iBiology.org Teaching Tools
Jack Szostak's Lecture Part 3:
Non-enzymatic Copying of Nucleic Acid Templates

Teaching Tools were prepared by Kelsey Hass and Jack Szostak.

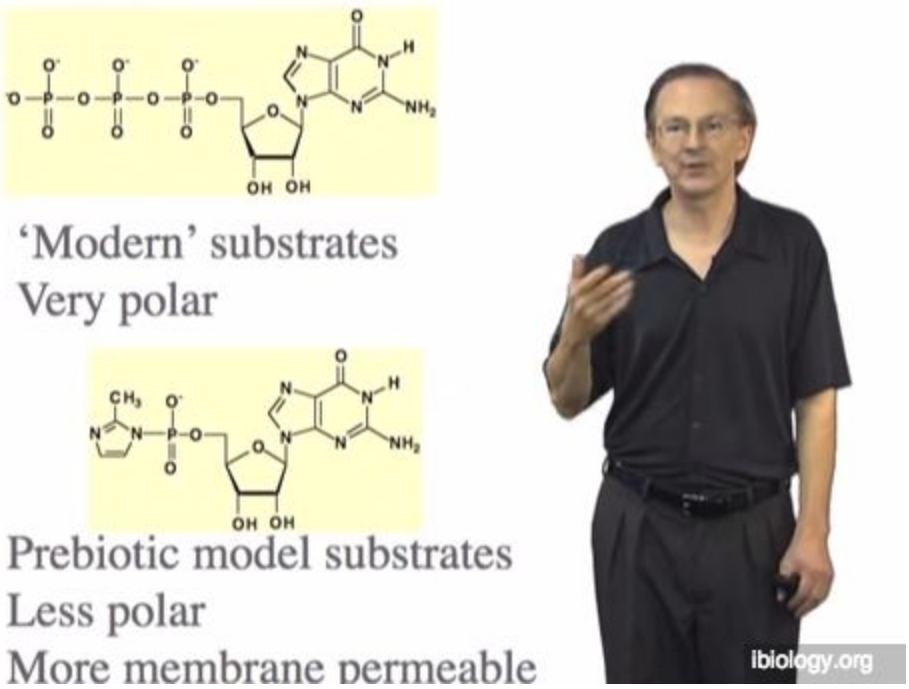
Contents

---

1. Keywords and Terms

RNA world hypothesis, prebiotic chemistry, protocell, imidazole nucleobase, self-polymerization

