365 nm or 254

nm light and isocyanide ligand exchange also occurs during exposure to 365 nm light (Mariani, Russell et al. 2018), consistent with the broad band irradiation from the young Sun (Kasting 2010). The formation of MeNC using longer wavelength UV irradiation suggests that the generation of MeNC is compatible with the presence of nucleotides. Another consideration is that an important ingredient in MeNC synthesis is methylamine, which others have suggested could have been delivered to the primordial Earth by comets, or formed by reduction of HCN with hypophosphite and sponge nickel (Aponte, Elsila et al. 2017, Mariani, Russell et al. 2018). On the basis of Ugi-type reactions (Domling 2002), a primary amine such as methylamine will readily react with an aldehyde, forming an iminium ion, however, this may not affect the nucleotide activation chemistry because MeNC will still add reversibly to the iminium ion to produce the nitrilium ion required for nucleotide activation (Sutherland, Mullen et al. 2008). In addition, 2-hydroxy-N-methylpropanamide, the side product of nucleotide activation, will eventually hydrolyze to regenerate methylamine, which can then take part in another round of MeNC synthesis (Mariani, Russell et al. 2018). The prebiotic plausibility of the remaining chemical ingredients has been previously discussed (Ritson and Sutherland 2012, Ritson and Sutherland 2013, Patel, Percivalle et al. 2015), but additional experiments will be needed to reconstitute in situ MeNC generation with simultaneous nucleotide activation.

Chapter 4

Endocytosis in protocells enables nutrient transport bypassing permeability limits

Reproduced and modified with permission from:

Zhang, S. J., Palapuravan, A., Krishnan, Y., Fai, T. G., Szostak, J. W., and Wang, A. Endocytosis in protocells enables nutrient transport bypassing permeability limits. Manuscript in preparation. S.J.Z., and J.W.S., and A.W. designed research; S.J.Z. performed research and analyzed data; A.P. and Y. K. contributed a novel membrane dye; and S.J.Z., T.G.F., A.W., and J.W.S. wrote the paper.

Abstract

Semipermeable membranes are a key feature of all living organisms. While specialized membrane transporters in cells can import otherwise impermeable nutrients, the earliest cells would have lacked a mechanism to import nutrients rapidly under nutrient-rich circumstances. Using both experiments and simulations, we find that a process akin to endocytosis can be recreated in model primitive cells. Molecules that are too impermeable to be absorbed can be captured in a matter of seconds in an endocytic vesicle. These molecules can then be slowly released over hours, into the main lumen or 'cytoplasm'. This work demonstrates one manner in which primitive life could have broken the symmetry of passive permeation well before the evolution of protein transporters.

Introduction

In extant life, membranes provide a selective barrier between the cell and its environment, which enables the inheritance of adaptive traits and ultimately leads to Darwinian evolution (Sole 2009, Schrum, Zhu et al. 2010). Life itself may have emerged from self-replicating informational molecules spatially localized by primitive membranes (Schrum, Zhu et al. 2010, Dzieciol and Mann 2012). Amongst various lipid candidates, fatty acids are particularly well suited as the components of primitive membranes due to their prebiotic relevance (Yuen, Lawless et al. 1981, Bonfio, Caumes et al. 2019) and dynamic exchange properties. Their probable involvement in the origin of life has been further demonstrated by their ability to grow and divide without complex biochemical machinery (Walde, Goto et al. 1994, Chen and Szostak 2004, Zhu and Szostak 2009) and encapsulate RNA

templates that are being nonenzymatically copied (Chen, Salehi-Ashtiani et al. 2005, Adamala and Szostak 2013).

In nascent life, the absence of protein transporters means that entities enveloped by lipid bilayers would have had to rely on passive diffusion for internalizing nutrients (Barton and Gunstone 1975, Blok, Vanderneutkok et al. 1975, Mansy 2010). Although the diffusive transport across the membrane is symmetric and driven purely by the concentration gradient of solutes across the membrane, the actual fluxes need not be the same, e.g. if the symmetry is broken by placing a nutrient sink within the cell, or by a changing external environment. However, the low solute permeability of primitive membranes, which is desired to prevent the loss of the encapsulated contents, stands in the way of efficiently acquiring nutrients under nutrient-rich circumstances. Thus a protocell would have to reside within a pool of nutrients for hours to absorb useful levels of polar materials (O'Flaherty, Kamat et al. 2018). Whether non-diffusive transport mechanisms can be fueled by physiochemical stimuli remains an important question because such mechanisms could break the trans-membrane symmetry in a way that could bias the inward nutrient flow, or expedite nutrient import.

One type of transport which is neither strictly 'active' or 'passive' in the traditional sense is endocytosis. Such a higher-order topological transformation has been previously demonstrated in phospholipid membranes via various pathways (Lipowsky 1999, Peterlin, Arrigler et al. 2009, Zong, Ma et al. 2017). However, the ability of primitive membranes to endocytose has yet to be definitively shown. Model primitive membranes typically consist of lipids with dynamic properties distinct from those of phospholipids and thus the routes to engender topological changes could potentially differ (Ruiz-Herrero, Fai et al. 2019). For instance, flip-flop of model primitive membrane lipids rapidly relaxes any curvature stress (Bruckner, Mansy et al. 2009).

Here, we demonstrate that primitive cell compartments composed of fatty acids can endocytose via a purely physicochemical process. A simultaneous reduction in volume and increase in surface area allows larger molecules to be imported into model protocells bypassing permeability limits, mimicking the process of endocytosis in complex eukaryotes. The process of 'endocytosis' takes several seconds. The protocells, if removed from the nutrient source, continue to retain the internalized solute, thereby providing time to then absorb the nutrients. Further, we construct a numerical model that captures such out-of-equilibrium topological changes. Taken together, we believe this marks an important step forward toward understanding the development of primitive cells.

Results and Discussions

Endocytosis in fatty acid vesicles

Here we study the shape transformations and subsequent topological transitions of fatty acid vesicles in response to volume and area perturbations. For a sphere to change shape, it must have an increase in surface area to volume ratio. In principle, this can be accomplished via two pathways: by reducing the volume or increasing the surface area. We explore shape changes obtained through combinations of these pathways. We use hyperosmotic shocks to draw water out of the vesicles and reduce the internal volume; the magnitude of the osmotic shock is denoted as a change in concentration of the salt (Na-bicine), ΔC_V, from its starting point of 50 mM. We achieve membrane area growth by adding alkaline oleate micelles to a buffered solution of vesicles; the increase in total membrane lipid is denoted as a change in lipid concentration, ΔC_A, relative to a starting lipid vesicle concentration of 0.5 mM. To observe shape changes in real-time with fluorescence microscopy, we prepared oleic acid/oleate (C18:1) giant unilamellar vesicles (GUVs) encapsulating water-soluble and membrane-impermeable fluorescent dyes (either 0.5 mM HPTS 8-hydroxypyrene-1,3,6-trisulfonic acid trisodium salt, or 1 mM calcein blue). These vesicles were then diluted 1:10 in buffer that was

identical in composition and pH to the original buffer to provide good contrast between encapsulated and unencapsulated dyes, enabling visualization of the dye encapsulated within vesicles.

We previously reported that an increase in lipid concentration by micelle addition to oleate GUVs results in outward budding, with the GUVs transforming into a series of connected pearls (Figure 4.1a) after micelle addition (ΔC_A = 1 mM) (Kindt, Szostak et al. 2020). Addition of a higher concentration of micelles (ΔC_A > 1 mM) resulted in vesicle division. In terms of a putative geological setting for such an event, tectonic and volcanic processes that drive hydrothermal fluid flow might have carried and mixed nutrients, mineral particulates, and lipid components (Phleger, Nichols et al. 1997, Phleger, Nelson et al. 2005, Phleger, Nelson et al. 2005, Govenar 2012). Such episodic events from fluid flow-induced geochemical processes might result in fluctuations in osmotic pressures and simultaneous release of nutrients and lipids. Because both an increase in surface area and reduction in internal volume led to increases in the surface area to volume ratio, we expected that doing both simultaneously (hereby called stimulus A; ΔC_V = 100 mM, ΔC_A = 1 mM) would also result in vesicle division.

Instead, a series of different dramatic shape transformations occurred (Figure 4.1b, Figure S4.1, and Movie S4.1). Specifically, GUVs appeared to elongate, then rapidly transform into a stomatocyte via invagination and then a sphere. We confirmed that the topological transformation was to a vesicle-in-vesicle structure rather than a torus by taking sectional image slices along the z-axis and reconstructing the volume in 3D (Movie S4.2). Further, we performed two distinct experiments to show that the internal vesicles are indeed fully separated from the external membrane. First, we added an aqueous dye, Alexa 594 hydrazide, that is exclusively aqueous and does not bind to membranes, to a final concentration of 10 μM. Alexa 594 hydrazide stayed completely outside of the GUV and did not enter the interior of the endocytosed vesicles, which demonstrates that the inner vesicles are not connected to the outside (Figure S4.2). Second, we

added the dye Voltair^{PM} (Saminathan, Devany et al. 2021)(final concentration: 250 nM) which localizes to the fatty acid membrane by its POPE (1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine) moiety, but cannot flip-flop across the external membrane and reach the inner compartment owing to its conjugation to polar DNA oligonucleotides. The observation that only the GUV outer membranes became labeled excludes the possibility that the inner vesicle membranes are connected to the outer GUV membrane (Figure S4.3). Therefore, instead of budding *outward* (stimulus B; ΔC_V = 0 mM, ΔC_A = 1 mM Figure 1b), the vesicles had budded *inward* followed by pinching off of a complete budded vesicle, presumably via the lateral tension exerted by the bud that is sufficient to break the neck of the complete bud (Seifert, Berndl et al. 1991, Lipowsky 2014).

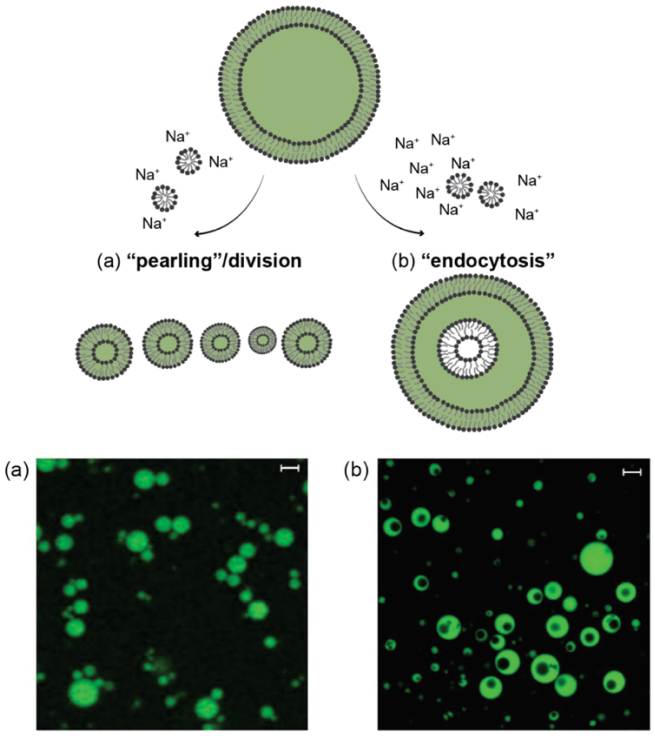


Figure 4.1. Schematic of topological transformations for an oleic acid/oleate giant unilamellar vesicle (GUV) induced by physical stimuli. a) "Pearling" and division of GUV was induced by micelle addition. Confocal micrographs of divided vesicles by 2 equivalents of micelles (stimulus B; ΔC_A = 1 mM). b) The GUV undergoes "endocytosis" forming a compartmentalized vesicle upon

hyperosmotic shock coupled with micelle addition (Movie S1). Confocal micrographs of compartmentalized vesicles formed via GUVs containing 0.05 M Na-bicine buffer solution inside triggered by 100 mM Na-bicine solution and 2 equivalents of micelles outside (stimulus A; ΔC_V = 100 mM, ΔC_A = 1 mM). The vesicles were labeled with 1 mol % HPTS (green fluorescence). Scale bars represent 5 μm.

Effects of osmotic shock and micelles on primitive cellular endocytosis

Micelle addition by itself leads only to outward budding. Interestingly, micelle addition (ΔC_A = 0.5 mM to 5 mM) accompanied by varying levels of osmotic shock (ΔC_V = 25 mM to 100 mM), consistently leads to inward budding. We found that the efficiency of internal compartment formation increased with the magnitude of the increase in surface area. As more micelles were added (ΔC_A = 0.5 mM to 5 mM), the fraction of vesicles with internal compartments increased, and this effect was evident across all magnitudes of osmotic shock (ΔC_V = 25 to 100 mM, Figure 4.2). For example, the population of vesicles with internal compartments increased from 7.3 % (ΔC_A = 0.5 mM; ΔC_V = 100 mM) to 88.5 % (ΔC_A = 5 mM; ΔC_V = 100 mM).

We also found that at any given concentration of added micelles, the number of internal compartments increased with the magnitude of the osmotic shock. At concentrations of added micelles, ΔC_A = 2.5 mM to 5 mM, increasing the osmotic shock increased both the number of the vesicles with compartments and the incidence of multi-compartment vesicles. For example, upon addition of 5 mM micelles, the fraction of vesicles with internal compartments increased from 70% (ΔC_V = 25 mM) to 88.5% (ΔC_V = 100 mM), while the fraction of multicompartment vesicles increased 3-fold from 12.3% (ΔC_V = 25 mM) to 36.2% (ΔC_V = 100 mM). Creating two- or multi-compartment vesicles required a higher concentration of added micelles for lower osmotic shocks. A significant yield of two- or multi-compartment vesicles required ΔC_A = 5 mM at lower osmotic

shock (ΔC_V = 25 mM), whereas it required ΔC_A = 2.5 mM at higher osmotic shock (ΔC_V = 100 mM). At lower concentrations of added micelles (ΔC_A = 0.5 mM to 1 mM), varying the magnitude of osmotic shock (ΔC_V = 25 mM to 100 mM) did not lead to significant changes in the number of internal compartments formed.

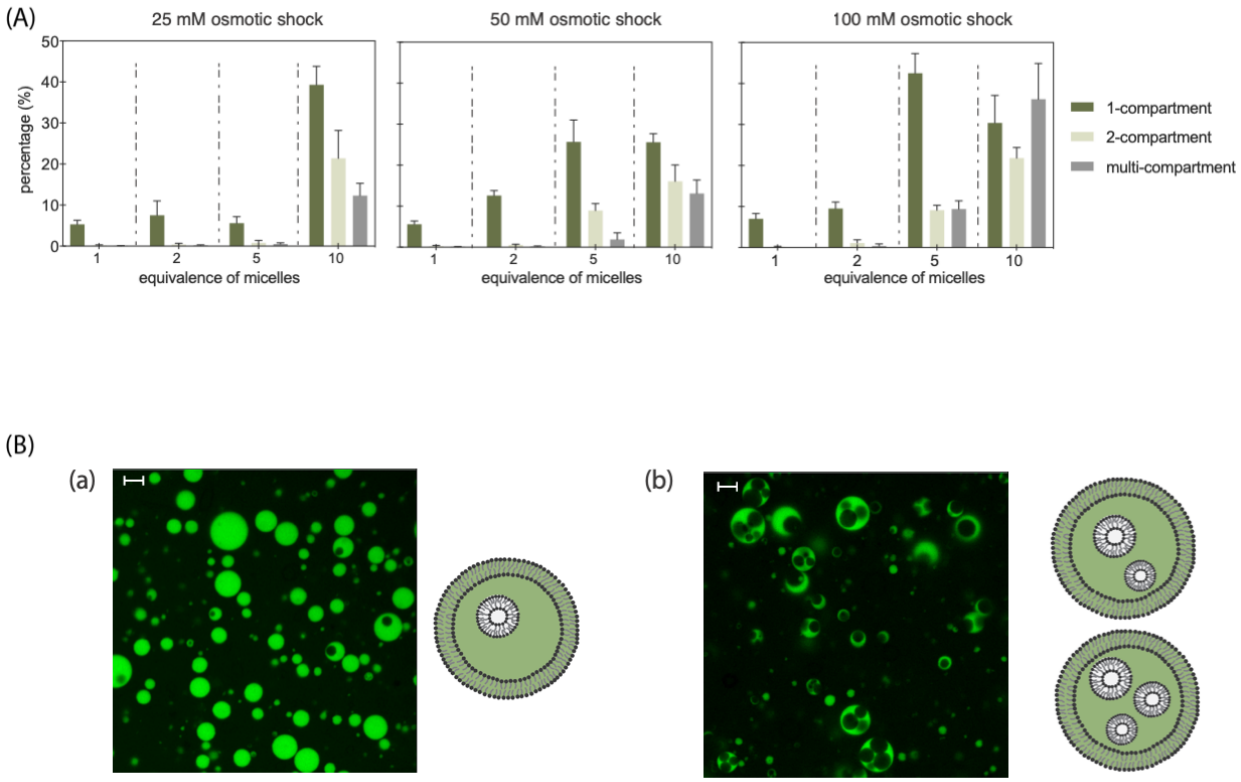


Figure 4.2. (A) Distribution of the number of compartments per vesicle among osmotic shocks and micelles of different magnitude (ΔC_V = 25, 50, 100 mM Na-bicine, ΔC_A = 0.5, 1, 2.5, 5 mM oleate for each cycle). (B) Representative confocal micrographs of vesicles with (a) single compartment and (b) the various number of compartments. Scale bars represent 5 μm.

Having determined that the inward budding is robust, we then asked whether this pathway could allow for the passage of large polar molecules into protocells prior to the evolution of complex protein machinery. We found that the formation of endocytosing vesicles allowed the long

DNA-dye conjugate (5'-Cy5-$C_{(10)}$ $A_{(18)}$-3') (Figure 4.3A) to be imported from the external medium into the vesicle itself. Specifically, Cy5 (magenta) labeled DNA 28-mers were introduced along with the simultaneous osmotic shock and micelle addition (stimulus A) into oleic acid/oleate GUVs encapsulating calcein blue. The resultant formation of magenta interior compartments indicates that the Cy5 labeled DNA 28-mers were successfully transported to the lumen of the internalized vesicles and internalized by the GUVs (Figure 4.3B(a,b)).

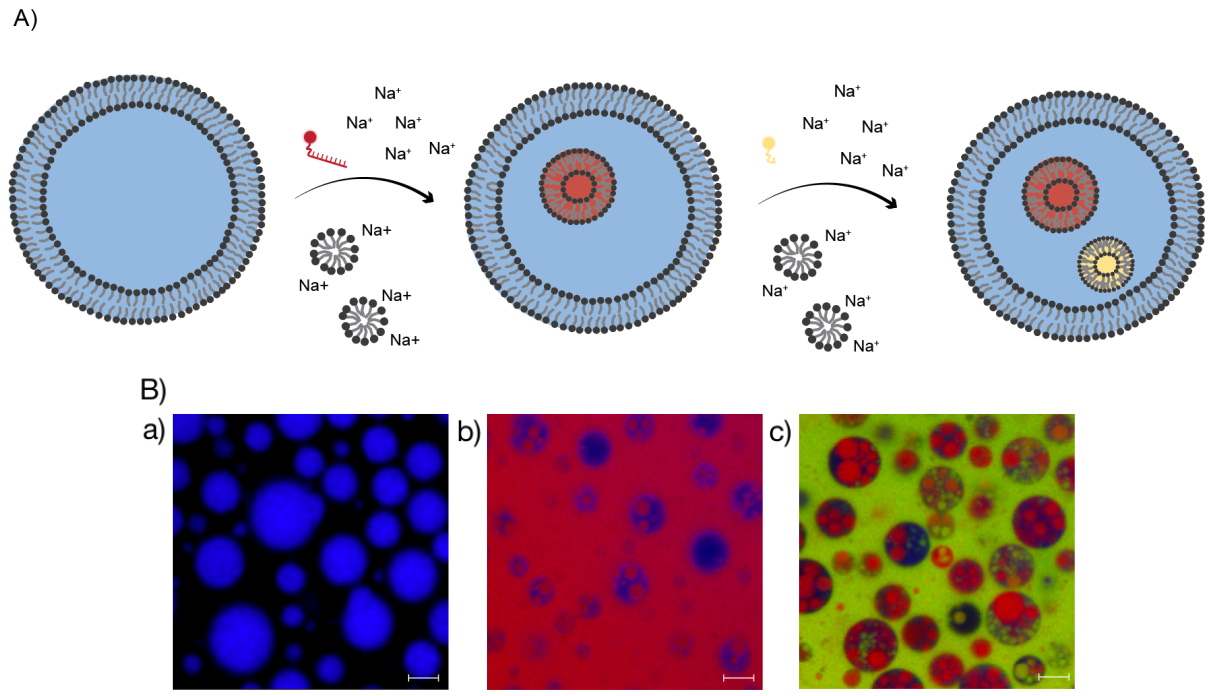


Figure 4.3. (A) Scheme of multiple endocytosis events induced by successive osmotic shock coupled with micelle addition. (B) Confocal micrographs of fatty acid vesicles with various numbers and sizes of interior compartments. (a) Vesicles encapsulate dye (calcein blue, blue). The inner compartments encapsulate dyed oligomers (5'-cy5-$C_{(10)}$ $A_{(18)}$-3', red), and dyed monomers (fluorescein-12-UTP, yellow) followed by one (b) or two (c) hyperosmotic shocks coupled with micelle addition with some mixing between compartments (ΔC_V = 100 mM Na-bicine, ΔC_A = 2.5 mM oleate for each cycle). Scale bars represent 5 μm.

The resultant vesicle-in-vesicle structures were not expected to be limited to a single inward budding event. We therefore asked whether repeated inward budding could be induced by successive rounds of simultaneous osmotic shock and micelle addition (stimulus A). Following the intake of Cy5-labeled DNA 28-mers, a second simultaneous osmotic shock and micelle addition event was induced in the presence of the yellow aqueous compound fluorescein-12-UTP in the exterior solution. This resulted in the blue vesicles containing both magenta and yellow inner compartments (Figure 4.3B c). With both stimulus events being in the regime where two or more internal compartments are likely (Fig. 4.2), the resultant vesicles displayed a high internal volume fraction of internalized vesicles. The excess micelles introduced also led to *de novo* vesicle formation: magenta vesicles containing green inner compartments are also occasionally observed.

This result is informative for two reasons. First, the presence of inner compartments that are a different color from the external medium (after two cycles of micelle addition in the presence of different dyes) indicates that the shape transformation of the vesicles does not end in the stage of stomatocytes (Figure S4.1a). Rather, inward vesiculation was completed and a full topological transformation had occurred (Figure S4.1b). Without the final topological transformation, all the compartments inside would be the same color as the external medium. Second, the presence of both yellow and magenta internal compartments is consistent with two separate rounds of endocytosis taking place. This confirms that the simultaneous osmotic shock and micelle addition event is able to trigger endocytosis.

Release of nutrients enclosed in endocytic compartment to main lumen

A substantial uptake of nutrients by passive diffusion requires protocells to reside within a pool of nutrients for hours. Endocytosis, however, enables the inward flow and retention of the nutrients upon transient interactions with a hyperosmotic environment rich in nutrients such as nucleotides and containing a high concentration of fatty acid micelles. Releasing nutrients stored

inside the endocytic compartments into the main lumen (or 'cytoplasm') would allow them to interact with other components within the protocell and is the final critical step in this pathway for nutrient uptake.

To test the ability of the internal endocytic compartment to release nutrients, we first used a relatively impermeable substrate, the fluorescein-labeled cyclic dinucleotide (cGAMPfluo) as a fluorescent tracer. cGAMPfluo was introduced along with a simultaneous osmotic shock and micelle addition (ΔC_V = 100 mM, ΔC_A = 1 mM) to oleic acid/oleate GUVs (Figure 4.4A). The endocytosed vesicles were then diluted five-fold into a 200 mM glucose buffer that was identical in composition and pH to the original buffer but lacked the lipids and encapsulated solute to remove the free cGAMPfluo from the external medium. We then took confocal microscopy images over time to measure the fluorescence intensities in the internalized compartments, the main lumen of the GUVs, and the external medium. The fluorescence intensity serves as a measure of relative concentrations of encapsulated material.