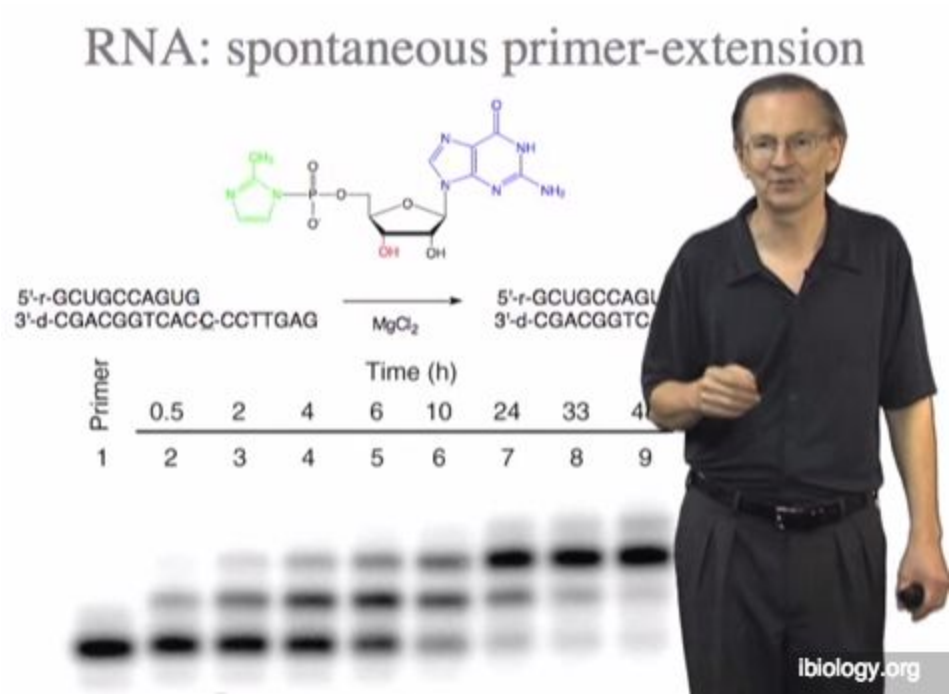
2. Lecture Notes



Time: 4:35

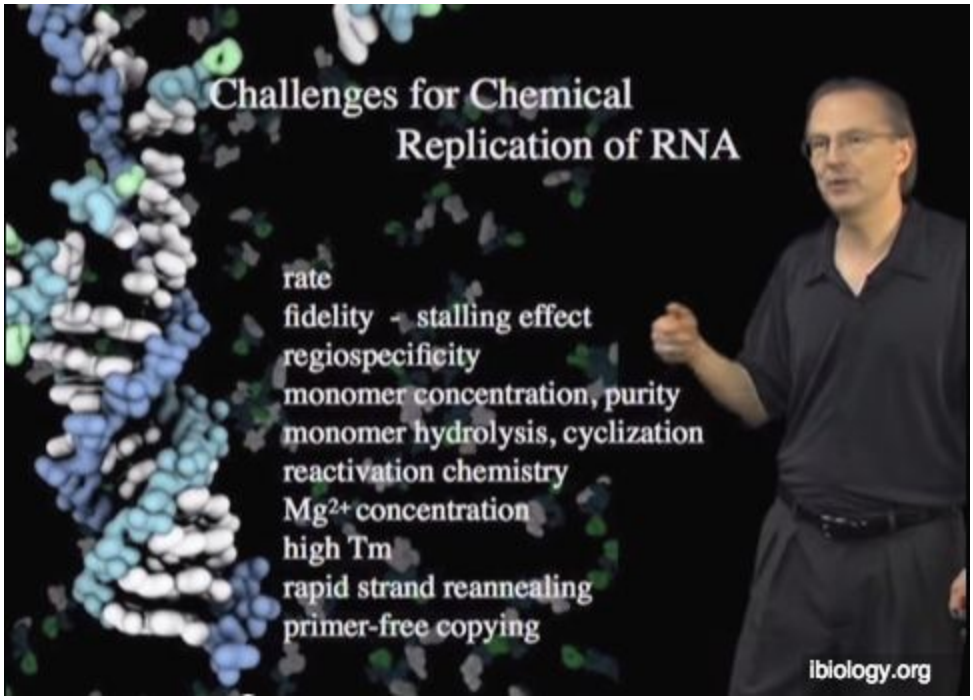
In this lecture, Jack Szostak describes characteristics of prebiotic nucleic acids, and experiments that support models for non-enzymatic replication.

Prebiotic nucleotide triphosphates have different chemical and structural properties compared to modern nucleotide triphosphates.