Tracking the fluorescence intensity of the internalized compartments relative to the external medium (Figure 4.4A), which we consider an infinite bath, showed a decrease from 6.2 $\pm$ 1.0 to 3.9 $\pm$ .9 (N=124) over 24 hours. This suggests that the oligomer was slowly released by the internalized compartment, into the main lumen. Indeed, the fluorescence intensity of the internalized compartments relative to the main lumen decreased more than two-fold, from 36.3 $\pm$ 1.7 to 14.3 $\pm$ 1.1 over 24 hours (Figure 4.4B, N=113). This indicates that the oligomer was slowly released from the internalized compartment into the main lumen. We found similar results for another representative molecule with low permeability to fatty acid membranes (Fujikawa, Chen et al. 2005), HPTS (Figure S4.5). Overall, these results demonstrate that internalized compartments can be used as nutrient storage depots, which would slowly release contents to the main lumen or 'cytoplasm' of a protocell.

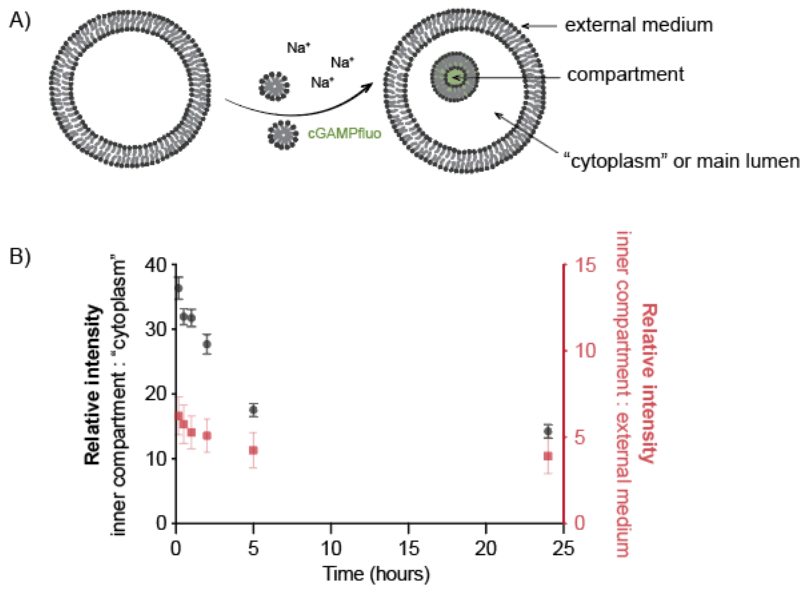


Figure 4.4. Internal compartments can slowly deliver nutrients of low permeability to the main lumen or 'cytoplasm' of the protocell. (A) Fluorescein labeled cyclic dinucleotide (cGAMPfluo), a relatively impermeable substrate, is encapsulated in the interior compartment via one hyperosmotic shock coupled with micelle addition (ΔC_V = 100 mM Na-bicine, ΔC_A = 2.5 mM oleate, $C_{cGAMOfluo}$ = 2 mol% of the total lipid). (B) The fluorescence intensity ratio of 'cytoplasm' over compartment increases and of external medium over compartment increases indicate that dinucleotide inside the compartment is slowly entering the 'cytoplasm'.

Modeling out-of-equilibrium topological changes

To understand the reason for inward vesiculation, we model the vesicle as an inextensible elastic material with prescribed bending modulus, spontaneous curvature, and permeability. The surrounding fluid is considered as an aqueous solution that contains amphiphiles that incorporate into the vesicle membrane at a given rate. As shown in Ruiz-Herrero et al. (Ruiz-Herrero, Fai et al. 2019), the resulting behavior may be described in terms of two dimensionless parameters that

govern the steady-state vesicle morphology. In addition, to capture the effect of osmotic shocks, here we include the possibility of osmotic pressures that drive flows across the membrane. This is done by adding an appropriate normal force to the membrane surface corresponding to the van't Hoff pressure $p = ck_BT$, where c is the difference in osmotic concentrations across the membrane.

After applying osmotic shocks in the presence of permeability and membrane growth, we observe that vesicles develop stomatocyte morphologies that progress spontaneously to develop inward buds (Figure 4.5A-B). The initial hyper-osmotic shock causes the vesicles to shrink in volume while adding additional surface area. This increases the reduced volume fraction over time. We use parameters listed in Table S1 and confirm that the timescales of volume loss and area increase are consistent with those of prior work (Sacerdote and Szostak 2005).

In terms of the morphological phase plane described in Ruiz-Herrero et al. (Ruiz-Herrero, Fai et al. 2019), the experimental parameter regime corresponds to inward vesiculation. We find that the added effect of osmotic pressures is parallel to a higher effective membrane growth rate and lowers the threshold for vesiculation (Figure 4.5C). Our model recapitulates several of the behaviors observed experimentally, including the spontaneous onset of inner vesiculation on the correct timescale of seconds. In terms of the resulting vesicle morphology, we show that the resulting effect from hyperosmotic shock corresponds to that of an increased effective membrane growth rate. This rationale is consistent with the lowered vesiculation threshold that is observed in experiments on vesicles under osmotic shocks.

Several open questions remain. First, positive spontaneous curvature favors outward budding, whereas negative spontaneous curvature favors inward budding. This suggests that negative spontaneous curvature is important for endocytosis to occur, yet the mechanism of how negative spontaneous curvature would arise remains unclear. One possibility is that the bilayer leaflets have an asymmetric interaction with water molecules or ions during water efflux and

membrane growth. Another hypothesis is that increased external salt bridges carboxylates, either decreasing effective membrane area or perhaps changing effective lipid shape so as to favor negative curvature. Ultimately, while coarse-grained models are effective for predicting shape changes, molecular dynamics simulations of membrane packing might be necessary to reveal the underlying mechanism and cause (Markvoort, Spijker et al. 2009). Second, in these simulations, only one inner vesicle is seen to form, whereas the experiments show that different numbers of inner vesicles may emerge (Figure 2B(b) and Figure 3B(c)). We speculate that it may be possible to increase the number of inner vesicles by lowering the spontaneous curvature and bending modulus. Although here we explicitly include the solution by solving for the hydrodynamics of the incompressible surrounding fluid, we have not studied the role of hydrodynamics in controlling the vesiculation behavior in detail. This would be an interesting direction for future study.

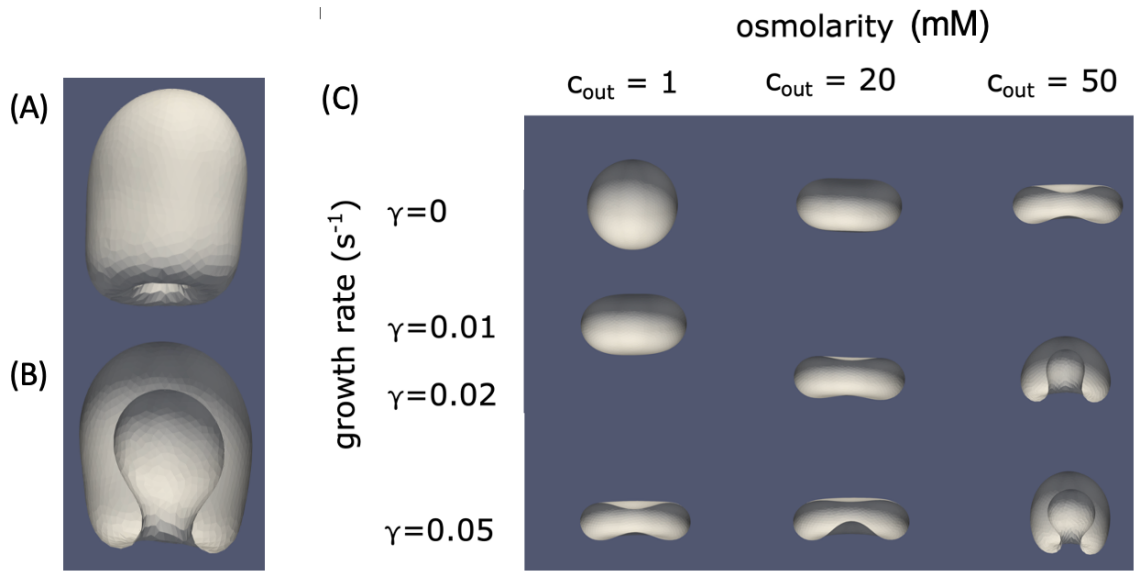


Figure 4.5. Modeling of vesicle topological changes. (A) Side view and (B) cutaway of the final morphology of an initially spherical vesicle after 2 s of growth, with spontaneous formation of an inner vesicle. (C) Final morphology of an initially spherical vesicle after 3 s of growth, for different values of the growth rate (g) and osmolarity ($\Delta c = c_{out} - c_{in}$). Increasing the osmotic shock results in an effective higher growth rate, thereby reducing the threshold for vesiculation.

Comparison to passive permeability

In light of our experimental demonstration of the release of both dinucleotides and HPTS from the endocytic compartment into the lumen, we also seek to further understand the utility of endocytosis as an inward-transport mechanism and estimate the time that is required for the vesicle's lumen to obtain various nutrients by passive diffusion across the compartment membrane. Such time scales can be estimated from the net flux (denoted by J, with units of number of molecules crossing unit area per unit time). The spontaneous transport of molecules along their concentration gradient across membranes is described by a simple equation

$$J = P \cdot \Delta C \ , \tag{1}$$

where P is the permeability coefficient and $\Delta C\ (C_{in} - C_{out})$ is the concentration difference across the membrane. The number of molecules that cross a given area per unit time can be determined by rearranging Eqn. 1 and

$$J = \frac{dN}{dt} \cdot \frac{1}{A} \tag{2}$$

into $dN = J \cdot dt \cdot A = P \cdot \Delta C \cdot dt \cdot A$ (3). Membrane permeability coefficients (P) through oleic acid membranes for short oligomers up to the size of trinucleotides, have been previously measured.[14] We consider trinucleotides, which can enhance nonenzymatic primer extension. The permeability of the oleic acid membrane to trimer is $P \approx 0.21 \cdot 10^{-12}$ cm/s. In the presence of a 1 mM trimer 'pool', the number of trimers that enter a GUV with radius of 2 μm from the external medium can be estimated from Equation (3) giving 0.06 trimers per second. In other words, it takes approximately 16 seconds for a single trimer to cross the oleic acid membranes when vesicles are surrounded by 1 mM trimers.

We assessed the flux for a range of permeabilities (Figure S4.6) spanning values for common nutrients (10^{-10} and 10^{-13} cm/s). We found that it takes anywhere between 0.03-33

seconds for a single molecule to cross 4-μm-diameter oleic acid vesicles. It is clear that a protocell would have to reside within such pool of nutrients for minutes to hours to acquire a substantial amount of the nutrients by passive diffusion.

This result is in drastic contrast to endocytosis, which transports a parcel of nutrients inwards within seconds. We find that a dramatic release event is not necessary for the nutrients to reach the 'cytoplasm'. Instead, the molecules in the inner compartment can be slowly released into the main lumen of the protocell, where they can then interact with other components. For an inner compartment ¼ the diameter of a 4-μm-diameter GUV, 90% of molecules with permeability comparable to trimers can be released into the lumen within 11 hours (Figure S4.5A). This indicates that protocells that endocytose can absorb a substantial amount of nutrients from a pool despite interacting with it only briefly.

This principle can be extended further to consider the transport of even larger oligomers. While permeabilities of longer oligomers through oleic acid membranes remain unknown, they are expected to have a lower permeability than trimers. After approximately 4.7 hours (Figure S4.6B), 10% of the molecules with a permeability of 10^{-13} cm/s are released from the interior compartment, with full release within 5 days. This result points to a remarkable function of endocytosis – a means of importing otherwise 'membrane-impermeant' molecules.

Conclusions

We have shown that model primitive cell membranes are capable of inward vesiculation, leading to a complete topological transition to a vesicle-in-vesicle morphology. The number of internal compartments can be controlled by the rate of surface area growth, with the osmotic shocks decreasing the rate of surface area growth necessary for vesiculation. We also recapitulated the main results and the relevant timescales in an out-of-equilibrium numerical model. We then found that

such inward vesiculation events could lead to internalization of nutrient solutes including mononucleotides and oligonucleotides, drawing further parallels to endocytosis. Such processes could have helped primitive cells capture nutrients that are otherwise impermeable.

Chapter 5

Conclusions and outlook

My thesis work identifies chemical and physical mechanisms that enables the prebiotically plausible pathway begins with ribonucleotide 5'-monophosphates, generates 2AI activated nucleotides and then the essential bridged dinucleotides, and finally yields templated polymerization products under mutually compatible conditions. As discussed previously, a route to selective phosphate activation with isocyanide, aldehyde, and imidazole, has been reported (Mariani, Russell et al. 2018). A major hurdle to the compatibility of this activation chemistry and RNA copying is that activation requires an excess of 2AI to drive the reaction forward; however, excess 2AI prevents nonenzymatic RNA polymerization by inhibiting the accumulation of bridged dinucleotides (Walton and Szostak 2016, Walton and Szostak 2017). We have identified a variant of isocyanide activation chemistry, named bridge-forming activation, that circumvents this issue by directly generating bridged dinucleotides from activated mononucleotides (Zhang, Duzdevich et al. 2020). The consequent high concentrations of bridged dinucleotide increase the yield of primer extension products.

Although bridge-forming activation enabled us to demonstrate the compatibility of a prebiotically plausible activation chemistry with nonenzymatic RNA polymerization, it nonetheless depends on a source of previously activated mononucleotides. Efficient *in-situ* activation of mononucleotides with 2AI was observed after repeatedly subjecting the solution to ice eutectic phase freezing, which temporarily brings the reactants into close physical proximity and obviates the need for a large excess of 2AI otherwise required to drive activation in relatively dilute aqueous phase reactions. The geochemically plausible conditions that we have identified enable a potentially prebiotic pathway from chemical nucleotide activation to nonenzymatic RNA copying in one reaction mixture.

My thesis work has also demonstrated how primitive cells composed of fatty acids endocytose via a purely physicochemical process upon encountering nutrients in conjunction with

the extra membrane material. Such inward vesiculation events could lead to the internalization of nutrient solutes including mononucleotides and oligonucleotides, drawing further parallels to endocytosis, and helping primitive cells capture nutrients that are otherwise impermeable. Our results also highlight new opportunities for the study of non-equilibrium phenomena in primitive cells, and provide potential environmental constraints for inducing complex behavior in the self-assembly of fatty acid vesicles.

Our elucidation of bridge-forming activation opens up the possibility of continuous re-activation of hydrolyzed nucleotides for further rounds of polymerization. The periodic introduction of methyl isocyanide is expected to continuously regenerate bridged intermediates and hence increase the yield of full-length copies of longer sequences. Future research could investigate how continuous re-activation can be achieved in the context of primer extension. One possible challenge is the accumulation of side products, which may inhibit primer extension by competing for binding sites on the template. A major question that has not been pursued yet is identifying alternative environmental conditions relevant to methyl isocyanide activation on the early Earth. To evaluate the likelihood of various sources, processes, and geochemical environments on early Earth, we need to consider various potential scenarios. (Note that even if there are multiple pathways to life, there are some basic constraints. For example, the extreme high or low temperatures that will denature the RNA are unlikely to occur.) As the methyl isocyanide is released upon irradiation, the possibility of spatially and temporally controlled activation chemistry is thus raised. The frequency of methyl isocyanide delivery in the system and how that will affect the activation yield require future work. The eutectic phase might have facilitated certain nucleotide activation, which indicates the presence of ice matrices on the early Earth. However, the hydrothermal field is also considered a candidate prebiotic Earth environment, where repeated dehydration-hydration cycles might come into play. The activation under partial dry and hydration cycles could also be studied. Additional processes that

might sequester or degrade 2AI after efficient nucleotide activation should also be investigated. For example, UV radiation, the presumptive energy source for producing isocyanide, photodegrades 2AI on the order of days (Todd, Szabla et al. 2019) although the photodegradation rates of activated nucleotides and bridged dinucleotides remain unknown.

To provide a more complete description of how primitive cells (protocells) could have emerged, compartmentalization will be required to recreate a prebiotically plausible protocell and understand how membranes and genetic polymers could have coexisted and coevolved. The effects of a reducing environment, freeze/thaw cycles, and wet/dry cycles on primitive membranes are not well understood. Additionally, the lamellarity, size, density of the vesicles, and presence of other ubiquitous solutes that might protect membranes against damage arising from freezing or dehydration will need to be assessed. Exploration of alternative environmental conditions and possible constraints might render possible early Earth conditions relevant to isocyanide activation chemistry.

Appendix

Chapter 2

Potentially prebiotic activation chemistry compatible with nonenzymatic RNA copying

Materials and methods
Supplementary figures S2.1 to S2.13
Supplementary schemes S2.1
Supplementary tables S2.1 to S2.4
Supplementary texts

Materials and methods

General information

Materials. Reagents and solvents were obtained with highest purity available from Acros Organics, Alfa Aesar, Fisher Scientific, Sigma-Aldrich, ThermoFisher Scientific, or Tokyo Chemical Industry Co., and were used without any further purification unless noted below. Nucleoside-5'-monophosphate, free acid, was purchased from Santa Cruz Biotechnology. 2-aminoimidazole hydrochloride and 2,2'-dipyridyldisulfide were purchased from Combi Blocks. RNA oligonucleotides were purchased from Integrated DNA Technologies. All reactions were carried out in DNase/RNase-free distilled water.

NMR spectroscopy. ^{1}H and ^{31}P-NMR spectra were obtained using a 400 MHz NMR spectrometer (Varian INOVA) operating at 400 MHz and 161 MHz respectively. Samples in H_2O/D_2O mixtures were analyzed using Wet1D suppression to collect ^{1}H-NMR data. Chemical shifts (δ) are shown in ppm. Coupling constants (J) are given in Hertz (Hz) and the notations s, d, t, and m represent the multiplicities of singlet, doublet, triplet, and multiplet, respectively.

pH measurements. pH values were determined by a micro pH probe (Orion 9863BN) equipped with a needle tip and a SevenCompact meter (Mettler Toledo S220).

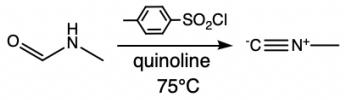
Mass spectrometry. All samples of nucleotides from isocyanide activation reactions were purified by acetone precipitation as described previously, prior to analysis by mass spectrometry (Walton and Szostak 2016). The nucleotides were precipitated by combining the activation reaction mixture with

a solution of 3 mL diethyl ether, 6 mL of acetone, and 0.23 g of sodium perchlorate. The precipitated material was washed with acetone and diluted to 200 μM in Milli-Q water with a few drops of acetonitrile immediately prior to analysis. Spectra were obtained by direct injection on an Esquire 6000 mass spectrometer (Bruker Daltonics), operated in the alternating ion mode.

Data analysis. All spectra were analyzed using MestReNova (version 12.0.3). The yields of conversion were determined by the relative integration of the signals in the ^{1}H or ^{31}P NMR spectra.

All data shown are representative of distinct samples, n = 3 replicates or greater.

Synthesis and storage of methyl isocyanide (R. E. Schuster 1966)



In a 500 mL 3-neck flask equipped with a pressure-equalizing dropping funnel, a sealed mechanical stirrer, a thermometer, and a receiver trap were placed 100 mL (800 mmoles) of quinoline, which was freshly distilled from zinc dust (Alfa Aesar), and 57.2 g (300 mmoles) of *p*-toluenesulfonyl chloride. The solution was heated to 75 °C in an oil bath and the system evacuated to a pressure of 15 mm. The receiver was cooled in a bath of liquid nitrogen. While the solution was vigorously stirred and maintained at this temperature, 11.8 g (200 mmoles) of N-methylformamide was added dropwise to maintain a smooth distillation rate. The addition was complete in 45-60 minutes. The material, which collected in the receiver, was distilled under vacuum. Methyl isocyanide was collected at 59-60°C.

Analysis by NMR indicates that the purity exceeded 98% (Figure S1); ^{1}H-NMR (400MHz, 10% D_2O in H_2O) δ 3.03 (t, ^{1}H, J =2.28 Hz). Methyl isocyanide was stored at pH 9 or greater directly after distillation in Teflon-backed screw-cap glass vials inside air-tight plastic containers.

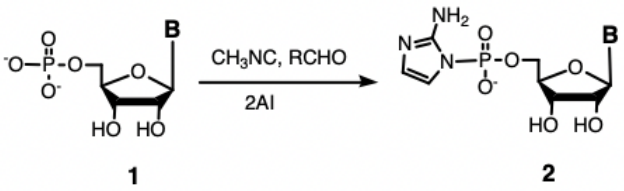
Preparation, storage, and concentration determination of stock solutions

Stock solutions of nucleoside-5'-monophosphate disodium salt, 2-aminoimidazole hydrochloride, and buffer, 2-(Bis (2-hydroxyethyl) amino) acetic acid (BICINE), were prepared by dissolving the corresponding reagent in DNase and RNase-free distilled water. After adjusting the pH to the reported values with NaOH/HCl, the stock solution was filter sterilized with 0.22-micron syringe filters (Millipore Sigma). Each stock solution was then aliquoted and kept at -20 °C until further use.

The exact concentrations of the nucleoside-5'-monophosphates were determined by analysis of serial dilutions on a spectrophotometer. The absolute concentrations of the other stock solutions were found by comparing the integrals of ^{1}H-NMR peaks of interest to the calibrant, adenosine-5'-monophosphate by NMR spectroscopy.

Nucleoside 5′-phosphoro-2-aminoimidazolides (2AImpN)

a. Generation by isocyanide nucleotide activation chemistry



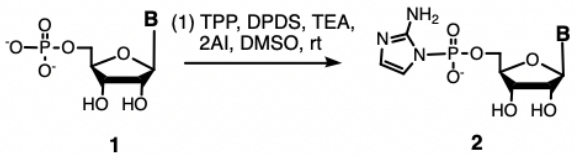
1 2

The nucleoside-5'-monophosphate (NMP), 2-aminoimidazole (2AI), the indicated aldehyde, magnesium chloride ($MgCl_2$), water, and/or buffer solution were added to their corresponding concentration listed in Table S1, in a total volume of 450 μL. Buffer was added to avoid changing pH of the solution during the reaction. Methyl isocyanide was added to the solution, which was briefly vortexed ensuring proper mixing. The total volume was brought to 450 μL. The reaction was allowed to sit for 6 hours, which was the optimal incubation time as determined by the time course (Figure 2), at room temperature. 50 μL D_2O (10%) was added before transferring 500 μL of this mixture to an NMR tube (Fisher) for NMR spectroscopy. The chemical shifts of 2AImpA were reported as they appeared in the ^{31}P NMR of the reaction mixture. Spectroscopic data (Figures 3a &

b) were in agreement with previous literature (Li, Prywes et al. 2017, Mariani, Russell et al. 2018); for 2AImpA, ^{31}P-NMR (161 MHz, 10% D_2O in H_2O, 1H-decoupled): δ -8.70.

The progress of the reaction was monitored by ^{31}P-NMR over the course of 15 hours in a Shigemi NMR microtube assembly (Sigma Aldrich) with trimethyl phosphate as the internal reference (δ 0.00 ppm). The entire time-course of the reaction was obtained through arrayed acquisition and tracked by monitoring peak areas.

b. Synthesis of synthetic standards for activated nucleotides



1 2

Nucleoside-5'-monophosphate free acid (1 equiv.) was dissolved along with 2-aminoimidazole hydrochloride (10 equiv.) in water and the pH was adjusted to 5.5 using NaOH/HCl. The solution was flash frozen in liquid nitrogen and lyophilized to yield a powder. To the solution of nucleoside-5'-mono-phosphate and 2-aminoimidazole in triethylamine (13 equiv., TEA) and DMSO (30 mL) was added 2,2'-dipyridyldisulfide (10 equiv., DPDS), triphenylphosphine (9 equiv., TPP) with stirring and under argon for 30 minutes. The reaction mixture was then added to a pre-chilled solution of acetone/diethyl ether/TEA (400/250/30 mL) to which 1.5mL of saturated $NaClO_4$ in acetone was added. After the precipitate had settled out, the majority of the supernatant was removed using pipette-suction and the remaining suspension was centrifuged at 4,000 rpm for 5 min. The pellets were washed first with a solution of acetone/diethyl ether/TEA (133/83/10 mL) and twice with acetone (10 mL). The product was purified by reverse-phase flash chromatography using gradient elution between (A) aqueous Milli-Q water and (B) acetonitrile. The sample was eluted between 0% and 15% B over 8 column volumes (CVs) with a flow rate of 40 mL/min.

Synthesis of adenosine 5'-phosphoro-2-methylimidazolides (2MImpA)

2MImpA was prepared as above but with the following minor difference: 2-methylimidazole was used instead of 2-aminoimidazole.

Primer extension reactions and analysis of polymerization products

The primer-template duplex was first annealed in a 40 μL solution containing 2 μM thiol-modified primer, 3.1 μM template, 416.7 mM HEPES (pH 8.0), 62.5 mM $MgCl_2$ by heating at 95 °C for 3 min. and cooling down to 23 °C at a rate of 0.1 °C/s. The reaction was initiated by the addition of activated monomers. The stock solutions (100 mM) of 2AI-activated monomers had a pH of around 9.6.

Aliquots (10 μL each) were removed at given time points and desalted using a ZYMO Oligo Clean & Concentrator spin column (ZYMO Research). The isolated material was resuspended in 30 uL 100 mM HEPES (pH 7.50) and disulfide bonds were reduced using a 10-fold molar excess of tris-(2-carboxyethyl) phosphine hydrochloride (TCEP). Alexa 488 C5 maleimide dissolved in anhydrous dimethyl sulfoxide (DMSO) at a concentration of 1 mM was added to the primer-template duplex dropwise, and the reaction was allowed to proceed at room temperature for 2 hours, protected from light. The labeled primer-template duplex was separated from free dye using ZYMO DNA Clean & Concentrator-5 spin columns, resuspended in 5 uL of 100 mM HEPES buffer, and mixed with 30 μL of quenching buffer containing 8.3 M urea, 1.3x Tris/Borate/EDTA (TBE) buffer (pH 8.0), 0.005% Bromphenol Blue, 0.04% Orange G, and 75 μM RNA complementary to the template. To denature the labeled product from the template prior to polyacrylamide gel analysis, and sequester the template with its complement (added in excess as a competitor), the sample was heated to 95 °C for 3 min. and cooled to 23 °C at 0.1 °C/s. The primers, templates, complementary DNAs and RNAs used in the primer extension assays are listed below (Table S4).

20% polyacrylamide gels were prepared using the SequaGel–UreaGel system (National Diagnostics, Atlanta, GA). The gels were allowed to pre-run at a constant power of 20 W for 40−45 mins. 15 μL aliquots of samples were separated by 20% (19:1) denaturing PAGE at a constant power of 5 W for 10 min. and 25 W for 1.5 hrs. Gels were imaged on a Typhoon 9410 scanner (GE Healthcare, Little Chalfont, Buckinghamshire, UK) and quantified using the accompanying ImageQuant TL software. All reactions were performed in triplicate or greater.

Supplementary Figures

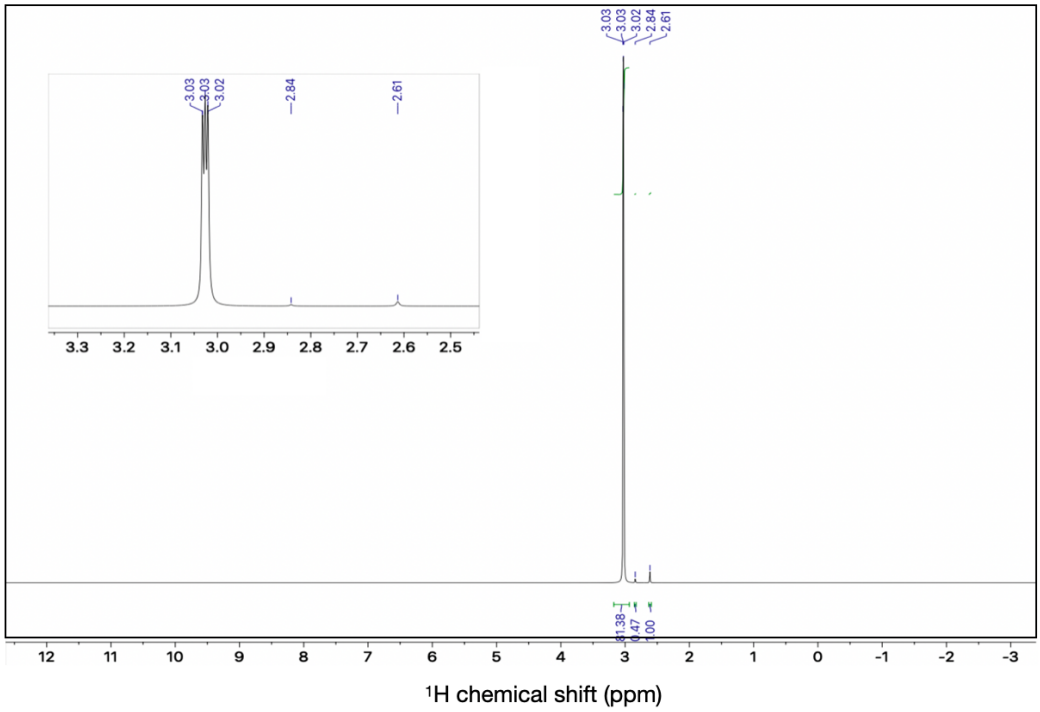


Figure S2.1. ^{1}H-NMR spectrum of synthesized methyl isocyanide in 10% D_2O. Inset is the zoomed-in view of the three peaks.

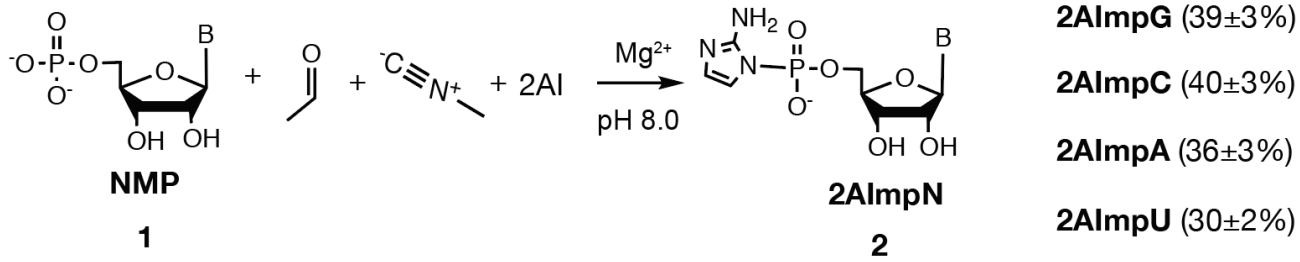


Figure S2.2a. Isocyanide-mediated phosphate activation with 2AI of the four canonical ribonucleotides. The relative yield of each 2AImpN is calculated based on the corresponding ^{31}P NMR peak integration normalized to the number of phosphorus atoms in a given species. Errors are standard deviations of the mean, n = 3 replicates.
Reaction conditions: AMP (10 mM), Mg^{2+} (30 mM, $MgCl_2$), methyl isocyanide (200 mM), acetaldehyde (200 mM), 2AI (100 mM), and HEPES (200 mM) at pH 8 and t = 6 hours (the timepoint at which the concentration of activated monomer plateaus) (Figure S2.2b).

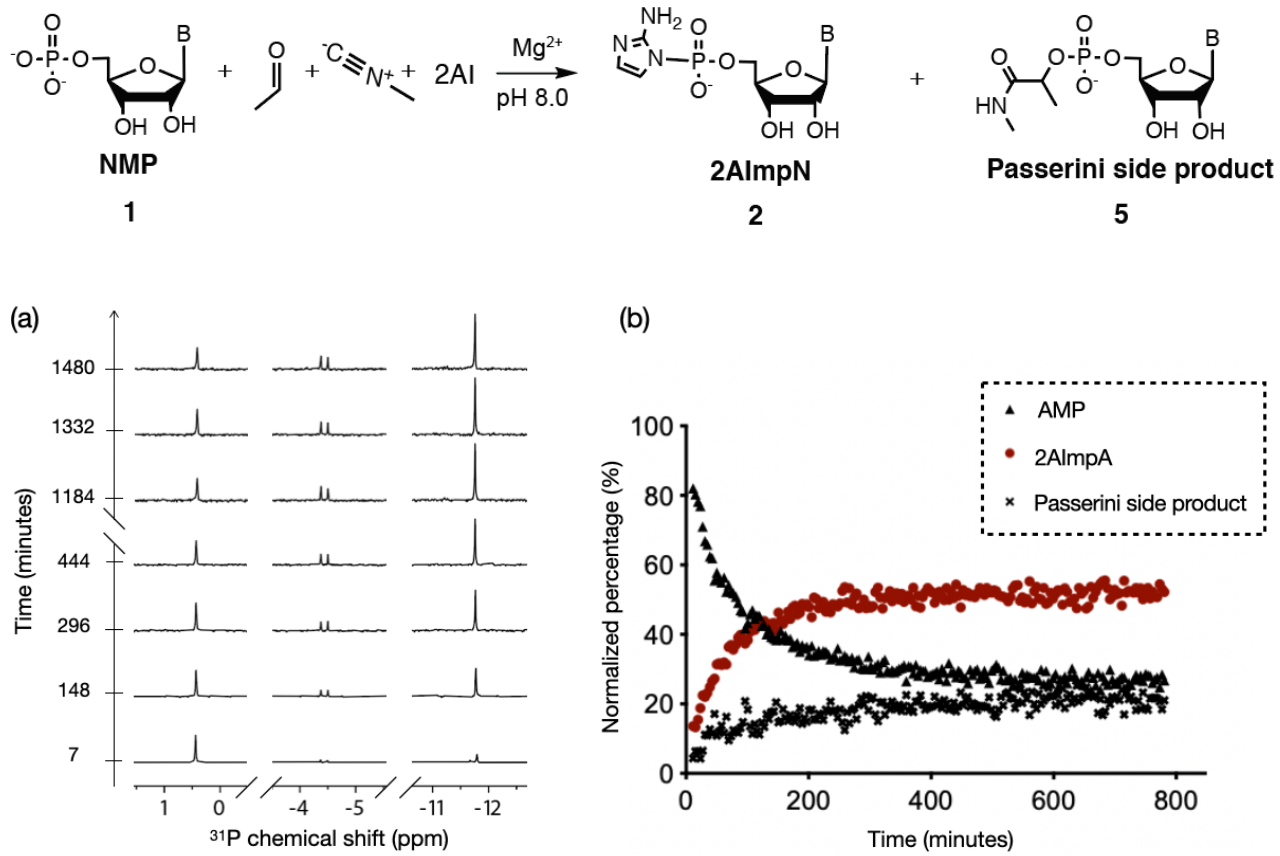
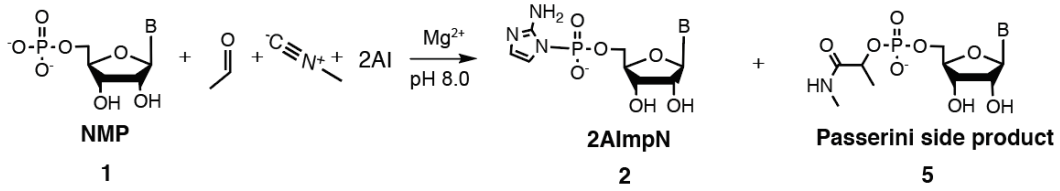


Figure S2.2b. ^{31}P-NMR time-course experiments for 2AImpA synthesis by isocyanide chemistry. (a) The stacked ^{31}P-NMR spectra of nucleotide activation. All spectra are referenced to the trimethyl phosphate peak at 0 ppm (not shown). (b) Quantification of nucleotide activation as a function of time based on (a).
Reaction conditions: AMP (10 mM), Mg^{2+} (10 mM, $MgCl_2$), HEPES (200 mM) at pH 8, 2AI (200 mM), acetaldehyde (200 mM), and methyl isocyanide (200 mM).



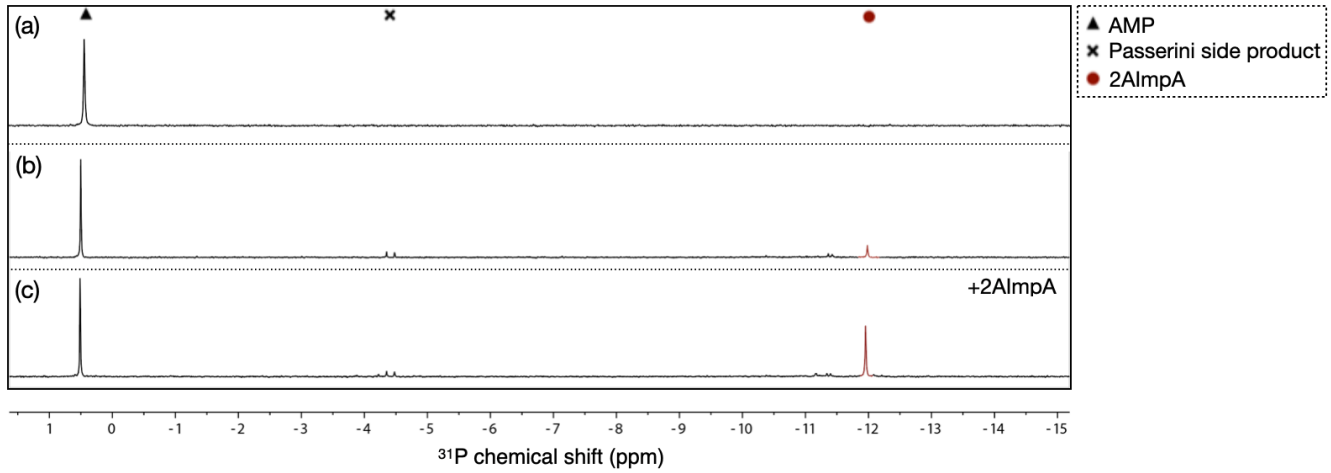


Figure S2.3. Formation of 2AImpA by isocyanide activation chemistry. ^{31}P NMR spectra of (a) AMP alone. (b) 2AImpA formed by isocyanide activation chemistry (c) as in (b) but with spiking of synthetic standard 2AImpA.
Initial reaction conditions:
(a) AMP (10 mM), Mg^{2+} (30 mM, $MgCl_2$), and HEPES (200 mM) at pH=8.0.
(b) AMP (10 mM), acetaldehyde (150 mM), methyl isocyanide (200 mM), Mg^{2+} (30 mM, $MgCl_2$), HEPES (200 mM) at pH 8, 2AI (100 mM).

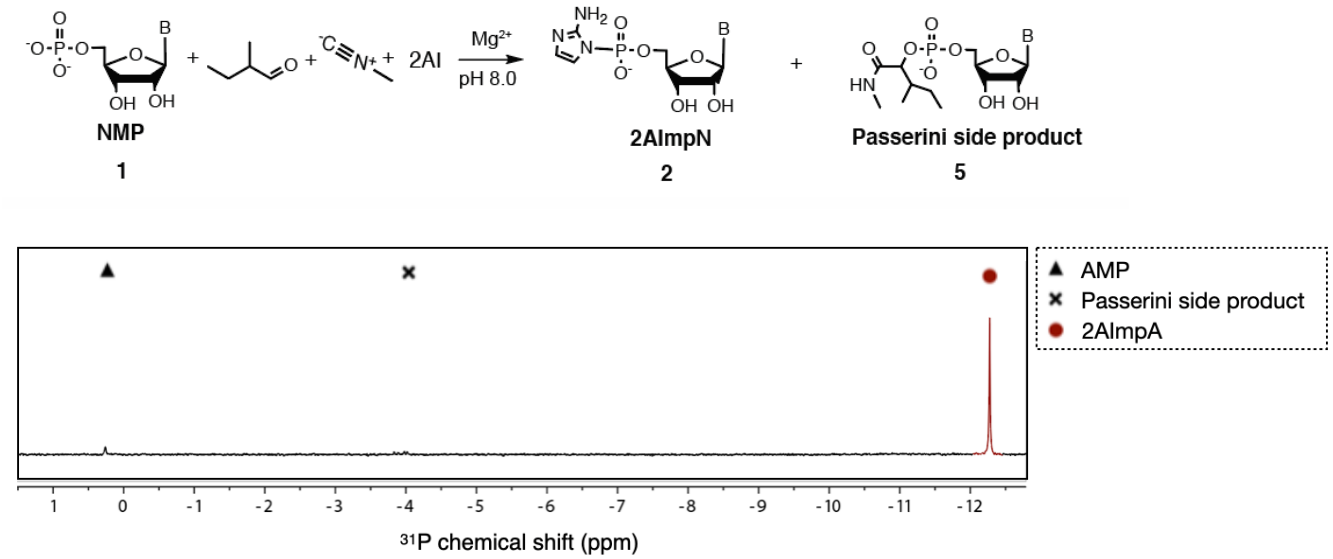
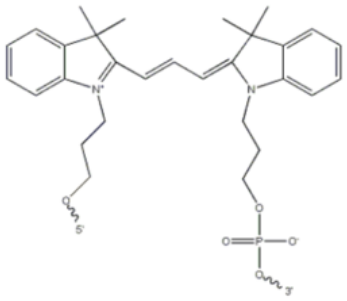
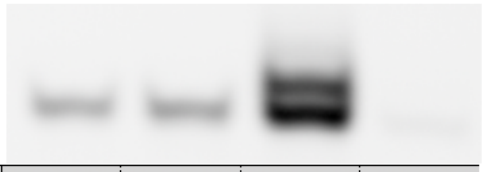


Figure S2.4. 2-methyl butyraldehyde (2MBA) yields efficient activation with minimal Passerini side product. A ^{31}P-NMR spectrum of the standard activation reaction mixture in 10% D_2O in H_2O at t = 6 hr.
Reaction conditions: AMP (10 mM), 2MBA (200 mM), methyl isocyanide (200 mM), Mg^{2+} (30 mM, $MgCl_2$), HEPES (200 mM) at pH 8, and 2AI (100 mM).

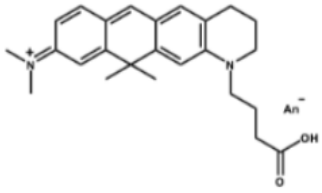
CY3

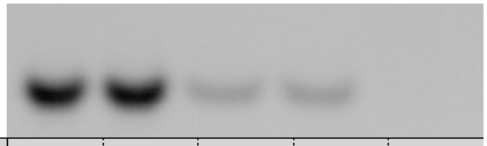




Primer	+	+	+	+
Methyl isocyanide		+	+	+
2-methyl butyraldheyde			+	+
2AI				+

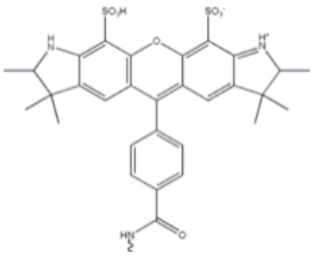
ATTO-610

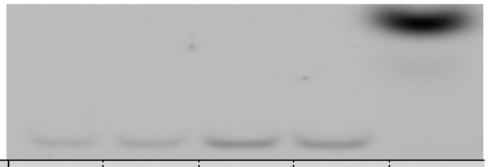




Primer	+	+	+	+	+
Methyl isocyanide		+		+	+
2-methyl butyraldheyde			+	+	+
2AI				+	+

ALEXA-532

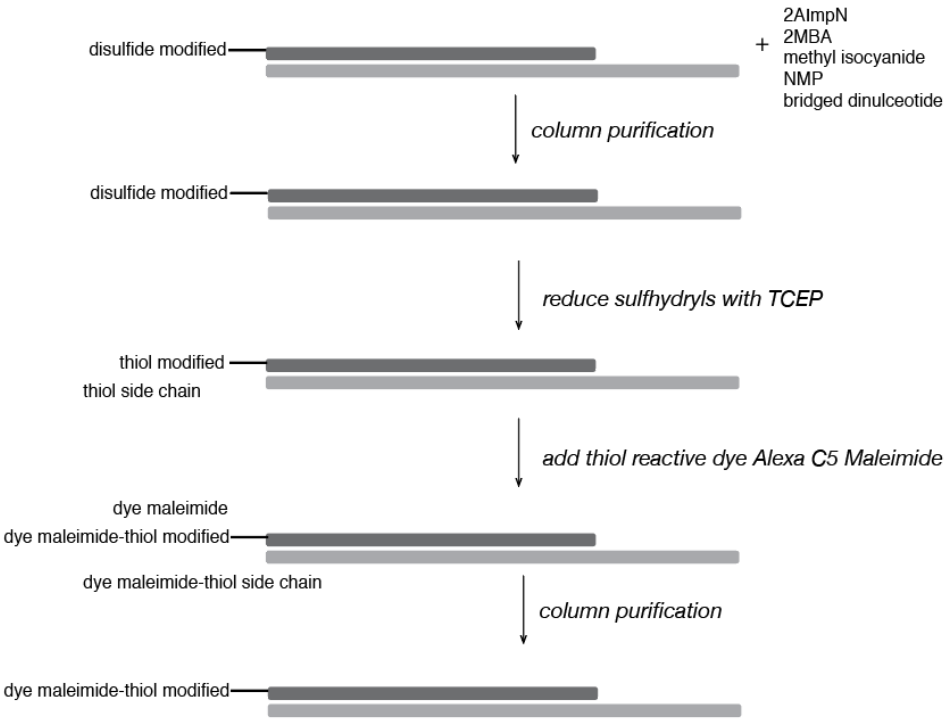


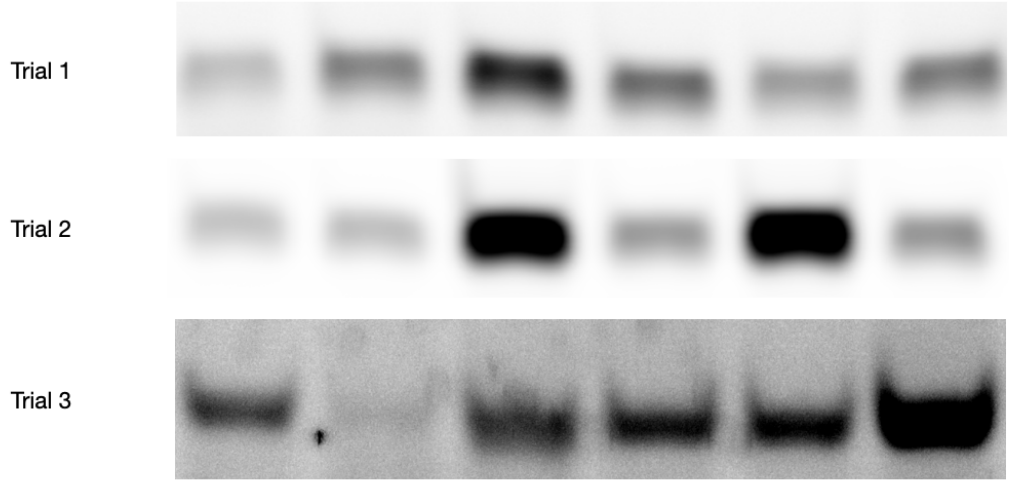


Primer	+	+	+	+	+
Methyl isocyanide		+		+	+
2-methyl butyraldheyde			+	+	+
2AI				+	+

DYE	OBSERVATIONS
CY3	An additional band was observed when the primers were treated with isocyanide and aldehyde.
ATTO-610	The disappearance of the primer band was observed on treatment with isocyanide and aldehyde, with or without 2AI.
ALEXA-532	A band shift was observed when the primers were treated with isocyanide, aldehyde, and 2AI.