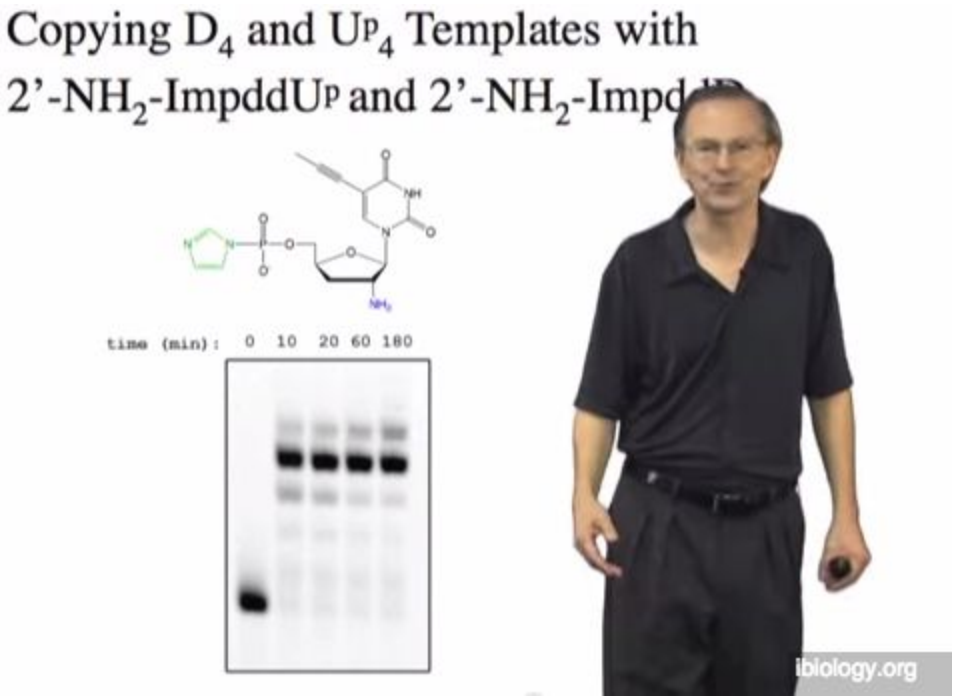
Time: 13:59

Imidazole derivatives of nucleotide triphosphates can undergo spontaneous chemical replication of RNA templates. It is an intrinsically slow process; the experiment to the left shows that synthesizing two bases can take about 24 hours. An additional problem is that the rates of synthesis and degradation are on similar time scales when the reaction proceeds in high concentrations of magnesium ions.



Time: 16:12

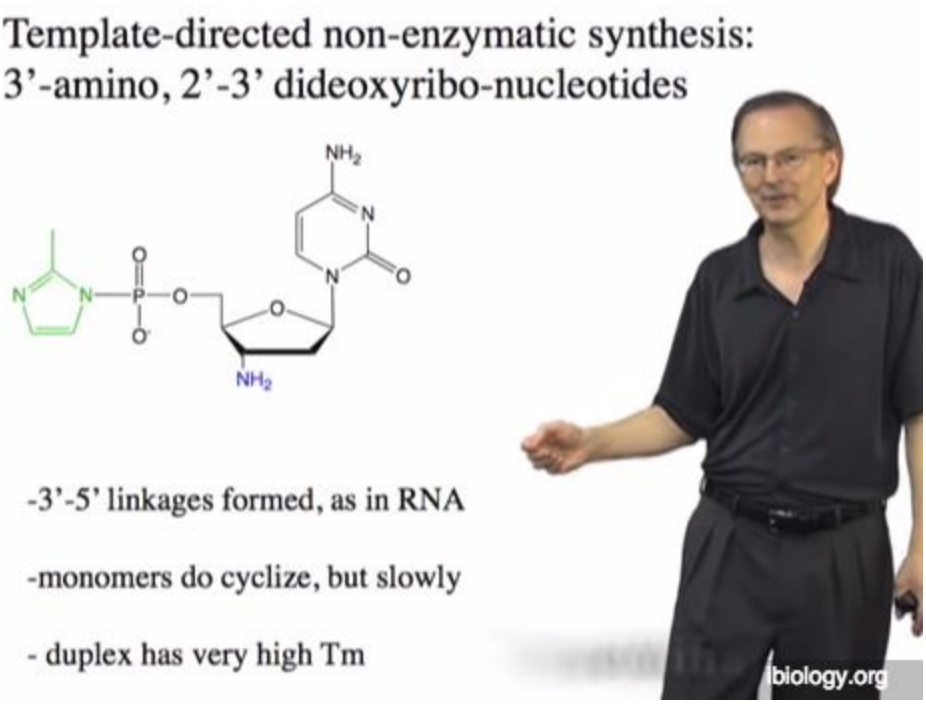
There are many additional challenges for chemical replication of RNA, summarized to the left.



Time: 29:22

To solve the fidelity problem, one area of research is investigating alternative nucleotide derivatives. There are many chemical additions or modifications that have been shown experimentally to increase the speed and/or fidelity of spontaneous RNA template copying.

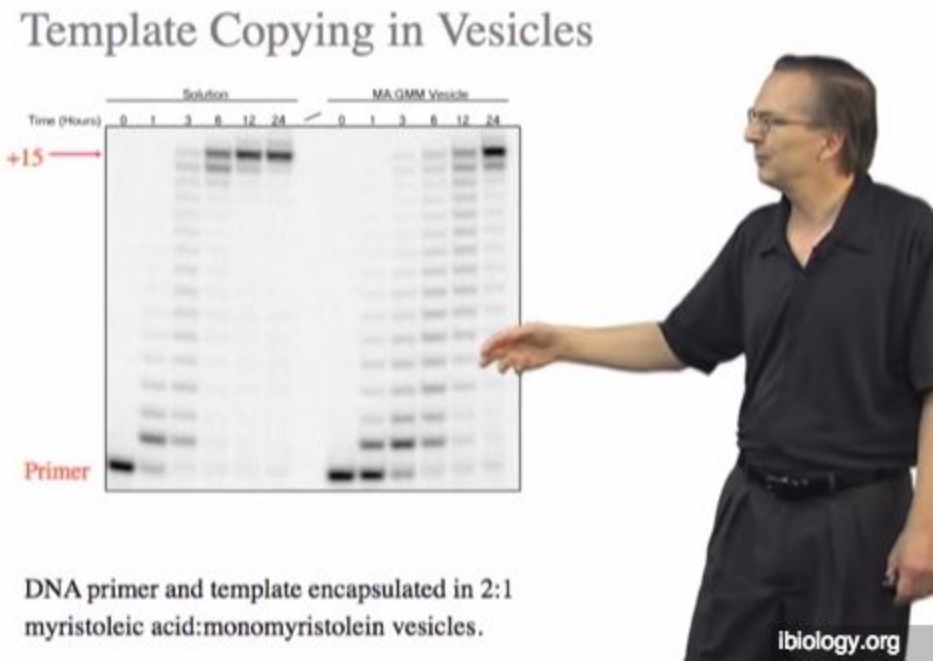
Here, amino-uracil derivatives are shown to rapidly polymerize RNA primers.



Time: 34:56

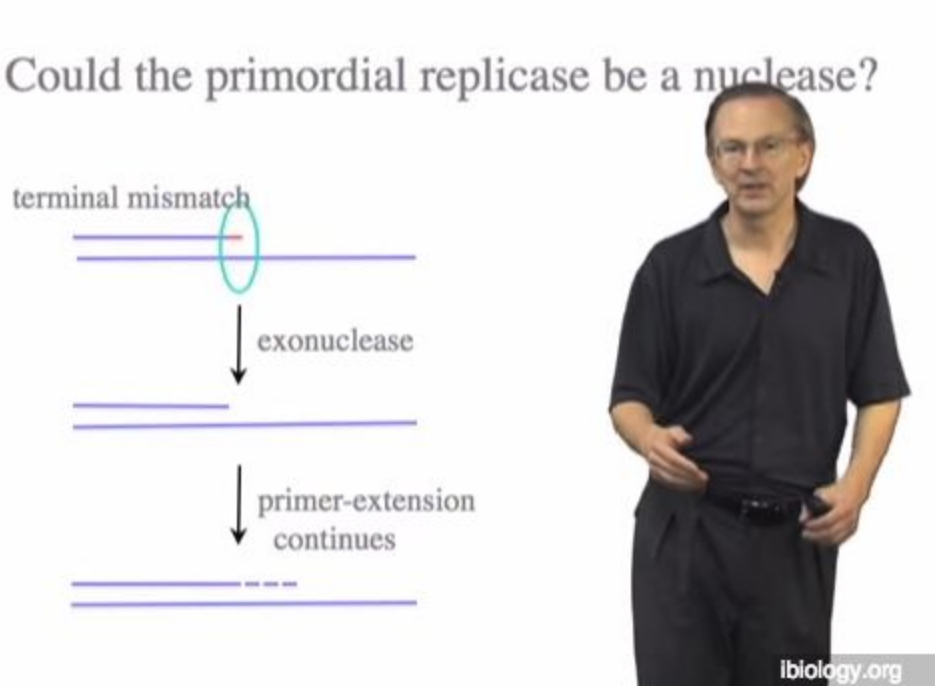
To solve the fidelity problem, one area of research is investigating alternative nucleotide derivatives. There are many chemical additions or modifications that have been shown experimentally to increase the speed and/or fidelity of spontaneous RNA template copying.

Here, imidazole-nucleotide derivatives are promising precursors for synthesizing growing RNA chains.



Time: 40:58

Prebiotic nucleotide precursors can diffuse across fatty acid membranes to form replicated RNA products, indicating the self-replication chemistry is compatible inside vesicles. One major area of research remaining is to combine the replication chemistry with growing/dividing vesicles to watch early life emerge.