Figure S2.5. Instability of dye-labeled primers to isocyanide nucleotide activation chemistry. Fluorescent dyes expected to be the least likely to react with the isocyanide nucleotide activation chemistry and within the relevant imaging wavelength range were tested. Most dyes contain either carboxyl and/or amine functional groups. We tested dyes that contain only hindered tertiary amines, which are less likely to participate in the chemistry. However, PAGE analysis revealed additional bands, band shifts, and band disappearances.





Primer	+	+	+	+	+	+
Methyl isocyanide		+			+	+
2-Methyl butyraldehyde			+		+	+
2AI				+		+

Figure S2.6. Schematic representation of post-labeling used to visualize the primer extension. From left to right, each primer was treated with: (1) nothing, (2) methyl isocyanide, (3) 2MBA, (4) 2AI, (5) methyl isocyanide and 2MBA, (6) methyl isocyanide, 2MBA, and 2AI. Three trials are shown. The band intensity varies from trial to trial, probably due to variability in purification and labeling efficiency, which should not affect internally normalized quantifications of primer extension efficiencies.

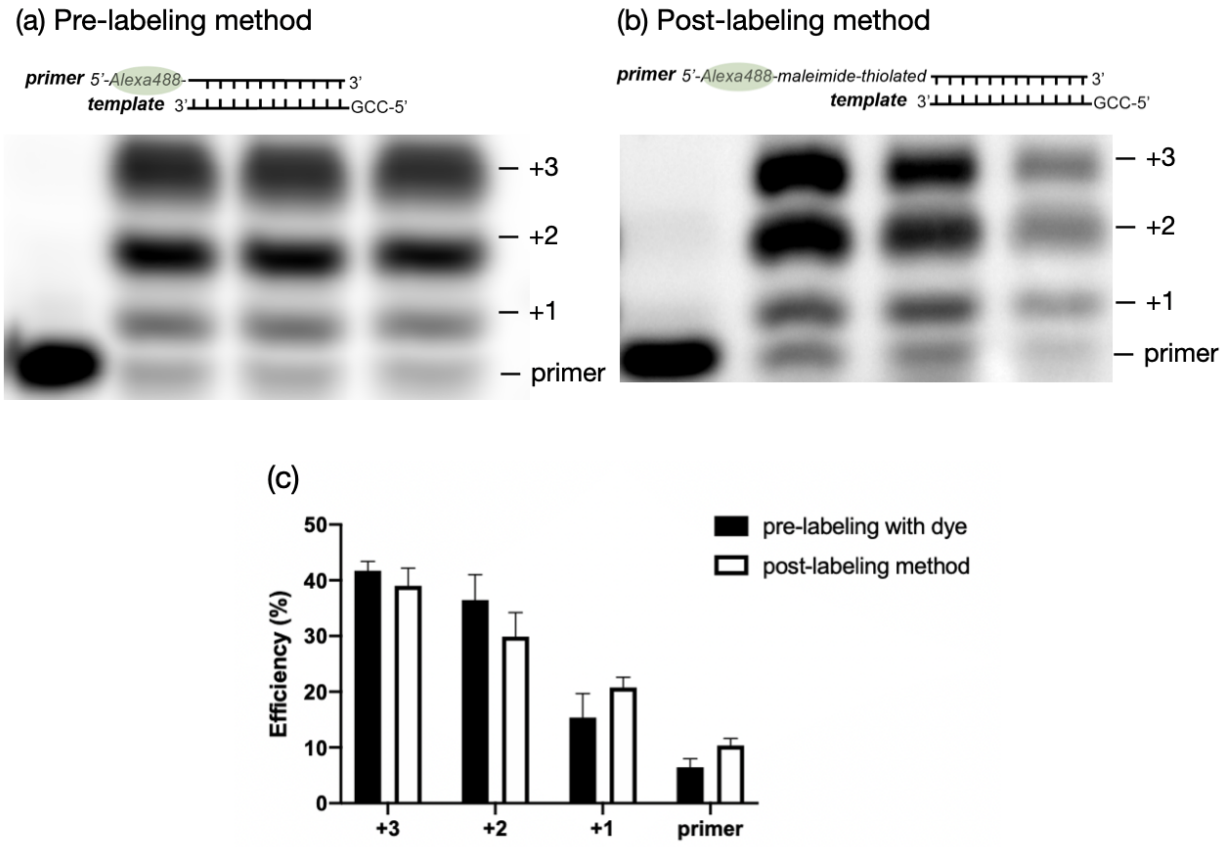
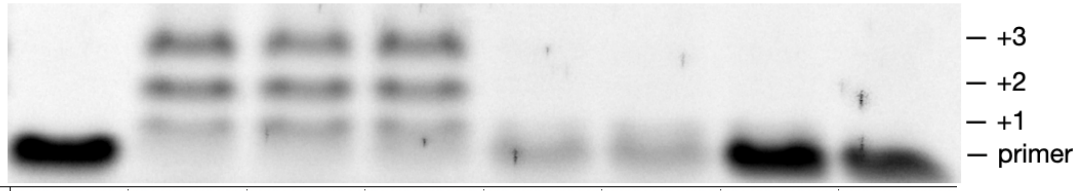


Figure S2.7. Primer extension efficiency measured by pre-labeling or post-labeling methods in the absence of isocyanide chemistry. Purified activated monomers were added to (a) the pre-labeled and (b) thiol-modified primer (triplicate data shown in both cases). A quantification of the results across the two methods (c) shows very good agreement. The variance in intensity seen in (b) is consistent with Figure S8.



	1	2	3	4	5	6	7	8
2AImpC&G (mM)	-	10 each	10 each	10 each	10 each	10 each	-	-
CMP&GMP (mM)	-	-	-	-	-	-	10 each	10 each
2AI (mM)	-	-	-	-	200	200	200	200
Isocyanide chemistry	-	-	-	-	-	-	+	+

Figure S2.8. Excess 2AI inhibits primer extension. In the presence of 200 mM 2AI, only minimal +1 product was observed in reactions with activated mononucleotides, or unactivated mononucleotides with isocyanide chemistry (lanes 5-8). This is in contrast to the positive control (lanes 2-4), where primers were extended to +3 in the absence of added 2AI.

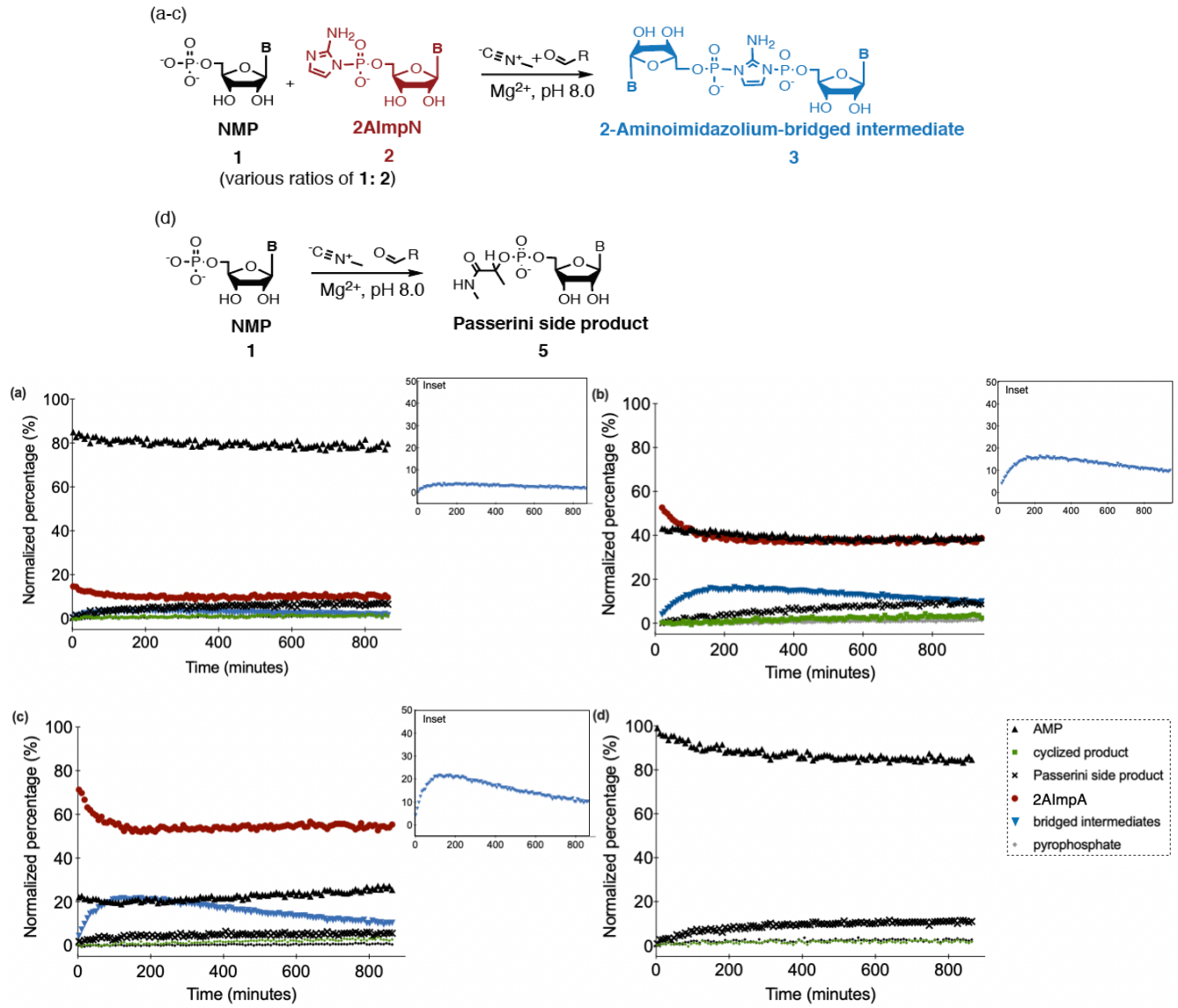


Figure S2.9. Isocyanide chemistry drives bridged dinucleotide formation with various ratios of NMP to 2AImpN. Bridged dinucleotides can be generated across various ratios of 2AImpA to AMP: (a) ~15:85, (b) ~50:50, (c) ~80:20. Analyses of the reactions over the course of 15 hours; insets are from the same data, highlighting the bridged dinucleotides. (d) No bridged dinucleotide in the absence of activated monomer at the ratio of 2AImpA to AMP at 0:100. The relative percentage of each species is calculated based on the corresponding peak integration normalized to the number of phosphorus atoms. Passerini side product formation exhibits no correlation with the ratio of 2AImpA to AMP, although 5% more of it forms than in the absence of activated monomer because the reaction proceeds via the imidoyl phosphate **4** (Scheme S2.1), which undergoes competing intramolecular attack by hydroxyl to form **5**.
Reaction conditions: varying ratio of 2AImpA/AMP (10 mM total nucleotides), Mg^{2+} (30 mM, $MgCl_2$), HEPES (200 mM) at pH 8 with or without methyl isocyanide (200 mM), and 2MBA (200 mM).

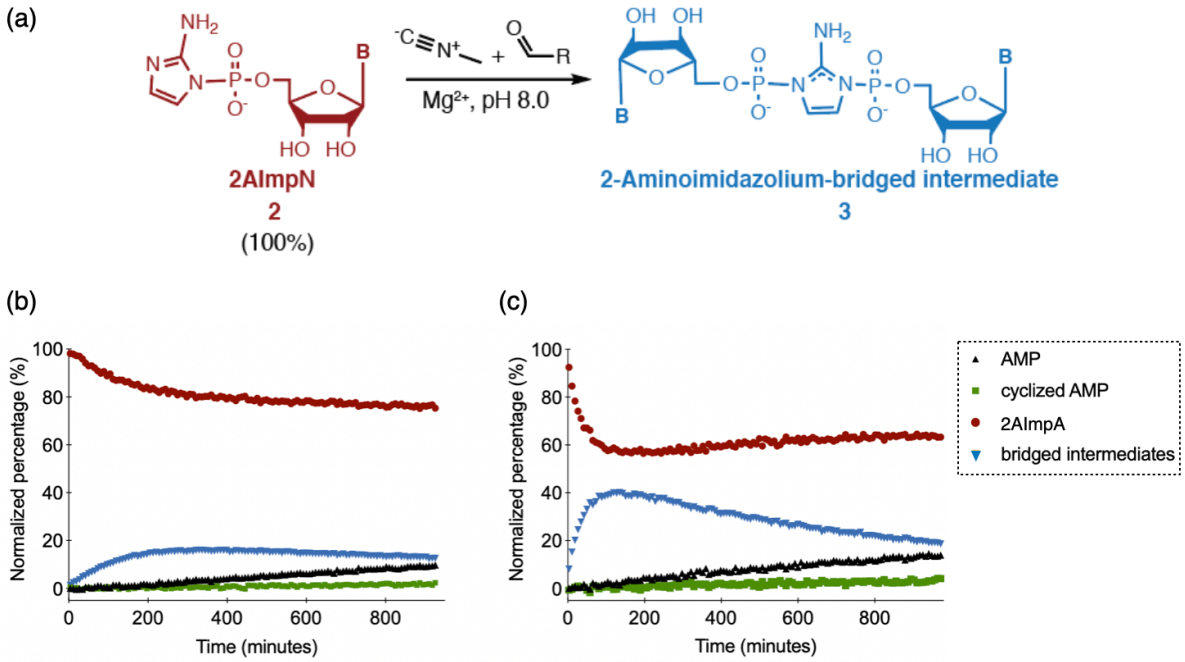


Figure S2.10. Isocyanide chemistry drives bridged dinucleotide formation even with only 2AI-activated mononucleotides as substrates. (a) The prebiotically plausible synthesis of bridged dinucleotide using only 2AI-activated monomers. Analyses of the reaction over the course of 15 hours (b) without and (c) with bridge-forming chemistry. The relative percentage of each species is calculated based on the corresponding peak integration normalized to the number of phosphorus atoms.
Reaction conditions: 2AImpA (10 mM), Mg^{2+} (30 mM, $MgCl_2$), HEPES (200 mM) at pH 8 with or without methyl isocyanide (200 mM), and 2MBA (200 mM).

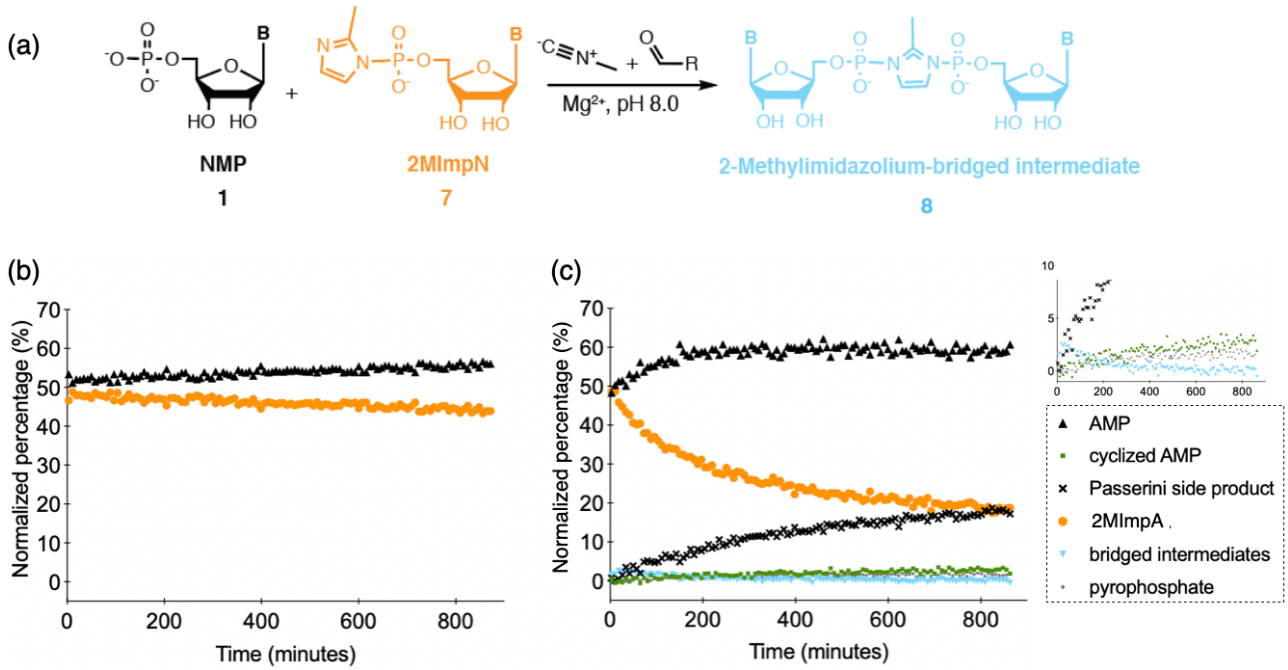


Figure S2.11. Bridge-forming chemistry acting on an equimolar mixture of AMP and 2MImpA. (a) Scheme for synthesis of bridged intermediates from a mixture of AMP and 2MImpA. Analyses of the reaction over the course of 15 hours (b) without and (c) with bridge-forming chemistry. The relative percentage of each species is calculated based on the corresponding peak integration normalized to the number of phosphorus atoms. (Much higher concentrations of activated mononucleotides are required to spontaneously form the bridged dinucleotide with 2MI, which explains the lack of bridged dinucleotide in the control case. (Walton and Szostak 2017)) Reaction conditions: 2MImpA (5 mM), AMP (5 mM), Mg^{2+} (30 mM, $MgCl_2$), HEPES (200 mM) at pH 8 with or without methyl isocyanide (200 mM), and 2MBA (200 mM).

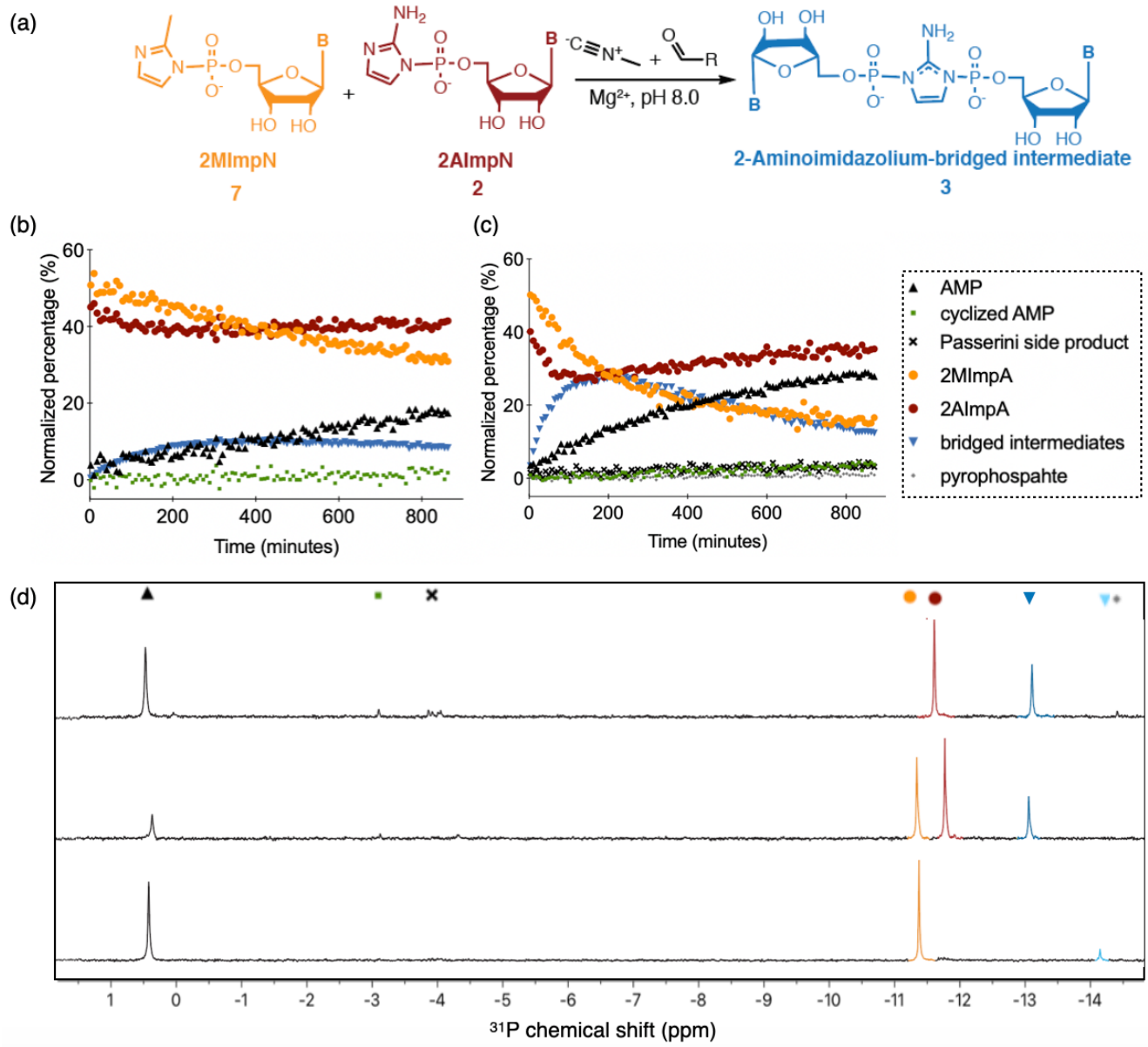


Figure S2.12. The selectivity of bridge-forming chemistry towards 2-aminoimidazolium bridged dinucleotides. (a) Scheme for synthesis of bridged dinucleotide using an equimolar mixture of 2AImpA and 2MImpA. Analyses of the reaction over the course of 15 hours (b) without and (c) with bridge-forming chemistry. The relative percentage of each species is calculated based on the corresponding peak integration normalized to the number of phosphorus atoms. (d) The only bridged species (δ -12.95) present in the synthesis can be assigned to 2AImpA because the shift of 2-aminoimidazolium bridged dinucleotide is at δ -13.0, whereas 2-methylimidazolium bridged dinucleotide is at δ -14.02.
Reaction conditions: 2AImpA (5 mM), 2MImpA (5 mM), Mg^{2+} (30 mM, $MgCl_2$), HEPES (200 mM) at pH 8 with or without methyl isocyanide (200 mM), 2MBA (200 mM).