Time: 50:20

Why could the primordial replicase be a nuclease? One hypothesis is that a nuclease could alleviate the fidelity stalling effect when an incorrect base is incorporated. If the primordial replicase was a nuclease, the incorrect base could be excised to maintain fidelity and the rate of synthesis.

3. Review Questions

1. Fill in the blanks: Prebiotic nucleotides are less _____ and more _____ than modern nucleotides.
2. List two plausible natural environments that could facilitate generating RNA chains.
3. What bond formation was considered a major roadblock for prebiotic pyrimidine synthesis?
4. What are two problems with the high magnesium concentrations needed for spontaneous RNA polymerization? Select all that apply.
    a. Magnesium is used for synthesis of RNA, but also for degradation of RNA.
    b. Magnesium ions were likely not found in prebiotic environments.

c. Magnesium ions need to be coordinated to another metal for RNA polymerization.
d. Magnesium can disrupt fatty acid vesicle membranes, causing fatty acid precipitation.

5. True or false: Given the right membrane permeability, prebiotic nucleotides could flow across a protocell membrane to be accessed for RNA self-replication.

6. What could be the primordial replicase?

    a. RNA polymerase
    b. DNA polymerase
    c. Nuclease
    d. Ribozyme

4. Answers to Review Questions

1. polar; membrane permeable

2. Polymerization of prebiotic nucleotides on clay and ice eutectic phase could facilitate generating RNA chains.

3. The glycosidic bond (or linkage) in pyrimidine nucleosides serves as a roadblock for pyrimidine synthesis; it cannot be formed from hypothesized prebiotic precursors. A different pathway using an alternative intermediate molecule was needed to generate pyrimidine nucleosides.

4. a and d

5. True

6. C

5. Discussion Questions

1. What is structurally and catalytically different about the substrates required for chemistry-driven RNA/DNA replication? How could they have been generated on early Earth?

2. Describe some of the challenges for chemical replication of protocell genetic material. What evidence strongly suggests chemical replication is possible?

3. Give examples of current research efforts to solve the fidelity problem with RNA replication.
4. What is the major takeaway about monomer (or nucleotide base) homogeneity?

6. Answers to Discussion Questions

1. The substrates for non-enzymatic chemistry-driven RNA/DNA replication are more reactive nucleophiles. This means the substrate contains a more reactive group (such as an imidazole group) that will attack the RNA/DNA backbone to spontaneously add another base to the growing RNA/DNA strand. In addition, these substrates are less polar and can more easily or rapidly diffuse across the protocell membrane.
    a. Hypotheses for generating nucleotides on early Earth include: self-assembly of ribose from five formaldehyde units; adenine from five cyanide units (pentamer of cyanide); and cytosine from cyanoacetyline and urea units. These simple prebiotic units can be spontaneously combined together into bigger macromolecular structures by UV light, intense heat/boiling, and other extreme environmental processes.
2. There are many challenges for chemical replication of protocell genetic material (students may have a variety of answers, including):
    a. No direct pathway established (yet) for purine synthesis (still an active topic of research)
    b. Unknown how to generate pure, concentrated pools of starting nucleotide monomers from spontaneous combinations of precursor units
    c. In high concentrations of magnesium ions, the rate of RNA synthesis and rate of RNA degradation are on the same time scale (synthesis possible, but not sustainable for the cell to replicate its genetic material)
    d. Slow synthesis rate
    e. Stalling effect for polymerizing a correct base after an incorrect base is added
    f. Lack of regiospecificity (2’ or 3’ OH linkages can form spontaneously)
    g. Activated monomers are unstable; monomers can hydrolyze or cyclize, and no longer be useful for long chain generation
    h. High melting temperature of RNA template and copied strand duplex; strands are hard to dissociate for further template copying
    i. Rapid RNA strand re-annealing (same problem as above)

j. GC base pairing works well in current research experiments, but AU base pairing is much slower or does not work at all
k. Despite these challenges, there are many example long-chain molecules in the lecture, such as GNA, TNA, and MoNA, that suggest chemical replication is possible. At the most basic level, there is a plausible pathway for pyrimidine synthesis from spontaneous combinations of prebiotic units. Simple precursors can combine to generate an intermediate that can undergo further spontaneous manipulations to become cytosine, or deaminated to become uracil. This is promising and suggests nucleotides and longer chain nucleic acids could be generated spontaneously. Polyacrylamide gel electrophoresis results showed that longer chains of RNA could be spontaneously synthesized over time from variations of different starting nucleotides (all with different nucleophile groups or slightly different chemical structures and reactivity). The prominent example in this lecture is 2'-amino, 2',3'-dideoxyribonucleotides (2'-NH2 ImpdG).