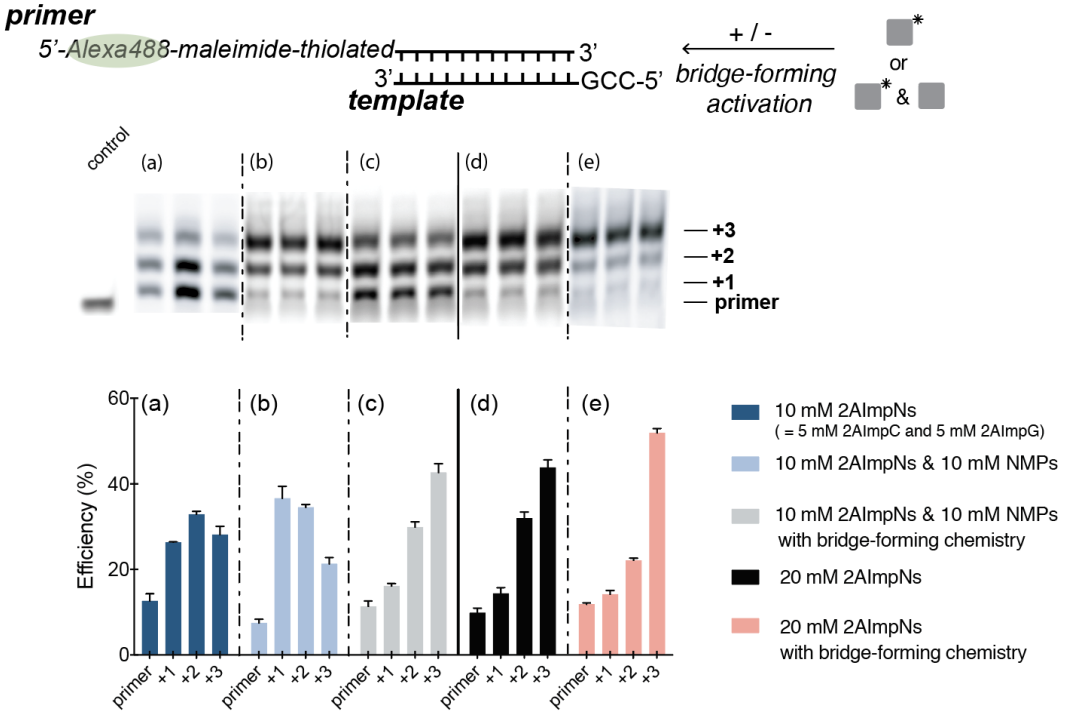
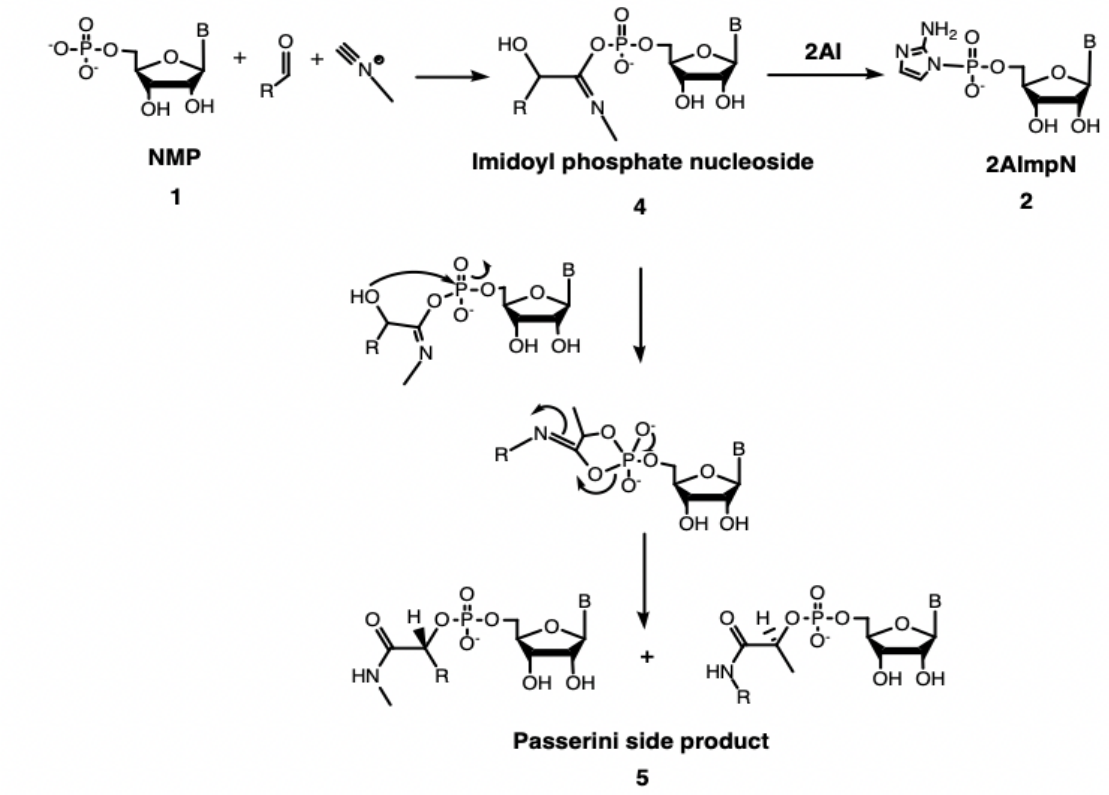


Figure S2.13. Raw PAGE data of primer extension products used for Figure 3 under various conditions at t = 24 hours. Positions of primer and +1 to +3 products are indicated. Reaction conditions: the corresponding amount of 2AImpNs and NMPs (as indicated below) were added to primer (1 μM, green), template (1.5 μM, black), HEPES (200 mM) pH 8, and Mg^{2+} (30 mM, $MgCl_2$) with isocyanide chemistry (light blue and pink) or without (dark blue, grey, and black). n = 3 replicates.

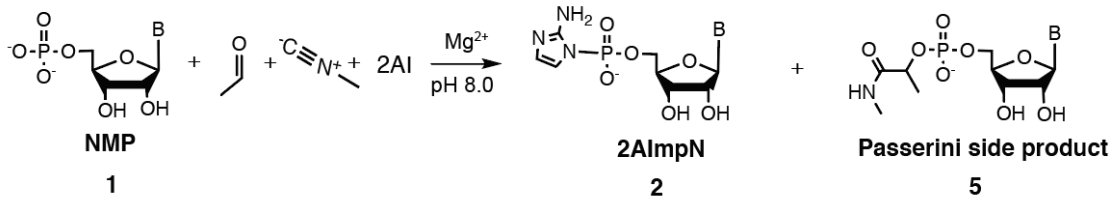
Supplementary scheme



Scheme S2.1. Proposed mechanism for Passerini side product **5** formation.

Supplementary tables

Table S2.1. Yields and ratios of 2AImpA **2** to Passerini side product **5** across different conditions, as measured by ^{31}P-NMR spectroscopy.

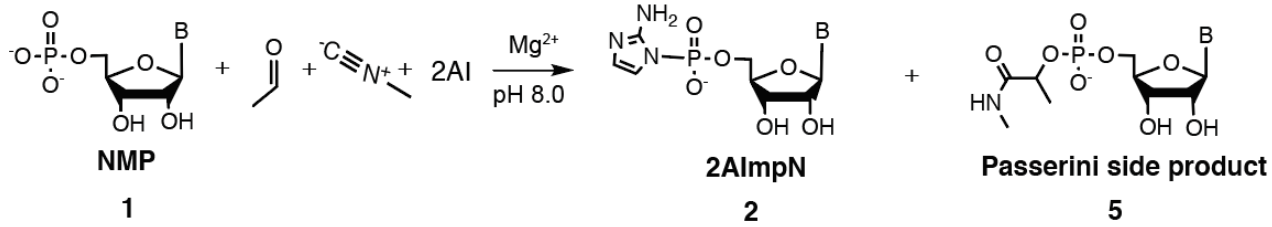


pH	[Acetaldehyde] mM	[Methyl isocyanide] mM	[Mg^{2+}] mM	Buffer	[2AImpN **2**]/[Passerini product **5**]	Activation (**2**) yield %
			Mg^{2+} Concentration			
8	200	200	0	HEPES	4.6±0.5	31±2
8	200	200	10	HEPES	2.9±0.3	48±2
8	200	200	20	HEPES	2.3±0.1	46±1
8	200	200	30	HEPES	1.9±0.1	33.0±0.5
8	200	200	50	HEPES	1.2±0.1	32±2
			pH			
6	200	200	10	MES	4.0±0.4	38±2
7	200	200	10	HEPES	4.2±0.4	57±2
8	200	200	10	HEPES	2.9±0.3	48±2
			Methyl Isocyanide Concentration			
8	200	200	30	HEPES	2.3±0.4	36±2
8	200	300	30	HEPES	2.9±0.2	45±4
8	200	400	30	HEPES	2.7±0.8	50±4
8	200	800	30	HEPES	5.0±0.9	58±4
			Acetaldehyde Concentration			
8	300	200	30	HEPES	2.3±0.4	45±2
8	400	200	30	HEPES	2.9±0.6	50±3
8	800	200	30	HEPES	2.2±0.3	36±1
			Acetaldehyde and Methyl Isocyanide Concentrations			
8	30	30	30	HEPES	6±6	0.6±0.6
8	50	50	30	HEPES	5.4±0.7	11.9±0.1
8	400	400	30	HEPES	3.5±0.2	62±2
8	600	600	30	HEPES	3.7±0.3	71±1
8	800	800	30	HEPES	3.5±0.8	69±4

(1) All reactions were carried out using AMP (10 mM), 2AI (100 mM), and buffer (200 mM) in addition to the materials listed in the table. Errors are standard deviations of the mean, n = 3 replicates.

(2) 200 mM buffer solutions were used to maintain nearly constant pH values over a given reaction course. We note that HEPES proved an ideal buffer because it exhibited no discernible chemical reactions with the species in the activation chemistry (Table S1). The carboxylate groups of BICINE will undergo typical Passerini-type chemistry, resulting in a lower concentration of the other components of the reaction, and hence less efficient activation. Similarly, the TRIS buffer interacts with aldehyde as described previously (Ogilvie and Whitaker 1976).

Table S2.2. Yields and ratios of 2AImpA **2** to Passerini side product **5** with various

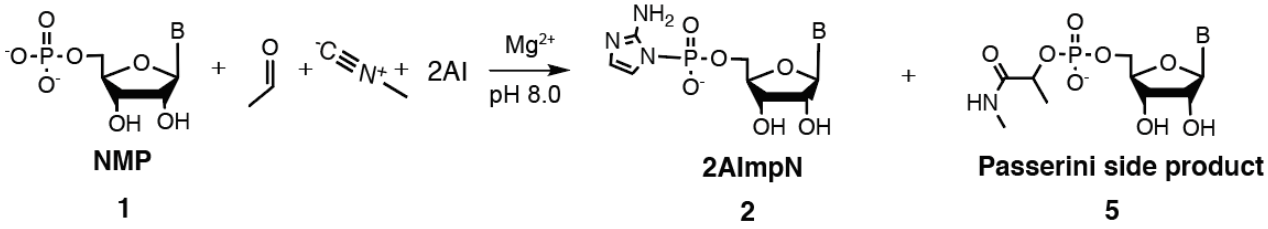


aldehydes/ketones measured by NMR spectroscopy.

Aldehyde or Ketone	Activation (**2**) yield%	[2AImpN **2**]/ [Passerini product **5**]
Glycolaldehyde	23±1	2.1±0.2
Acetaldehyde	31.0±0.7	3.0±0.5
Propionaldehyde	50±1	6±1
Isobutyraldehyde	63.0±0.6	12.7±0.3
2-Methylbutyraldehyde	64±2	18±3
2-Methylbutyraldehyde*	89±1	23±3
Acetone	0	NA

All reactions were carried out using AMP (10 mM), aldehyde/ketone (200 mM), methyl isocyanide (200 mM), 2AI (200 mM), Mg2+ ($MgCl_2$, 10 mM), and BICINE (200 mM) at pH 8.0. Errors are standard deviations of the mean, n = 3 replicates.
*Other than BICINE being replaced by HEPES buffer, the reaction conditions were the same. A corresponding NMR spectrum for 2-methylbutyraldehyde can be found in Figure S5.

Table S2.3. Yields and ratios of 2AImpN to Passerini side product across different conditions measured by NMR spectroscopy. All reactions were carried out using NMP (10 mM), acetaldehyde (200 mM), methyl isocyanide (200 mM), 2AI (100 mM), and buffer (200 mM) in addition to the materials listed in the table. n = 3 replicates.



NMP	[Mg^{2+}] mM	Buffer	[2AImpN]/ [Passerini product]	Activation yield %
		Nucleotide Identity		
GMP	30	HEPES	2.2±0.3	39±3
CMP	30	HEPES	2.0±0.2	40±3
UMP	30	HEPES	1.5±0.3	30±3
AMP	30	HEPES	2.3±0.4	36±2
		Buffer Identity		
AMP	10	HEPES	2.9±0.3	48±2

AMP	10	TRIS	6±1	33±4
AMP	10	BICINE	3.0±0.5	31.0±0.7

200mM buffer solutions were used to maintain nearly constant pH values over a given reaction course. We note that HEPES proved an ideal buffer because it exhibited no discernible chemical reactions with the species in the activation chemistry (Table S1). The carboxylate groups of BICINE will undergo typical Passerini-type chemistry, resulting in a lower concentration of the other components of the reaction, and hence less efficient activation. Similarly, the TRIS buffer interacts with aldehyde as described previously [6]

Table S2.4. Sequences of oligonucleotides used in this study.

Name	RNA Sequence (5' → 3')
control primer	/Alexa488/ AGU GAG UAA CUC
thiolated primer	/5ThioMC6-D/ AGU GAG UAA CUC
template	CCG GAG UUA CUC ACU
complementary	AGU GAG UAA CUC CGG

Supplementary text

a. Identification of the Dead-end Nucleotide Side Product
b. Development of a New Primer Extension Assay
c. Characterization of Bridge-forming Activation

a. Identification of the Dead-end Nucleotide Side Product

A solution of 200 mM isocyanide, 200 mM acetaldehyde, 100 mM 2AI, and 10 mM AMP was monitored in 10% D_2O by ^{31}P NMR over 12 hours. Besides the monophosphate starting material **1** and activated mononucleotides **2**, which were confirmed by the spike-in of their corresponding authentic standards, a pronounced nucleotide-containing side product was found (Figures S3, S4). Isocyanides, aldehydes, and phosphates are known to undergo Passerini-type addition reactions. (Sutherland, Mullen et al. 2008) Direct inject electrospray mass spectrometry (ESI-MS) of these reaction mixtures showed an m/z consistent with a Passerini product **5** (Scheme S2.1). Specifically, an otherwise unattributed m/z peak at 431.1 is found for the reaction using AMP, which is consistent with the full Passerini reaction product **5** (exact expected m/z = 431.11) in which intramolecular attack of the hydroxy group at the phosphorus atom is followed by a rearrangement (Scheme S2.1). In addition, the unique chiral center of **5** is in agreement with the observed diastereomers in the ^{31}P NMR spectrum (Figure S2.3). The resultant amidoyl moiety of **5** is not a likely substrate for nucleophilic attack given the lack of resonance from oxygen. This is in contrast with the transient imidoyl phosphate intermediate **4** (Scheme S2.1). We therefore ascribe the diastereomers in ^{31}P NMR spectrum to the Passerini reaction product **5**, which we deem undesirable because it is not expected to contribute to RNA polymerization and may even inhibit the reaction by competing for binding sites on the template.

b. Development of a New Primer Extension Assay

Typical experiments utilize a fluorescent dye-labeled primer which is then characterized by polyacrylamide gel electrophoresis (PAGE) in which different length primer extension products

yield a characteristic banding pattern on the gel. However, many dyes contain functional groups such as carboxyls and amines that are expected to be incompatible with the isocyanide activation chemistry. We screened a variety of labeled primers representing multiple dye families against the activation chemistry and its constituents. In all cases, an additional band or band shift was observed due to chemical modification of the dye, in many cases probably via Passerini-type reactions (Figure S2.6). We therefore sought a post-labeling strategy in which a dye is introduced after the reactions being probed have taken place. A thiol-modified primer was anticipated to be unreactive to isocyanide activation chemistry, thereby allowing the subsequent conjugation of a fluorophore-maleimide through the sulfhydryl-maleimide coupling. With this strategy, we observed no additional bands or band shifts compared to the control primer among samples that had been exposed to various combinations of isocyanide activation chemistry constituents for 24 hours (Figure S2.7). Although absolute band intensities vary due to differences in purification and labeling efficiencies, only relative intensities within each lane are used to quantify the primer extension results. We next tested the accuracy of this post-labeling method by preparing a standard primer extension reaction with the template oligo 5′-CCGGAGUUACUCACU-3′, in which the template sequence 5′-CCG-3′ is copied in the presence of a primer and purified 2AImpG and 2AImpC. The quantification of primer extension using a thiol-modified primer is in very good agreement with the quantification using the standard dye-labeled primer (Figure S2.8). We conclude that this post-labeling approach is an appropriate alternative technique to quantifying primer extension under otherwise dye-prohibitive conditions.

c. Characterization of Bridge-forming Activation

We note that the pool of unactivated mononucleotides **1** contributes to bridged dinucleotide **3** formation because the change in activated mononucleotide **2** concentration over the course of a reaction is insufficient to account for all the formed bridged dinucleotide **3**. For example, in the case

of 50:50 2AImpN:NMP (Figure 2.2b), the activated mononucleotide **2** concentration decreased by 13% from t = 12 min. to t = 229 min. Therefore, a maximum of 6.5% of the bridged intermediate species **3** could come solely from the activated mononucleotides **2** (two mononucleotides being required to generate a single bridged dinucleotide **3**). (Note that the self-reaction of activated monomers **2** only formed less than 3% bridged dinucleotides **3** (Figure 2.2c), which cannot account for the 6.5% increase). However, over the same time period, there was a 12% increase in the bridged dinucleotide **3** concentration. Recalling that there is no free 2AI here, these observations are consistent with bridged dinucleotides forming from (i) the activation chemistry generating an imidoyl phosphate **4** from an initially unactivated mononucleotide **1**, (ii) displacement of the imidoyl moiety by N3 of 2AI(mpN) **2**, and (iii) the background self-reaction of activated mononucleotides **2** as per the established pathway. (i) and (ii) are analogous to the isocyanide-mediated activation of mononucleotides in the presence of free 2AI, except in this case the 2AI is already attached to another mononucleotide (Figure 2.1).

Chapter 3

Freeze-thaw cycles enable a prebiotically plausible and continuous pathway from nucleotide activation to nonenzymatic RNA copying

Materials and methods
Supplementary figures S3.1 to S3.7

Materials and methods

General information

Materials. Reagents and solvents were obtained at the highest available purity from Acros Organics, Alfa Aesar, Fisher Scientific, Sigma-Aldrich, ThermoFisher Scientific, or Tokyo Chemical Industry Co., and were used without further purification unless noted below. Nucleoside-5'-monophosphates, free acid, were purchased from Santa Cruz Biotechnology. 2-aminoimidazole hydrochloride and 2,2'-dipyridyldisulfide were purchased from Combi Blocks. RNA oligonucleotides for standard primer extension experiments were purchased from Integrated DNA Technologies. The synthesis and storage of methyl isocyanide was as previously described (Zhang, Duzdevich et al. 2020). All reactions were carried out in DNase/RNase-free distilled water.

NMR spectroscopy. ^{1}H and ^{31}P-NMR spectra were obtained using a Varian INOVA NMR spectrometer operating at 400 MHz and 161 MHz respectively. Samples in H_2O/D_2O mixtures were analyzed using Wet1D suppression to collect ^{1}H-NMR data. Chemical shifts (δ) are shown in ppm. Coupling constants (J) are given in Hertz (Hz) and the notations s, d, t, and m represent singlet, doublet, triplet, and multiplet multiplicities respectively.

pH measurements. pH values were determined by a micro pH probe (Orion 9863BN) equipped with a needle tip and a SevenCompact meter (Mettler Toledo S220).

Data analysis. All NMR spectra were analyzed using MestReNova (version 12.0.3). The yields of conversion were determined by the relative integration of the signals in the ^{1}H or ^{31}P NMR spectra.

Preparation, storage, and concentration determination of stock solutions

Stock solution of nucleoside-5'-monophosphate disodium salt was prepared in DNase and RNase-free distilled water. After adjusting the pH to the reported values with NaOH/HCl, the stock solution was filtered with 0.22-micron syringe filters (Millipore Sigma). Each stock solution was then aliquoted and kept at -20 °C until further use. Stock solution of 2-aminoimidazole hydrochloride was prepared in DNase and RNase-free distilled water, filtered with 0.22-micron syringe filters (Millipore Sigma), and kept at -20 °C until further use. The pH was adjusted to the reported values with NaOH/HCl right before the reaction to prevent possible polymerization at basic pH.

The concentrations of the nucleoside-5'-monophosphate solutions were determined by analysis of serial dilutions on a UV spectrophotometer. The absolute concentrations of other stock solutions were determined by comparing the integrals of ^{1}H-NMR peaks of interest to the calibrant, trimethyl phosphate, by NMR spectroscopy.

Sample preparation for ice eutectic phase experiments

Reaction mixtures were first flash-frozen by immersion in liquid nitrogen, followed by incubation at -13 °C in a standard laboratory freezer unit for the indicated time (Kanavarioti, Monnard et al. 2001, Monnard, Deamer et al. 2004, Attwater, Wochner et al. 2013).

In figure 1, reaction conditions are listed as follows: 10 mM pA, 10 mM 2AI, 100 mM 2MBA, 30 mM $MgCl_2$, 50 mM Na^+-HEPES pH 8, plus (C) one addition of 200 mM MeNC at ice eutectic phase or room temperature, or (D) periodic addition of the MeNC beyond the initial 50 mM in three aliquots of 50 mM each under ice eutectic phase at -13 °C or (E) at room temperature.

Primer extension reactions using only monomers

Control primer extension reactions. The primer-template duplex was first annealed in a 40 μl solution containing 2 μM thiol-modified primer (/5ThioMC6-D/AGUGAGUAACUC, IDT; see 1.6 below for use of the thiol to dye-label products), 3.1 μM template (CCGGAGUUACUCACU, IDT), 125 mM Na^+-HEPES (pH 8.0), and 62.5 mM $MgCl_2$ by heating at 95 °C for 3 min followed by cooling to 23 °C at a rate of 0.1 °C/s. The reaction was initiated by the addition of 10 μl 100 mM activated mononucleotides. *Eutectic phase primer extension reactions.* The primer-template duplex was first annealed in a 100 μl solution containing 2.4 μM thiol-modified primer (/5ThioMC6-D/AGUGAGUAACUC, IDT), 3.0 μM template (CCGGAGUUACUCACU, IDT), 125 mM Na^+-HEPES (pH 8.0) by heating at 95 °C for 3 min followed by cooling to 23 °C at a rate of 0.1 °C/s. The reaction was initiated by the addition of 100 μl mixture of 10 mM pN, 10 mM 2AI, 100 mM 2MBA, 30 mM $MgCl_2$, 50 mM Na^+-HEPES pH 8, after four cycles of ice eutectic phase activation (50 mM MeNC each addition). For bridge-forming activation, the reaction mixture was incubated at room temperature with an additional 100 mM 2MBA and 200mM MeNC for 24 hours.

Trimer-assisted primer extension reactions

Control primer extension reactions. The primer-template duplex was first annealed in a 40 μl solution containing 3.6 μM thiol-modified primer (/5ThioMC6-D/ CGCUCGACUG, IDT), 4.5 μM template (5' GCGCCUCAUCAGUCGAGCG, IDT), 125 mM Na^+-HEPES (pH 8.0), and 62.5 mM $MgCl_2$ by heating at 95 °C for 3 min followed by cooling to 23 °C at a rate of 0.1 °C /s. The final reaction mixture contained 0.5 mM each activated trimer and 5 mM each activated monomer. *Eutectic phase primer extension reactions.* The primer-template duplex was first annealed in a 100 μl solution containing 2.4 μM thiol-modified primer (/5ThioMC6-D/AGUGAGUAACUC, IDT), 3.0 μM template (CCGGAGUUACUCACU, IDT), 125 mM Na^+-HEPES (pH 8.0) by heating at 95 °C for 3 min followed by cooling to 23 °C at a rate of 0.1 °C/s. The reaction was initiated by the

addition of 100 µl mixture of 5 mM each pN, 0.5 mM each $(pN)_3$, 23 mM 2AI, 100 mM 2MBA, 50 mM $MgCl_2$, 50 mM Na^+-HEPES pH 8, and periodic addition of MeNC in four aliquots of 50 mM each under ice eutectic phase. The reaction mixture was then further incubated with an additional 100 mM 2MBA and 100mM MeNC on day 5 and day 6.

Sample preparation and product analysis

Sample Purification and preparation (adapted from reference (Zhang, Duzdevich et al. 2020)*).* 10 µl aliquots were removed at given time points and desalted using a ZYMO Oligo Clean & Concentrator spin column (ZYMO Research). The isolated material was resuspended in 30 µl 100 mM Na^+-HEPES (pH 7.50) and disulfide bonds were reduced using a 10-fold molar excess of tris-(2-carboxyethyl) phosphine hydrochloride (TCEP). Alexa 488 C5 maleimide dissolved in anhydrous dimethyl sulfoxide (DMSO) at a concentration of 1 mM was added to the primer-template duplex dropwise, and the coupling reaction was allowed to proceed in the dark at room temperature for 2 hours. The labeled primer-template duplex was separated from free dye using ZYMO DNA Clean & Concentrator spin columns, resuspended in 5 µl of 100 mM Na^+-HEPES buffer, and mixed with 30 µl of quenching buffer containing 8.3 M urea, 1.3x Tris/Borate/EDTA (TBE) buffer (pH 8.0), 0.005% Bromphenol Blue, 0.04% Orange G, and 75 µM RNA complementary to the template. To denature the labeled product from the template prior to polyacrylamide gel analysis, and sequester the template with its complement (added in excess as a competitor), the sample was heated to 95 °C for 3 min and cooled to 23 °C at 0.1 °C/s.

Analysis of the polymerization products. 20% polyacrylamide gels were prepared using the SequaGel–UreaGel system (National Diagnostics, Atlanta, GA). The gels were allowed to pre-run at a constant power of 20 W for 40−45 mins. 15 µl aliquots of samples were separated by 20% (19:1) denaturing PAGE at a constant power of 5 W for 10 min. and then 25 W for 1.5 hrs. Gels were imaged on a Typhoon 9410 scanner (GE Healthcare, Little Chalfont, Buckinghamshire, UK) and quantified using

the accompanying ImageQuant TL software. All reactions were performed in duplicate or greater as indicated.

Supplementary figures

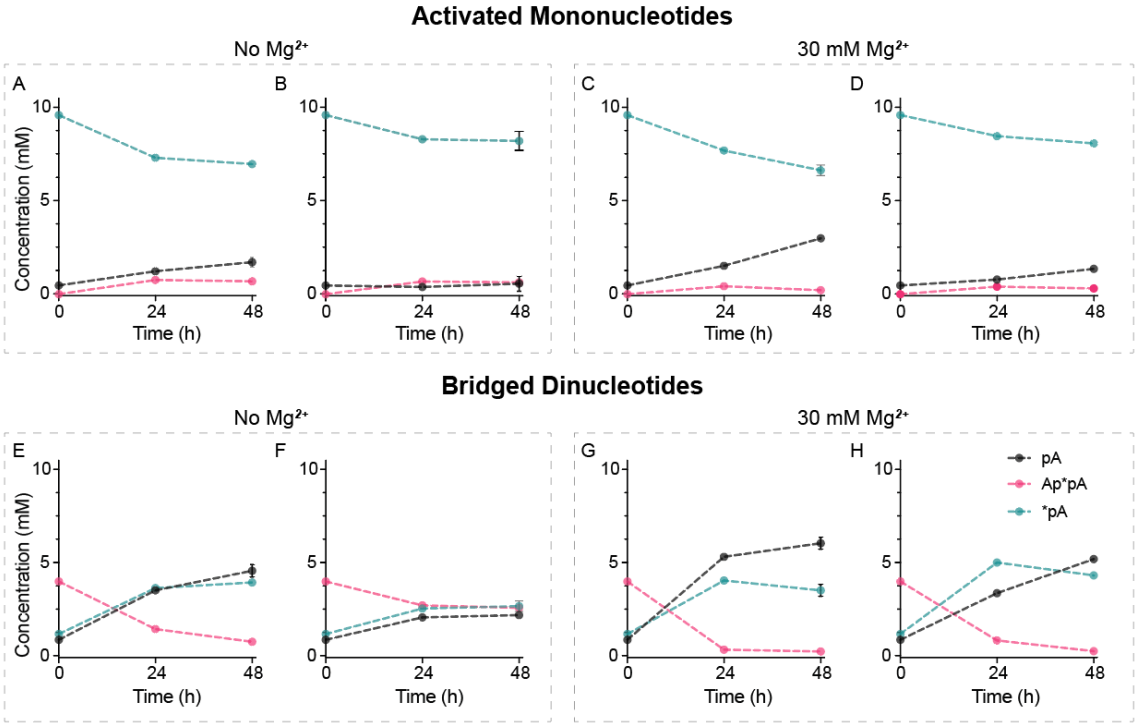


Figure S3.1. Activated nucleotides (*pA) are less susceptible to hydrolysis under ice eutectic conditions at -13 °C compared to room temperature. In the absence of Mg^{2+}, more *pA hydrolyzed to pA at (A) room temperature compared to (B) ice eutectic phase. A similar level of bridged dinucleotides is formed under either condition. With 30 mM Mg^{2+}, more *pA hydrolyzed to pA at (C) room temperature compared to (D) ice eutectic phase. In addition, the presence of Mg^{2+} accelerated the hydrolysis either under room temperature (compare (A) and (C)) or ice eutectic phase (compare (B) and (D)). In the absence of Mg^{2+}, there is almost no Ap*pA left under (E) room temperature while over 50% of Ap*pA remained under (F) ice eutectic phase. The presence of 30 mM Mg^{2+} significantly increased the hydrolysis rate of Ap*pA under either (G) room temperature or (H) ice eutectic phase. However, importantly, at the 24-hr timepoint, less Ap*pA has been hydrolyzed under the ice eutectic phase (H) compared to (G) room temperature. Note that the effective concentration of Mg^{2+} in the ice eutectic phase might be much higher than 30 mM. *Reaction conditions*: 10 mM *pA, 200 mM Na^{+}-HEPES at pH 8 under (A) room temperature and (B) ice eutectic phase at -13 °C. 10 mM *pA, 30 mM Mg^{2+}, 200 mM Na^{+}-HEPES at pH 8 under (C) room temperature and (D) ice eutectic phase. 10 mM Ap*pA, 200 mM Na^{+}-HEPES at pH 8 under (E) room temperature and (F) ice eutectic phase. 10 mM Ap*pA, 30 mM Mg^{2+}, 200 mM Na^{+}-HEPES at pH 8 under (G) room temperature and (H) ice eutectic phase.
Errors are standard deviations of the mean, $n \geq 2$ replicates.

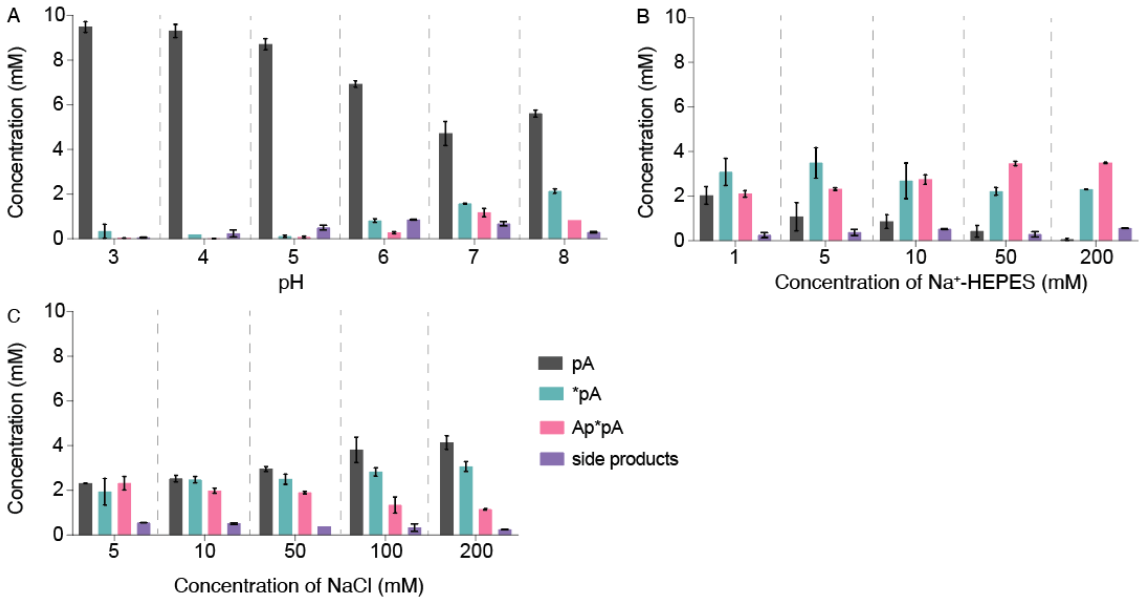


Figure S3.2. The importance of pH effects under ice eutectic phase. (A) Eutectic activation was performed across a range of initial pH values (3 to 8). The solution was brought to the indicated pH by addition of concentrated HCl or NaOH. No buffering agent was used. The optimal initial pH is 7. The low yields of activation compared to the case in which Na^+-HEPES buffer is used (compare to (B)) suggests that the initial adjustment of the pH is not sufficient to regulate the pH during the incubation under ice eutectic phase. (B) Increasing Na^+-HEPES (up to 200 mM) increased the activation yield. Note that significant activation is observed with only 1 mM Na^+-HEPES, much less than the concentrations of MeNC and 2MBA, suggesting that Na^+-HEPES is not actively participating in the isocyanide activation. (Note that the imidazolide moiety of 2AI can provide buffering, however, 2AI is consumed during the reaction over time.) (C) To confirm that the increase in reaction yield in the presence of Na^+-HEPES is not due to a simple salt effect, we titrated NaCl into the reaction. Increasing the concentration of NaCl leads to an opposite trend compared to that observed when increasing Na^+-HEPES, suggesting that simple ionic effects are not responsible for facilitating the reaction. We note that chloride can strongly affect the pH in the eutectic phase since it is retained in the ice, drawing H^+ into the ice for charge balance, making the liquid phase more alkaline (Hayyan, Mjalli et al. 2012). *Reaction conditions*: all experiments were performed using the same approach with four cycles as optimized for pA and shown in Figure 1D. (A) 10 mM pA, 10 mM 2AI, 100 mM 2MBA, 30 mM $MgCl_2$, 50 mM MeNC plus periodic addition of the MeNC under ice eutectic phase. (B) 10 mM pA, 10 mM 2AI, 100 mM 2MBA, 30 mM $MgCl_2$, 50 mM MeNC plus periodic addition of the MeNC, with indicated Na^+-HEPES concentration at pH 8 under ice eutectic phase. (C) 10 mM pA, 10 mM 2AI, 100 mM 2MBA, 30 mM $MgCl_2$, 50 mM MeNC plus periodic addition of the MeNC, 5 mM Na^+-HEPES pH 8, and indicated amount of NaCl under ice eutectic phase. Errors are standard deviations of the mean, $n \geq 2$ replicates.

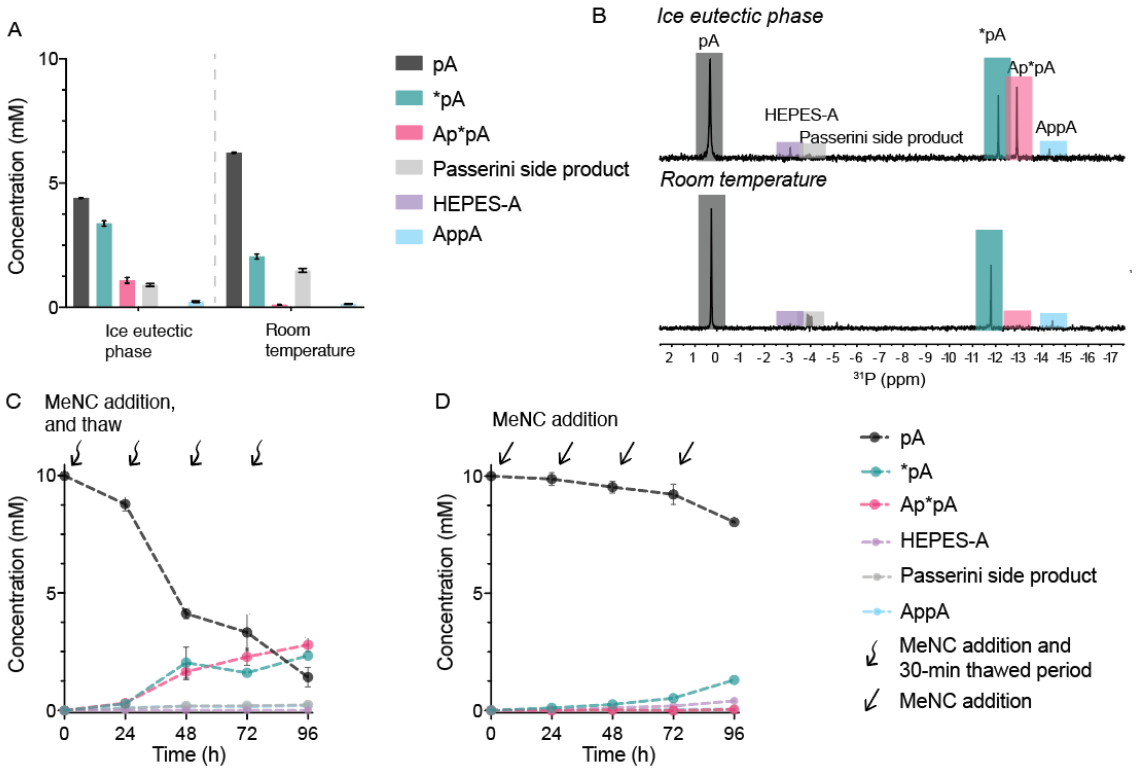


Figure S3.3. Ice eutectic phase significantly enhances the yield of activated nucleotides (*pN) and promotes formation of bridged dinucleotides (Np*pN). The same data as in Figure 1 C-E, but with a complete breakdown of the identities of the side products. (A) Effect of ice eutectic phase on nucleotide activation. Nucleotide activation at ice eutectic phase (left) and room temperature (right) with one round of MeNC addition and equimolar concentrations of initially unactivated nucleotides and 2AI. (B) Representative NMR spectra of (A). (C) Cycles of MeNC addition and eutectic freezing drive significant nucleotide activation. Controlled addition of MeNC to the thawed solution enhances the yield of activated nucleotides (*pA). The curvy arrows indicate the addition of MeNC after thawing of the reaction mixture. Note that the reaction mixture is frozen as eutectic ice for the entire reaction course except during the brief thawing steps (~30 minutes at room temperature for the mixture to thaw completely). (D) Activation is inefficient in the absence of ice eutectic freezing, at room temperature. The straight arrows indicate the addition of MeNC at room temperature. *Reaction conditions*: 10 mM pA, 10 mM 2AI, 100 mM 2MBA, 30 mM $MgCl_2$, 50 mM Na^+-HEPES pH 8, plus (A) one addition of 200 mM MeNC under ice eutectic phase or room temperature, or (C) periodic addition of the MeNC beyond the initial 50 mM in three aliquots of 50 mM each under ice eutectic phase at -13 °C or (D) at room temperature. Errors are standard deviations of the mean, n ≥ 2 replicates.

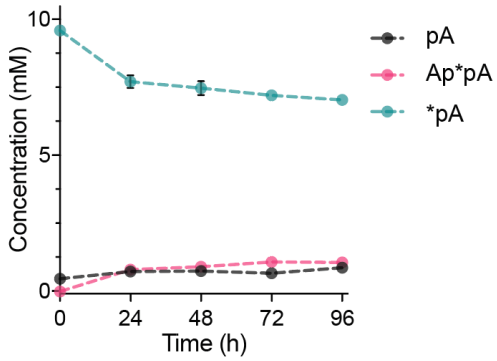


Figure S3.4. Spontaneous bridged dinucleotide formation from pure activated nucleotide (*pA) under ice eutectic phase. The concentration of bridged dinucleotides (Ap*pA) slowly increased over time as 10 mM of *pA was incubated under ice eutectic phase in the absence of isocyanide-driven activation chemistry (Figure S3B). The reaction mixture was under ice eutectic phase conditions except during the four brief thawing steps to make measurements. *Reaction conditions:* 10 mM *pA, 50 mM Na^+-HEPES pH 8, 30 mM $MgCl_2$, under ice eutectic phase. Errors are standard deviations of the mean, n = 2 replicates.

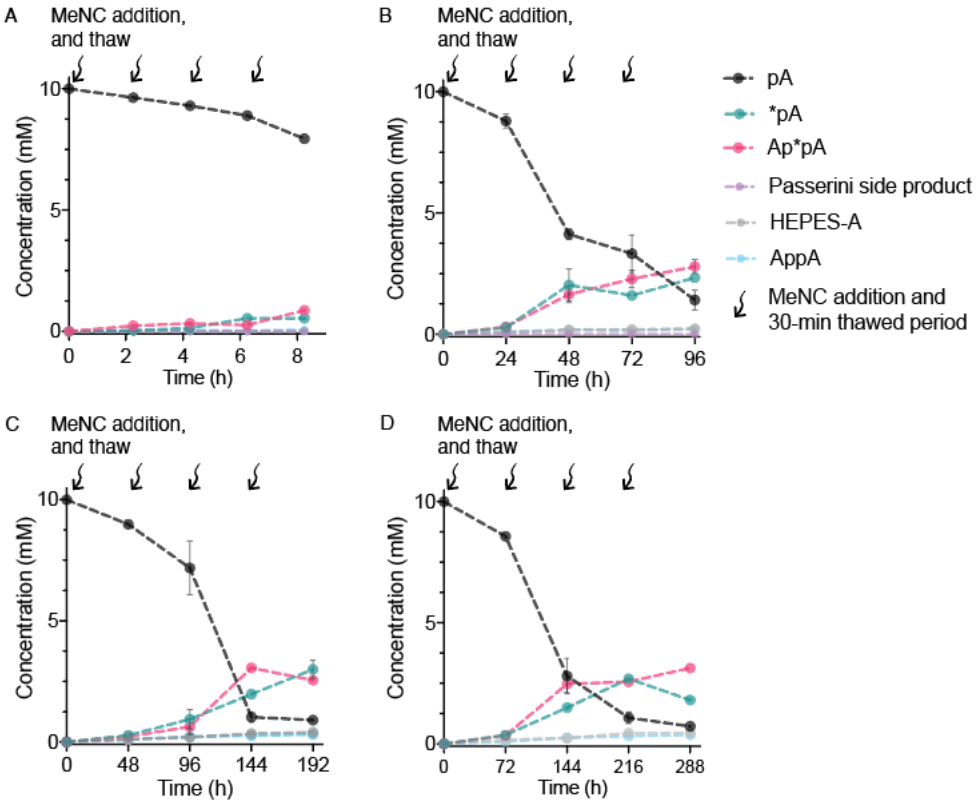


Figure S3.5. The frequency of MeNC delivery at long time intervals does not affect the overall yield of activated products. The same data as in Figure 2, but with a complete breakdown of the identities of the side products. (A) Yields of activated species are low with time intervals on the order of hours in between MeNC additions. (B-D) However, lengthening of the time interval beyond 24 hours does not significantly affect the final yield. The complete product breakdown is included in Figure S5. *Reaction conditions*: 10 mM pA, 10 mM 2AI, 100 mM 2MBA, 50 mM MeNC, 30 mM $MgCl_2$, 50 mM Na^+-HEPES pH 8, and subsequent periodic addition of MeNC in three aliquots of 50 mM. Errors are standard deviations of the mean, n ≥ 2 replicates.

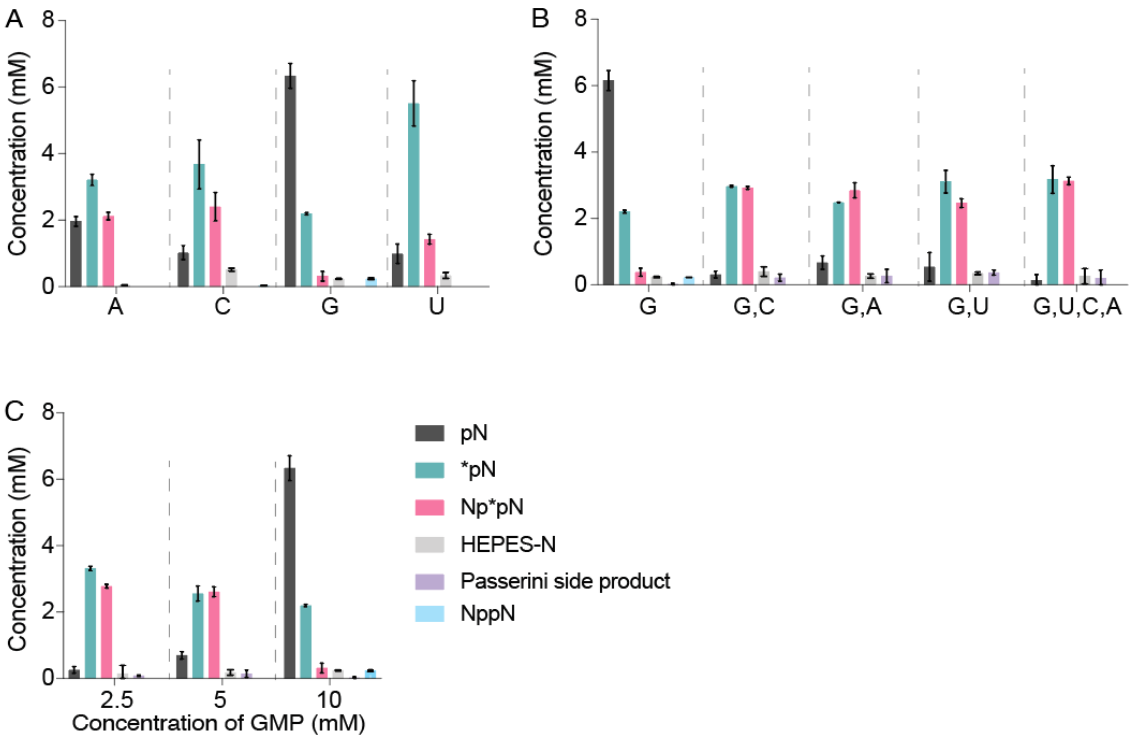


Figure S3.6. Efficient eutectic ice phase activation of all four canonical nucleotides. The same data as in Figure 3, but with a complete breakdown of the identities of the side products. (A) Each of the four canonical ribonucleotides could be activated using the same approach as optimized for pA in Figure 1D. However, pG exhibited lower activation compared to the other three. 10 mM pN, 10 mM 2AI, 100 mM 2MBA, 50 mM MeNC, 30 mM $MgCl_2$, 50 mM Na^+-HEPES pH 8, and periodic addition of MeNC in three aliquots of 50 mM each. (B) All four nucleotides can be activated to almost full yield with minimal side products when incubated simultaneously. Indicated concentrations are sum totals across all the nucleotides. Total [pN] = 10 mM, 10 mM 2AI, 100 mM 2MBA, 50mM MeNC, 30 mM $MgCl_2$, 50 mM Na^+-HEPES pH 8, and periodic addition of MeNC in three aliquots of 50 mM each. (C) GMP activation yields decrease as GMP concentration is increased, suggesting that its low activation is partly due to its low solubility in aqueous phase. Total [pN] = 10 mM, 10 mM 2AI, 100 mM 2MBA, 50 mM MeNC, 30 mM $MgCl_2$, 50 mM Na^+-HEPES pH 8, and periodic addition of MeNC in three aliquots of 50mM each. Errors are standard deviations of the mean, $n \geq 2$ replicates.

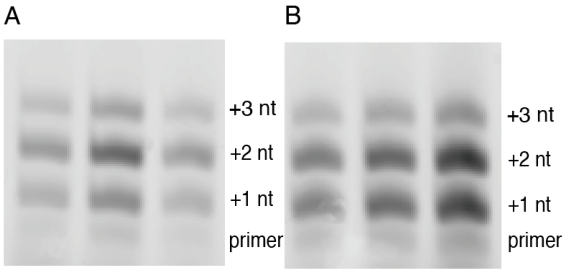


Figure S3.7. Eutectic ice phase completes a pathway from unactivated mononucleotides through activation to primer extension. Replicates of primer extension reactions using (A) pure activated mononucleotides (5 mM *pG and 5 mM *pC) under room temperature and (B) pNs (pN = pC and pG) generated with isocyanide activation chemistry under ice eutectic phase. Positions of primer and +1 to +3-nucleotide (nt) extension products, which represents +C, +CG, +CGG, are indicated. *Reaction conditions:* (A) 5 mM *pG and 5 mM *pC were added to 1 μM primer, 1.5 μM template, 200 mM Na^+-HEPES pH 8.0, and 30 mM $MgCl_2$. (B) 1 μM primer and 1.5 μM template were added to a mixture of 10 mM pN, 10 mM 2AI, 100 mM 2MBA, 30 mM $MgCl_2$, 50 mM Na^+-HEPES pH 8, after four cycles of eutectic phase activation (50 mM MeNC each addition).

Chapter 4

Endocytosis in protocells enables nutrient transport bypassing permeability limits

Materials and methods

General Information

Materials. Oleic acid (C18:1) was purchased from NuChek Prep (Elysian, MN). RNA labeled with fluorescein (5'-cy5- $C_{(10)}$ $A_{(18)}$-3') was synthesized at IDT (Coralville, IA). All other chemicals were purchased from Sigma-Aldrich (St. Louis, MO) and were used without any further purification. Data analysis was performed using ImageJ (version 1.53a) and GraphPad Prism (version 8.4.0).

Preparation of Micelles and Giant Unilamellar Vesicles (GUVs)

100 mM oleate micelles were prepared by dissolving 50 μmol of neat oleic acid oil in 1 equivalent of NaOH solution to a volume of 500 μL with Millipore water (18.2 M cmΩ). Fatty acid GUV sample was made by resuspending micelle solution in buffer stock (1 M Na-bicine, pH 8.45), and Millipore water to the final concentration of 5mM oleic acid, 50 mM Na-bicine buffer, and 200 mM sucrose as described previously. [22] Encapsulation of fluorescent dyes (1 mol% relative to the amount of neat oil) was achieved by mixing the solute with the resuspension buffer before adding the neat oil.

Vesicle Endocytosis

Vesicles labeled with dyes were diluted to 0.5 mM amphiphile with a buffer consisting of 50 mM bicine, pH 8.45, and 200 mM glucose and a 100 μL aliquot of the vesicle suspension was carefully transferred into a tube. (These vesicles were subsequently diluted 1: 9 into the buffer to a concentration of 0.5 mM allowing good contrast of vesicles to the background.) Vesicle endocytosis

was initiated by adding micelles from a 100 mM stock solution and Na-bicine buffer from a 1 M stock to their corresponding concentration (summarized and listed in following Table) with mixing by inverting the tube for ~5 s. Vesicle suspensions were allowed to equilibrate for at least 1 h before microscopy.

V(Na-bicine); V(micelle) (uL)	micelles (relative to the fatty acid concentration)			
osmotic shock [Na-bicine] (mM)	1eqv.	2 eqv.	5 eqv.	10 eqv.
25	2.7; 0.5	2.7; 1	2.7; 2.5	2.7; 5
50	5.56; 0.5	5.56; 1	5.56; 2.5	5.56; 5
100	11.76; 0.5	11.76; 1	11.76; 2.5	11.76; 5

Washing Endocytosed Vesicles

To remove the free HPTS during imaging, after endocytosis, the vesicles were diluted 1:10 into a 200 mM glucose buffer that was identical in composition and pH to the original buffer but lacked the lipids and encapsulated solute. After centrifugation at 2000 g for 30 s, the top 90% of volume was removed by pipetting. The remaining solution was then agitated to resuspend the vesicles.

Confocal Microscopy

Confocal images were collected using a Nikon A1R HD25 confocal laser scanning microscope equipped with LU-N4/N4S 4-laser unit.

Immersed boundary method simulations

To simulate growing, permeable vesicles under osmotic shock, we use the immersed boundary method to simulate the coupled fluid-structure interaction of permeable vesicles growing in an aqueous incompressible solution (Wu, Fai et al. 2015). Given the complexity of possible changes in vesicle morphology, it is important to perform simulations in three dimensions, and for computational efficiency we use the parallel GPU-accelerated code described in Fai et al. (Fai, Griffith et al. 2013).

For simplicity, we hold the inner and outer osmoloyte concentrations constant and do not account for solute exchange across the (i.e. we set $P_s = 0$ in the notation of Sacerdote et al. (Sacerdote and Szostak 2005). Given that the rate of solute exchange is typically much lower than that of solvent, this approximation is expected to be reasonable on the relatively short timescales of interest here.

	Symbol	**Value**	**Units**
Initial diameter	R	1	mm
Permeability	P_f	14e-3	cm/s
Growth rate	g	0-0.05	s^{-1}
Viscosity	m	0.01	dyn/cm×s
Density	r	1	g/cm^3
Osmolarity	Dc	1-50	mM
Specific volume	v_w	20	cm^3/mol
Spontaneous curvature	c_0	-1	mm
Bending modulus	k_B	4	k_BT
Total time	T	1-3	s

Table S4.1: Parameters used in simulations of growing permeable vesicles under osmotic shock.

Supplementary figures

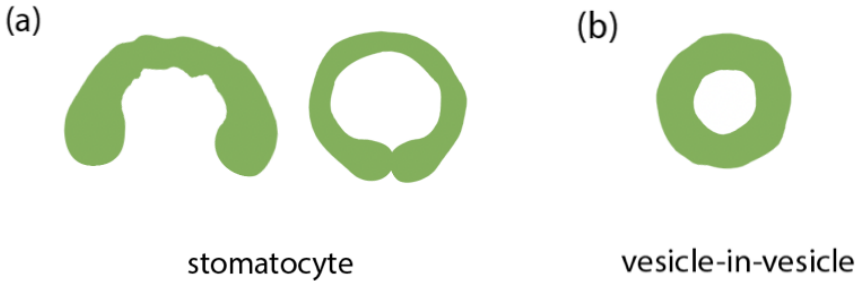


Figure S4.1. Schematic of topological transformations from a GUV into a stomatocyte shape (a), or a vesicle-in-vesicle structure (b).

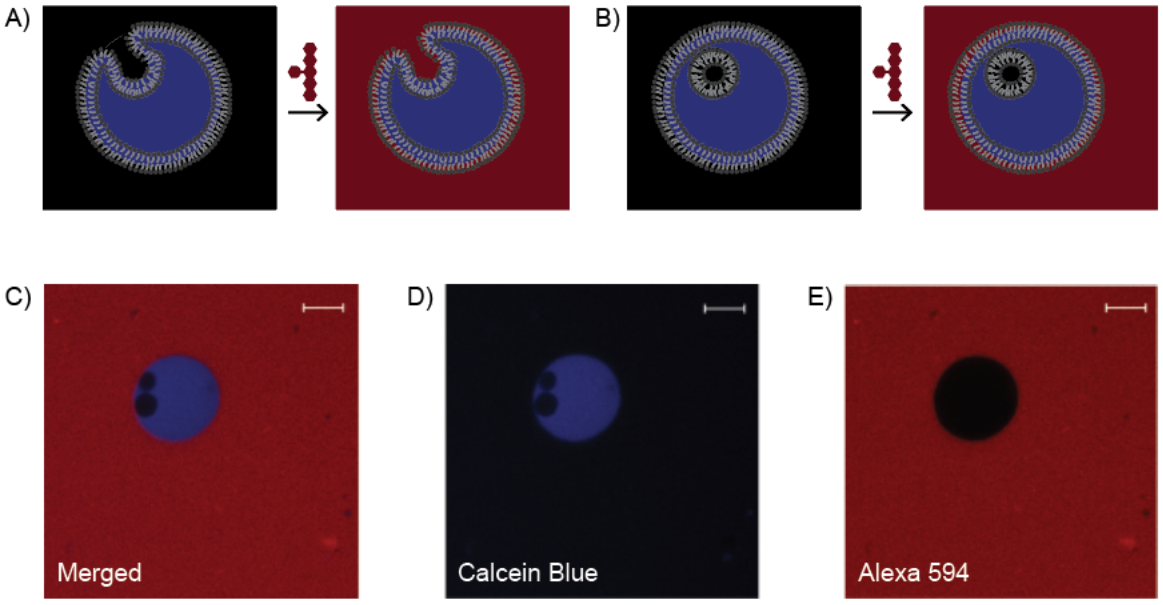


Figure S4.2. Schematic of adding an aqueous dye (Alexa 594 hydrazide, red) to the endocytosed vesicles (encapsulating calcein blue, blue) that are either connected to the outside (A) or not (B). Confocal micrographs of the compartmentalized vesicles after addition of aqueous dye, including merged image (C) of two separate channel calcein blue (D) and Alexa 594 (E). Scale bards represent 5 μm.

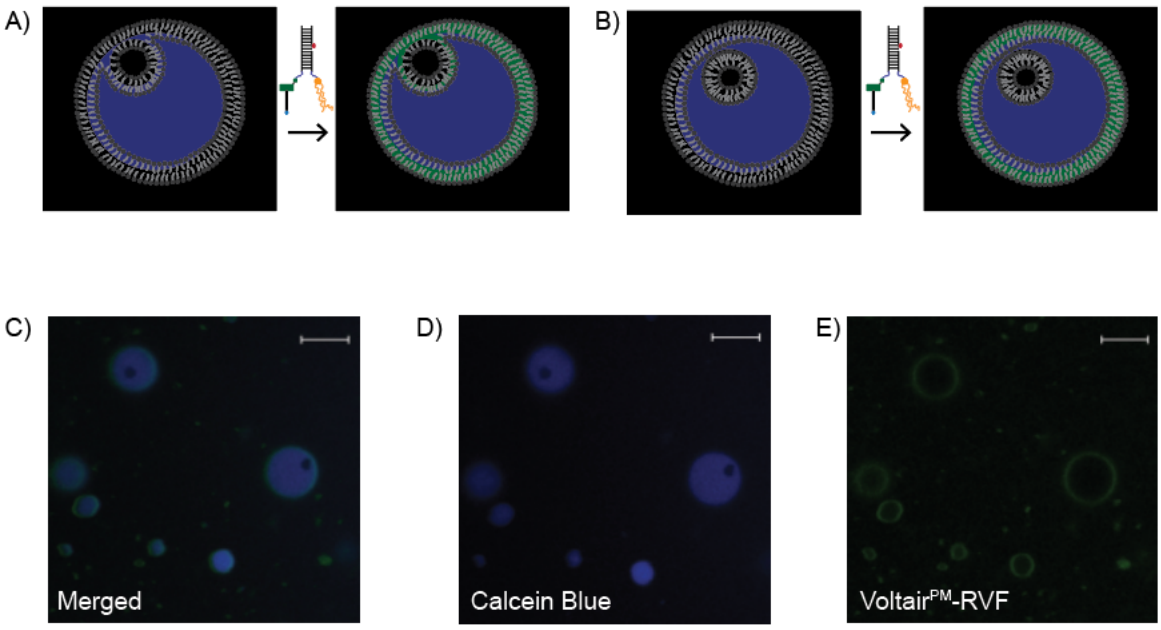


Figure S4.3. Schematic of adding Voltair^{PM} (green) to the endocytosed vesicles (encapsulating calcein blue, blue) that are either connected to the outer membranes (A) or not (B). Confocal micrographs of the compartmentalized vesicles after addition of aqueous dye, including merged image (C) of two separate channels, calcein blue (D) and Voltair^{PM} (green) (E). Scale bards represent 5 μm.

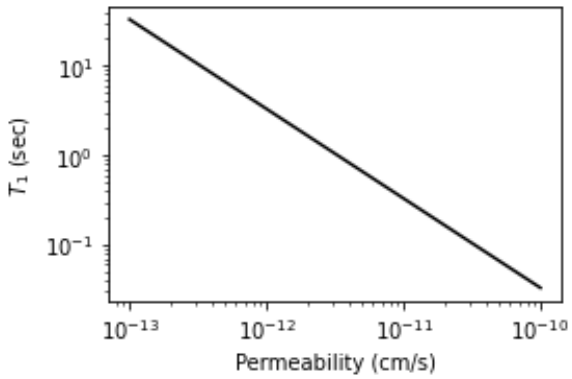


Figure S4.4. The average time T_1 required for a single molecule to enter a 4-μm-diameter vesicle, for a range of permeabilities spanning values for common nutrients (10^{-10} and 10^{-13} cm/s).

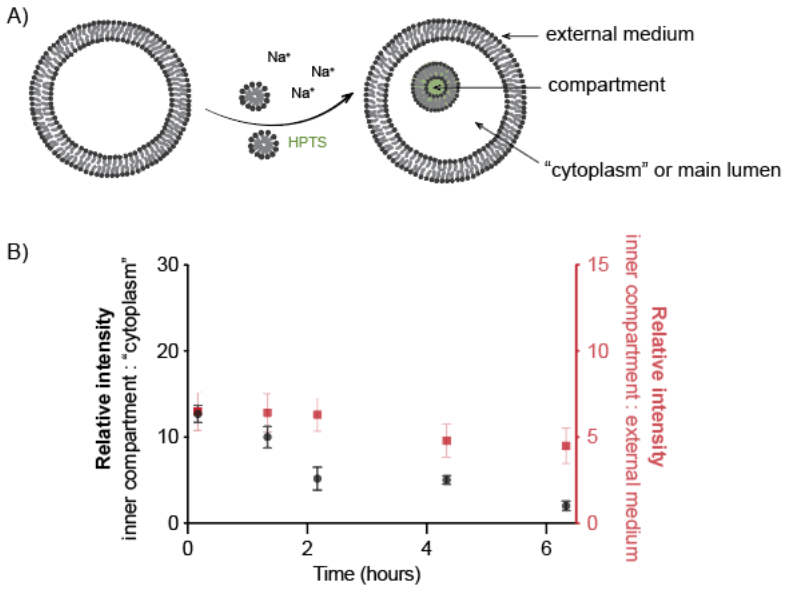


Figure S4.5. Internal compartments can slowly deliver nutrients of low permeability to the main lumen or 'cytoplasm' of the protocell. (A) HPTS, a relatively impermeant molecule, is encapsulated in the interior compartment via one hyperosmotic shock coupled with micelle addition (ΔCV = 100 mM Na-bicine, ΔCA = 2.5 mM oleate, CHPTS = 2 mol% of the total lipid). (B) The fluorescence intensity ratio of 'cytoplasm' over compartment increases and of compartment over external medium decreases indicate that HPTS inside the compartment is slowly entering the 'cytoplasm'.

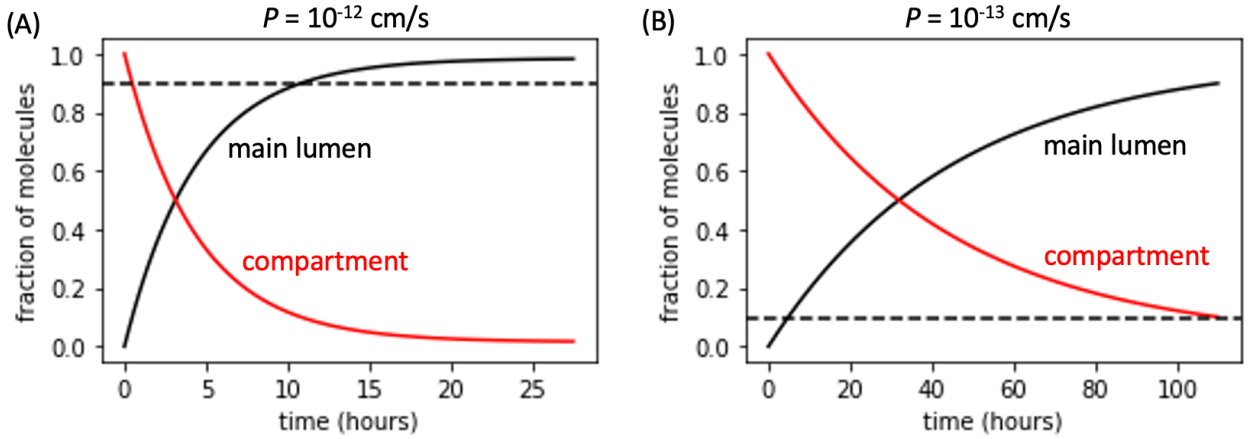
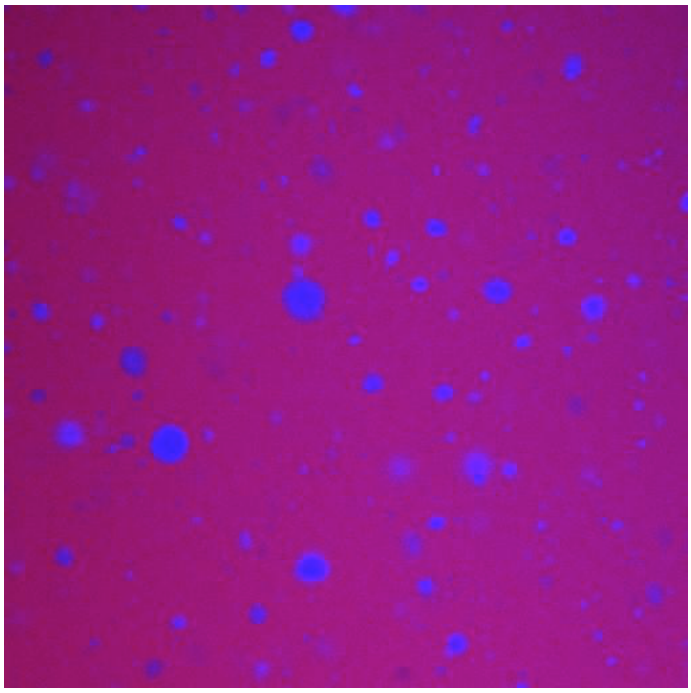
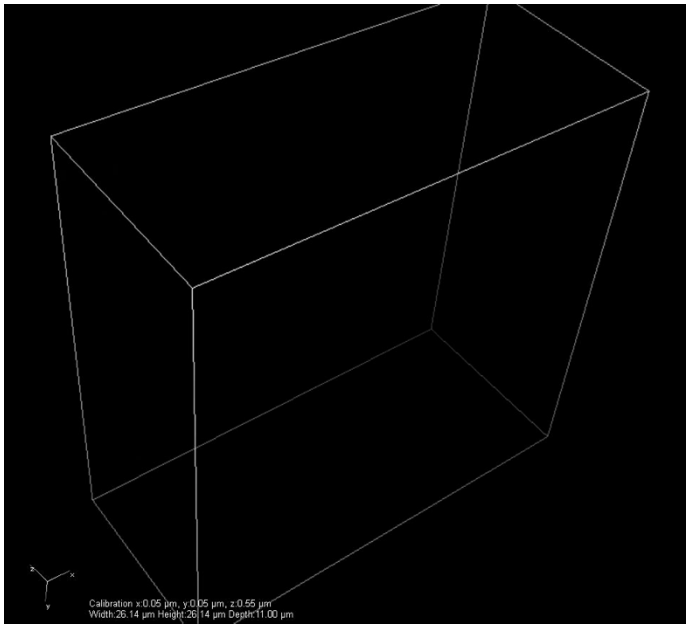


Figure S4.6. The release and retention of molecules with permeabilities of 10^{-12} and 10^{-13} cm/s across the membranes as a function of time. (A) The intersection of black dashed line (y = 0.9) and the blue line (the molecules released in lumen of the protocell) indicates that 90% of the molecules with $P = 10^{-12}$ cm/s will enter the main lumen of the protocells within 11 hours. (B) The intersection of black dashed line (y = 0.1) and the blue line (the molecules released in lumen of the protocell) indicates that 10% of the molecules with $P = 10^{-13}$ cm/s will enter the main lumen of the protocells within 5 hours.

Supplementary movies



Movie S4.1. This real-time movie shows that the topological transformation (prolate–oblate–stomatocyte-sphere) upon osmotic shock and micelle addition. A prolate vesicle kept its shape for ~1 s and then transformed into a stomatocyte shape via an oblate shape rapidly (in approximately 2s). Subsequently, the stomatocyte vesicle deformed to a sphere via invagination. Each micrograph is 126 μm by 126 μm.



Movie S4.2. 3D reconstruction of vesicle-in-vesicle structure from sectional image slices along the z-axis.

References

Acevedo, O. L. and L. E. Orgel (1987). "Nonenzymatic Transcription of an Oligodeoxynucleotide-14 Residues Long." Journal of Molecular Biology **197**(2): 187-193.

Adamala, K. and J. W. Szostak (2013). "Nonenzymatic Template-Directed RNA Synthesis Inside Model Protocells." Science **342**(6162): 1098-1100.

Aponte, J. C., J. E. Elsila, D. P. Glavin, S. N. Milam, S. B. Charnley and J. P. Dworkin (2017). "Pathways to Meteoritic Glycine and Methylamine." Acs Earth and Space Chemistry **1**(1): 3-13.

Attwater, J., A. Wochner and P. Holliger (2013). "In-ice evolution of RNA polymerase ribozyme activity." Nature Chemistry **5**(12): 1011-1018.

Bada, J. L., C. Bigham and S. L. Miller (1994). "Impact Melting of Frozen Oceans on the Early Earth - Implications for the Origin of Life." Proceedings of the National Academy of Sciences of the United States of America **91**(4): 1248-1250.

Barton, P. G. and F. D. Gunstone (1975). "Hydrocarbon Chain Packing and Molecular-Motion in Phospholipid Bilayers Formed from Unsaturated Lecithins - Synthesis and Properties of 16 Positional Isomers of 1,2-Dioctadecenoyl-Sn-Glycero-3-Phosphorylcholine." Journal of Biological Chemistry **250**(12): 4470-4476.

Biron, J. P. and R. Pascal (2004). "Amino acid N-carboxyanhydrides: Activated peptide monomers behaving as phosphate-activating agents in aqueous solution." Journal of the American Chemical Society **126**(30): 9198-9199.

Blain, J. C. and J. W. Szostak (2014). "Progress Toward Synthetic Cells." Annual Review of Biochemistry, Vol 83 **83**: 615-640.

Blok, M. C., E. C. M. Vanderneutkok, L. L. M. Vandeenen and J. Degier (1975). "Effect of Chain-Length and Lipid Phase-Transitions on Selective Permeability Properties of Liposomes." Biochimica Et Biophysica Acta **406**(2): 187-196.

Bonfio, C., C. Caumes, C. D. Duffy, B. H. Patel, C. Percivalle, M. Tsanakopoulou and J. D. Sutherland (2019). "Length-Selective Synthesis of Acylglycerol-Phosphates through Energy-Dissipative Cycling." Journal of the American Chemical Society **141**(9): 3934-3939.

Breslow, R. and D. L. Huang (1991). "Effects of Metal-Ions, Including Mg2+ and Lanthanides, on the Cleavage of Ribonucleotides and Rna Model Compounds." Proceedings of the National Academy of Sciences of the United States of America **88**(10): 4080-4083.

Bruckner, R. J., S. S. Mansy, A. Ricardo, L. Mahadevan and J. W. Szostak (2009). "Flip-Flop-Induced Relaxation of Bending Energy: Implications for Membrane Remodeling." Biophysical Journal **97**(12): 3113-3122.

Budin, I., A. Debnath and J. W. Szostak (2012). "Concentration-Driven Growth of Model Protocell Membranes." Journal of the American Chemical Society **134**(51): 20812-20819.

Budin, I. and J. W. Szostak (2013). "Physical Models for Early Membrane Growth and Evolution." Biophysical Journal **104**(2): 32a-32a.

Carell, T., C. Brandmayr, A. Hienzsch, M. Muller, D. Pearson, V. Reiter, I. Thoma, P. Thumbs and M. Wagner (2012). "Structure and Function of Noncanonical Nucleobases." Angewandte Chemie-International Edition **51**(29): 7110-7131.

Cavaluzzi, M. J. and P. N. Borer (2004). "Revised UV extinction coefficients for nucleoside-5'-monophosphates and unpaired DNA and RNA." Nucleic Acids Res **32**(1): e13.

Chen, I. A., R. W. Roberts and J. W. Szostak (2004). "The emergence of competition between model protocells." Science **305**(5689): 1474-1476.

Chen, I. A., K. Salehi-Ashtiani and J. W. Szostak (2005). "RNA catalysis in model protocell vesicles." Journal of the American Chemical Society **127**(38): 13213-13219.

Chen, I. A. and J. W. Szostak (2004). "A kinetic study of the growth of fatty acid vesicles." Biophysical Journal **87**(2): 988-998.
Crick, F. H. C. (1968). "Origin of Genetic Code." Journal of Molecular Biology **38**(3): 367-&.
Damer, B. and D. Deamer (2020). "The Hot Spring Hypothesis for an Origin of Life." Astrobiology **20**(4): 429-452.
Decatur, W. A. and M. J. Fournier (2002). "rRNA modifications and ribosome function." Trends in Biochemical Sciences **27**(7): 344-351.
Ding, D. A., L. J. Zhou, C. Giurgiu and J. W. Szostak (2022). "Kinetic explanations for the sequence biases observed in the nonenzymatic copying of RNA templates." Nucleic Acids Research **50**(1): 35-45.
Domling, A. (2002). "Recent advances in isocyanide-based multicomponent chemistry." Curr Opin Chem Biol **6**(3): 306-313.
Duzdevich, D., C. E. Carr, D. Ding, S. J. Zhang, T. S. Walton and J. W. Szostak (2021). "Competition between bridged dinucleotides and activated mononucleotides determines the error frequency of nonenzymatic RNA primer extension." Nucleic Acids Research **49**(7): 3681-3691.
Dzieciol, A. J. and S. Mann (2012). "Designs for life: protocell models in the laboratory." Chemical Society Reviews **41**(1): 79-85.
Fahrenbach, A. C., C. Giurgiu, C. P. Tam, L. Li, Y. Hongo, M. Aono and J. W. Szostak (2017). "Common and Potentially Prebiotic Origin for Precursors of Nucleotide Synthesis and Activation." Journal of the American Chemical Society **139**(26): 8780-8783.
Fai, T. G., B. E. Griffith, Y. Mori and C. S. Peskin (2013). "Immersed Boundary Method for Variable Viscosity and Variable Density Problems Using Fast Constant-Coefficient Linear Solvers I: Numerical Method and Results." Siam Journal on Scientific Computing **35**(5): B1132-B1161.
Ferris, J. P. (2002). "Montmorillonite catalysis of 30-50 mer oligonucleotides: Laboratory demonstration of potential steps in the origin of the RNA world." Origins of Life and Evolution of the Biosphere **32**(4): 311-332.
Ferris, J. P. and W. J. Hagan (1984). "Hcn and Chemical Evolution - the Possible Role of Cyano Compounds in Prebiotic Synthesis." Tetrahedron **40**(7): 1093-1120.
Ferris, J. P., A. R. Hill, R. H. Liu and L. E. Orgel (1996). "Synthesis of long prebiotic oligomers on mineral surfaces." Nature **381**(6577): 59-61.
Ferris, J. P., R. A. Sanchez and L. E. Orgel (1968). "Studies in Prebiotic Synthesis .3. Synthesis of Pyrimidines from Cyanoacetylene and Cyanate." Journal of Molecular Biology **33**(3): 693-&.
Fialho, D. M., T. P. Roche and N. V. Hud (2020). "Prebiotic Syntheses of Noncanonical Nucleosides and Nucleotides." Chemical Reviews **120**(11): 4806-4830.
Fox, G. E. (2010). "Origin and Evolution of the Ribosome." Cold Spring Harbor Perspectives in Biology **2**(9).
Fujikawa, S. M., I. A. Chen and J. W. Szostak (2005). "Shrink-wrap vesicles." Langmuir **21**(26): 12124-12129.
Gibb, E. L., D. C. B. Whittet, A. C. A. Boogert and A. G. G. M. Tielens (2004). "Interstellar ice: The Infrared Space Observatory legacy." Astrophysical Journal Supplement Series **151**(1): 35-73.
Gilbert, W. (1986). "Origin of Life - the Rna World." Nature **319**(6055): 618-618.
Govenar, B. (2012). "Energy Transfer Through Food Webs at Hydrothermal Vents: Linking the Lithosphere to the Biosphere." Oceanography **25**(1): 246-255.
Guerriertakada, C., K. Gardiner, T. Marsh, N. Pace and S. Altman (1983). "The Rna Moiety of Ribonuclease-P Is the Catalytic Subunit of the Enzyme." Cell **35**(3): 849-857.
Hanczyc, M. M. and J. W. Szostak (2004). "Replicating vesicles as models of primitive cell growth and division." Current Opinion in Chemical Biology **8**(6): 660-664.

Harada, K. and L. E. Orgel (1991). "The Cyclization of Arabinosyladenine-5'-Phosphorimidazolide." Journal of Molecular Evolution **32**(5): 358-359.
Hayyan, A., F. S. Mjalli, I. M. AlNashef, T. Al-Wahaibi, Y. M. Al-Wahaibi and M. A. Hashim (2012). "Fruit sugar-based deep eutectic solvents and their physical properties." Thermochimica Acta **541**: 70-75.
Horning, D. P., S. Bala, J. C. Chaput and G. F. Joyce (2019). "RNA-Catalyzed Polymerization of Deoxyribose, Threose, and Arabinose Nucleic Acids." ACS Synthetic Biology **8**(5): 955-961.
Inoue, T. and L. E. Orgel (1982). "Oligomerization of (Guanosine 5'-Phosphor)-2-Methylimidazolide on Poly(C) - an Rna-Polymerase Model." Journal of Molecular Biology **162**(1): 201-217.
James, K. D. and A. D. Ellington (1997). "Surprising fidelity of template-directed chemical ligation of oligonucleotides." Chemistry & Biology **4**(8): 595-605.
Johnston, W. K., P. J. Unrau, M. S. Lawrence, M. E. Glasner and D. P. Bartel (2001). "RNA-catalyzed RNA polymerization: Accurate and general RNA-templated primer extension." Science **292**(5520): 1319-1325.
Joyce, G. F. and L. E. Orgel (2006). "Progress toward Understanding the Origin of the RNA World." Rna World, Third Edition **43**: 23-56.
Joyce, G. F., A. W. Schwartz, S. L. Miller and L. E. Orgel (1987). "The Case for an Ancestral Genetic System Involving Simple Analogs of the Nucleotides." Proceedings of the National Academy of Sciences of the United States of America **84**(13): 4398-4402.
Joyce, G. F. and J. W. Szostak (2018). "Protocells and RNA Self-Replication." Cold Spring Harbor Perspectives in Biology **10**(9).
Kanavarioti, A., P. A. Monnard and D. W. Deamer (2001). "Eutectic Phases in Ice Facilitate Nonenzymatic Nucleic Acid Synthesis." Astrobiology **1**(3): 271-281.
Kasting, J. F. (2010). "EARLY EARTH Faint young Sun redux." Nature **464**(7289): 687-689.
Kim, S. C., L. Zhou, W. Zhang, D. K. O'Flaherty, V. Rondo-Brovetto and J. W. Szostak (2020). "A Model for the Emergence of RNA from a Prebiotically Plausible Mixture of Ribonucleotides, Arabinonucleotides, and 2 '-Deoxynucleotides." Journal of the American Chemical Society **142**(5): 2317-2326.
Kindt, J. T., J. W. Szostak and A. N. Wang (2020). "Bulk Self-Assembly of Giant, Unilamellar Vesicles." Acs Nano **14**(11): 14627-14634.
Kruger, K., P. J. Grabowski, A. J. Zaug, J. Sands, D. E. Gottschling and T. R. Cech (1982). "Self-Splicing Rna - Auto-Excision and Auto-Cyclization of the Ribosomal-Rna Intervening Sequence of Tetrahymena." Cell **31**(1): 147-157.
Lawless, J. G. and G. U. Yuen (1979). "Quantification of Monocarboxylic Acids in the Murchison Carbonaceous Meteorite." Nature **282**(5737): 396-398.
Leman, L., L. Orgel and M. R. Ghadiri (2004). "Amino acid condensation mediated by carbonyl sulfide." Abstracts of Papers of the American Chemical Society **228**: U693-U693.
Leman, L., L. Orgel and M. R. Ghadiri (2004). "Carbonyl sulfide-mediated prebiotic formation of peptides." Science **306**(5694): 283-286.
Li, L., N. Prywes, C. P. Tam, D. K. O'Flaherty, V. S. Lelyveld, E. C. Izgu, A. Pal and J. W. Szostak (2017). "Enhanced Nonenzymatic RNA Copying with 2-Aminoimidazole Activated Nucleotides." Journal of the American Chemical Society **139**(5): 1810-1813.
Lilley, D. M. J. and F. Eckstein (2007). Ribozymes and RNA catalysis. Cambridge, UK, The Royal Society of Chemistry.
Lipowsky, R. (1999). From membranes to membrane machines, Berlin, Heidelberg, Springer Berlin Heidelberg.

Lipowsky, R. (2014). "Remodeling of membrane compartments: some consequences of membrane fluidity." Biological Chemistry **395**(3): 253-274.
Lunine, J. I. (2006). "Physical conditions on the early Earth." Philosophical Transactions of the Royal Society B-Biological Sciences **361**(1474): 1721-1731.
Mansy, S. S. (2010). "Membrane Transport in Primitive Cells." Cold Spring Harbor Perspectives in Biology **2**(8).
Mariani, A., D. A. Russell, T. Javelle and J. D. Sutherland (2018). "A Light-Releasable Potentially Prebiotic Nucleotide Activating Agent." Journal of the American Chemical Society **140**(28): 8657-8661.
Markvoort, A. J., P. Spijker, A. F. Smeijers, K. Pieterse, R. A. van Santen and P. A. J. Hilbers (2009). "Vesicle Deformation by Draining: Geometrical and Topological Shape Changes." Journal of Physical Chemistry B **113**(25): 8731-8737.
McCollom, T. M., G. Ritter and B. R. T. Simoneit (1999). "Lipid synthesis under hydrothermal conditions by Fischer-Tropsch-type reactions." Origins of Life and Evolution of the Biosphere **29**(2): 153-166.
Menor-Salvan, C. and M. R. Marin-Yaseli (2012). "Prebiotic chemistry in eutectic solutions at the water-ice matrix." Chemical Society Reviews **41**(16): 5404-5415.
Monnard, P. A., D. W. Deamer and J. W. Szostak (2004). "The eutectic phase in water ice: A microenvironment conducive to nonenzymatic biopolymer formation." Proceedings of the Iii European Workshop on Exo-Astrobiology **545**: 145-148.
Monnard, P. A., A. Kanavarioti and D. W. Deamer (2003). "Eutectic phase polymerization of activated ribonucleotide mixtures yields quasi-equimolar incorporation of purine and pyrimidine nucleobases." J Am Chem Soc **125**(45): 13734-13740.
Muller, F., L. Escobar, F. Xu, E. Wegrzyn, M. Nainyte, T. Amatov, C. Y. Chan, A. Pichler and T. Carell (2022). "A prebiotically plausible scenario of an RNA-peptide world." Nature **605**(7909): 279-+.
O'Flaherty, D. K., N. P. Kamat, F. N. Mirza, L. Li, N. Prywes and J. W. Szostak (2018). "Copying of Mixed-Sequence RNA Templates inside Model Protocells." Journal of the American Chemical Society **140**(15): 5171-5178.
Ogilvie, J. W. and S. C. Whitaker (1976). "Effect of Tris on Reactions Catalyzed by Homoserine Dehydrogenase and Glyceraldehyde-3-Phosphate Dehydrogenase." Biochimica Et Biophysica Acta **445**(3): 525-536.
Orgel, L. E. (1968). "Evolution of Genetic Apparatus." Journal of Molecular Biology **38**(3): 381-&.
Orgel, L. E. (1992). "Molecular Replication." Nature **358**(6383): 203-209.
Orgel, L. E. (2004). "Prebiotic chemistry and the origin of the RNA world." Critical Reviews in Biochemistry and Molecular Biology **39**(2): 99-123.
Oro, J., B. Basile, S. Cortes, C. Shen and T. Yamrom (1984). "The Prebiotic Synthesis and Catalytic Role of Imidazoles and Other Condensing Agents." Origins of Life and Evolution of the Biosphere **14**(1-4): 237-242.
Patel, B. H., C. Percivalle, D. J. Ritson, C. D. Duffy and J. D. Sutherland (2015). "Common origins of RNA, protein and lipid precursors in a cyanosulfidic protometabolism." Nature Chemistry **7**(4): 301-307.
Peterlin, P., V. Arrigler, K. Kogej, S. Svetina and P. Walde (2009). "Growth and shape transformations of giant phospholipid vesicles upon interaction with an aqueous oleic acid suspension." Chemistry and Physics of Lipids **159**(2): 67-76.
Phleger, C. F., M. M. Nelson, A. K. Groce, S. C. Cary, K. Coyne, J. A. E. Gibson and P. D. Nichols (2005). "Lipid biomarkers of deep-sea hydrothermal vent polychaetes - Alvinella pompejana, A-

caudata, Paralvinella grasslei and Hesiolyra bergii." Deep-Sea Research Part I-Oceanographic Research Papers **52**(12): 2333-2352.
Phleger, C. F., M. M. Nelson, A. K. Groce, S. C. Cary, K. J. Coyne and P. D. Nichols (2005). "Lipid composition of deep-sea hydrothermal vent tubeworm Riftia packyptila, crabs Munidopsis subsquatnosa and Bythograea thermydron, mussels Bathymodiolus sp and limpets Lepetodrilus spp." Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology **141**(2): 196-210.
Phleger, C. F., P. D. Nichols and P. Virtue (1997). "The lipid, fatty acid and fatty alcohol composition of the myctophid fish Electrona antarctica: high level of wax esters and food-chain implications." Antarctic Science **9**(3): 258-265.
Powner, M. W., B. Gerland and J. D. Sutherland (2009). "Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions." Nature **459**(7244): 239-242.
Prywes, N., J. C. Blain, F. Del Frate and J. W. Szostak (2016). "Nonenzymatic copying of RNA templates containing all four letters is catalyzed by activated oligonucleotides." Elife **5**.
R. E. Schuster, J. E. S., and Joseph Casanova, Jr. (1966). Organic Syntheses **46**: 75.
Ritson, D. and J. D. Sutherland (2012). "Prebiotic synthesis of simple sugars by photoredox systems chemistry." Nat Chem **4**(11): 895-899.
Ritson, D. J. and J. D. Sutherland (2013). "Synthesis of Aldehydic Ribonucleotide and Amino Acid Precursors by Photoredox Chemistry." Angewandte Chemie-International Edition **52**(22): 5845-5847.
Rohatgi, R., D. P. Bartel and J. W. Szostak (1996). "Nonenzymatic, template-directed ligation of oligoribonucleotides is highly regioselective for the formation of 3'-5' phosphodiester bonds." Journal of the American Chemical Society **118**(14): 3340-3344.
Ruiz-Herrero, T., T. G. Fai and L. Mahadevan (2019). "Dynamics of Growth and Form in Prebiotic Vesicles." Physical Review Letters **123**(3).
Sacerdote, M. G. and J. W. Szostak (2005). "Semipermeable lipid bilayers exhibit diastereoselectivity favoring ribose." Proceedings of the National Academy of Sciences of the United States of America **102**(17): 6004-6008.
Saminathan, A., J. Devany, A. T. Veetil, B. Suresh, K. S. Pillai, M. Schwake and Y. Krishnan (2021). "A DNA-based voltmeter for organelles." Nature Nanotechnology **16**(1): 96-103.
Schrum, J. P., T. F. Zhu and J. W. Szostak (2010). "The Origins of Cellular Life." Cold Spring Harbor Perspectives in Biology **2**(9).
Seifert, U., K. Berndl and R. Lipowsky (1991). "Shape Transformations of Vesicles - Phase-Diagram for Spontaneous-Curvature and Bilayer-Coupling Models." Physical Review A **44**(2): 1182-1202.
Sole, R. V. (2009). "Evolution and self-assembly of protocells." International Journal of Biochemistry & Cell Biology **41**(2): 274-284.
Song, E. Y., E. I. Jimenez, H. Lin, K. Le Vay, R. Krishnamurthy and H. Mutschler (2021). "Prebiotically Plausible RNA Activation Compatible with Ribozyme-Catalyzed Ligation." Angewandte Chemie-International Edition **60**(6): 2952-2957.
Steitz, T. A. and P. B. Moore (2003). "RNA, the first macromolecular catalyst: the ribosome is a ribozyme." Trends in Biochemical Sciences **28**(8): 411-418.
Sulston, J., R. Lohrmann, L. E. Orgel and H. T. Miles (1968). "Nonenzymatic Synthesis of Oligoadenylates on a Polyuridylic Acid Template." Proceedings of the National Academy of Sciences of the United States of America **59**(3): 726-&.
Sutherland, J. D., L. B. Mullen and F. F. Buchet (2008). "Potentially prebiotic Passerini-type reactions of phosphates." Synlett(14): 2161-2163.
Szostak, J. W. (2011). "An optimal degree of physical and chemical heterogeneity for the origin of life?" Philosophical Transactions of the Royal Society B-Biological Sciences **366**(1580): 2894-2901.

Szostak, J. W. (2012). "The eightfold path to non-enzymatic RNA replication." Journal of Systems Chemistry **3**(1): 2.
Szostak, J. W. (2017). "The Narrow Road to the Deep Past: In Search of the Chemistry of the Origin of Life." Angewandte Chemie-International Edition **56**(37): 11037-11043.
Szostak, J. W., D. P. Bartel and P. L. Luisi (2001). "Synthesizing life." Nature **409**(6818): 387-390.
Tjhung, K. F., M. N. Shokhirev, D. P. Horning and G. F. Joyce (2020). "An RNA polymerase ribozyme that synthesizes its own ancestor." Proceedings of the National Academy of Sciences of the United States of America **117**(6): 2906-2913.
Todd, Z. R., R. Szabla, J. W. Szostak and D. D. Sasselov (2019). "UV photostability of three 2-aminoazoles with key roles in prebiotic chemistry on the early earth." Chem Commun (Camb) **55**(70): 10388-10391.
Toparlak, D., M. Karki, V. Egas Ortuno, R. Krishnamurthy and S. Mansy (2019). Cyclophospholipids Increase Protocellular Stability to Metal Ions, ChemRxiv.
Vazquez-Salazar, A., G. Tan, A. Stockton, R. Fani, A. Becerra and A. Lazcano (2017). "Can an Imidazole Be Formed from an Alanyl-Seryl-Glycine Tripeptide under Possible Prebiotic Conditions?" Origins of Life and Evolution of Biospheres **47**(3): 345-354.
Vogel, S. R., C. Deck and C. Richert (2005). "Accelerating chemical replication steps of RNA involving activated ribonucleotides and downstream-binding elements." Chemical Communications(39): 4922-4924.
Wachowius, F., J. Attwater and P. Holliger (2017). "Nucleic acids: function and potential for abiogenesis." Quarterly Reviews of Biophysics **50**.
Walde, P., A. Goto, P. A. Monnard, M. Wessicken and P. L. Luisi (1994). "Oparins Reactions Revisited - Enzymatic-Synthesis of Poly(Adenylic Acid) in Micelles and Self-Reproducing Vesicles." Journal of the American Chemical Society **116**(17): 7541-7547.
Walker, S. I., N. Packard and G. D. Cody (2017). "Re-conceptualizing the origins of life." Philosophical Transactions of the Royal Society a-Mathematical Physical and Engineering Sciences **375**(2109).
Walton, T., S. DasGupta, D. Duzdevich, S. S. Oh and J. W. Szostak (2021). "In vitro selection of ribozyme ligases that use prebiotically plausible 2-aminoimidazole-activated substrates (vol 117, pg 5741, 2020)." Proceedings of the National Academy of Sciences of the United States of America **118**(41).
Walton, T., L. Pazienza and J. W. Szostak (2019). "Template-Directed Catalysis of a Multistep Reaction Pathway for Nonenzymatic RNA Primer Extension." Biochemistry **58**(6): 755-762.
Walton, T. and J. W. Szostak (2016). "A Highly Reactive Imidazolium-Bridged Dinucleotide Intermediate in Nonenzymatic RNA Primer Extension." Journal of the American Chemical Society **138**(36): 11996-12002.
Walton, T. and J. W. Szostak (2017). "A Kinetic Model of Nonenzymatic RNA Polymerization by Cytidine-5 '-phosphoro-2-aminoimidazolide." Biochemistry **56**(43): 5739-5747.
Walton, T., W. Zhang, L. Li, C. P. Tam and J. W. Szostak (2019). "The Mechanism of Nonenzymatic Template Copying with Imidazole-Activated Nucleotides." Angew Chem Int Ed Engl **58**(32): 10812-10819.
Weimann, B. J., R. Lohrmann, L. E. Orgel, Schneide.H and J. E. Sulston (1968). "Template-Directed Synthesis with Adenosine-5'-Phosphorimidazolide." Science **161**(3839): 387-&.
Westheimer, F. H. (1987). "Why Nature Chose Phosphates." Science **235**(4793): 1173-1178.
Whitaker, D. and M. W. Powner (2018). "Prebiotic nucleic acids need space to grow." Nature Communications **9**.
Woese, C. R. and G. E. Fox (1977). "Concept of Cellular Evolution." Journal of Molecular Evolution **10**(1): 1-6.

Wordsworth, R., A. H. Knoll, J. Hurowitz, M. Baum, B. L. Ehlmann, J. W. Head and K. Steakley (2021). "A coupled model of episodic warming, oxidation and geochemical transitions on early Mars." Nature Geoscience **14**(3): 127-+.
Worth, R. J., S. Sigurdsson and C. H. House (2013). "Seeding Life on the Moons of the Outer Planets via Lithopanspermia." Astrobiology **13**(12): 1155-1165.
Wu, C. H., T. G. Fai, P. J. Atzberger and C. S. Peskin (2015). "Simulation of Osmotic Swelling by the Stochastic Immersed Boundary Method." Siam Journal on Scientific Computing **37**(4): B660-B688.
Wu, T. and L. E. Orgel (1992). "Nonenzymatic Template-Directed Synthesis on Hairpin Oligonucleotides .3. Incorporation of Adenosine and Uridine Residues." Journal of the American Chemical Society **114**(21): 7963-7969.
Xu, J. F., D. J. Ritson, S. Ranjan, Z. R. Todd, D. D. Sasselov and J. D. Sutherland (2018). "Photochemical reductive homologation of hydrogen cyanide using sulfite and ferrocyanide." Chemical Communications **54**(44): 5566-5569.
Yadav, M., R. Kumar and R. Krishnamurthy (2020). "Chemistry of Abiotic Nucleotide Synthesis." Chemical Reviews **120**(11): 4766-4805.
Yuen, G. U., J. G. Lawless and E. H. Edelson (1981). "Quantification of Monocarboxylic Acids from a Spark Discharge Synthesis." Journal of Molecular Evolution **17**(1): 43-47.
Zhang, S. J., D. Duzdevich, D. Ding and J. W. Szostak (2022). "Freeze-thaw cycles enable a prebiotically plausible and continuous pathway from nucleotide activation to nonenzymatic RNA copying." Proceedings of the National Academy of Sciences of the United States of America **119**(17).
Zhang, S. J., D. Duzdevich and J. W. Szostak (2020). "Potentially Prebiotic Activation Chemistry Compatible with Nonenzymatic RNA Copying." Journal of the American Chemical Society **142**(35): 14810-14813.
Zhang, W., S. C. Kim, C. P. Tam, V. S. Lelyveld, S. Bala, J. C. Chaput and J. W. Szostak (2021). "Structural interpretation of the effects of threo-nucleotides on nonenzymatic template-directed polymerization." Nucleic Acids Research **49**(2).
Zhang, W., T. Walton, L. Li and J. W. Szostak (2018). "Crystallographic observation of nonenzymatic RNA primer extension." Elife **7**.
Zhu, T. F. and J. W. Szostak (2009). "Coupled Growth and Division of Model Protocell Membranes." Journal of the American Chemical Society **131**(15): 5705-5713.
Zong, W., S. H. Ma, X. N. Zhang, X. J. Wang, Q. C. Li and X. J. Han (2017). "A Fissionable Artificial Eukaryote-like Cell Model." Journal of the American Chemical Society **139**(29): 9955-9960.