3. Examples of current research efforts include (students may have a variety of answers):

    a. Using another base analog, 2-thiol-uracil, that has a more reactive sulfur group and promotes tighter (higher fidelity) adenine/uracil base pairing
    b. Synthesizing conformationally constrained RNA long-chain backbone structures (ex 3'-phosphoramidate DNA; threose nucleic acid or TNA; morpholino nucleic acid or MoNA) that may favor the incorporation of the right base simply by being constrained to favor one conformation
    c. Investigating the possibility that ribozymes, small molecules or peptides may catalyze long-chain formation (i.e. using a catalyst that favors specific conformations of monomers and backbone products, instead of spontaneous non-enzymatic chemistry)

4. Monomer homogeneity is specific to the lab setting; the simplest experiments to investigate scientific questions use one type of pure monomer at high concentrations. However, in "real prebiotic life", there would be a heterogeneous mixture of monomers created from all possible random, spontaneous combinations of prebiotic precursor units. Therefore, a mixture of DNA or RNA bonds could form within the same molecule from combinations of these heterogeneous monomers. In fact, as shown in the lecture, a mixture of 2'-5' and 3'-5' linkages within the same molecule reduced the melting temperature for RNA duplexes, and may have been an essential characteristic to promote strand separation for RNA self-replication.

7. Explain or Teach These Concepts to a Friend

1. Teach a friend why modern nucleotide triphosphates and phospholipid membranes are not suitable candidates for nucleic acid self-replication chemistry in a primitive protocell. Describe what properties of nucleotides and lipids would have to change in order for these molecules to be suitable for a prebiotic protocell.
2. Tackle one of the challenges for chemical replication of protocell genetic material mentioned in the lecture. What experiments would you perform to investigate this challenge?

   Note: This may require students to do a bit of literature research. One advised starting point: look up the articles cited in this lecture on PubMed, and read other relevant research or review articles. Students should be encouraged to choose questions/experiments not yet conducted in published research articles.

8. Paper for Journal Club

Discussion Paper

Adamala, K. and Szostak, J.W. “Nonenzymatic Template-Directed RNA Synthesis Inside Model Protocells.” 2013. Science. 342: 1098-1100.

Discussion Questions

Note: Students are not limited to answers provided below; answers serve as one example for the discussion section. Students are encouraged to pursue other avenues of discussion for each question.

1. What was the “roadblock” in non-enzymatic RNA replication chemistry that the authors addressed?

The authors’ model for prebiotic cells has two major components: fatty acid membranes and RNA molecules that self-replicate. The chemistry to accomplish RNA copying requires a high concentration of magnesium ions, among other divalent (2+ charged) ions. However, in high concentrations of magnesium, the fatty acids become

water-insoluble and aggregate into clumps (precipitate out). This destroys spontaneous arrangement of fatty acids into vesicles to form the prebiotic cell. The major roadblock addressed in this paper is resolving the incompatibility between fatty acid integrity as vesicles and successful RNA copying in high concentrations of magnesium.

Other avenues of discussion:

Why do you think magnesium ions are required for RNA replication chemistry?

[Magnesium ions are necessary for catalysis chemistry in linking free-floating nucleotide triphosphates to the growing RNA strand backbone.]

If you were given this “roadblock”, what experiments would you do to find a solution?

2. Talk through Figure 2. What did they test, and how? Describe the main conclusion(s). What could be some biases in their primer-extension assay? What properties of chelators are beneficial for a primitive protocell?

In Figure 2, the authors tested the rate of RNA synthesis in the presence of Mg^{2+} and/or the chelators identified in Figure 1 that protected fatty acid membranes.

The rate of RNA synthesis was tested by primer-extension assay. Here, RNA duplexes were pre-formed such that two strands, a template and primer, were stably base-paired into a duplex at room temperature.

The template strand contained an additional stretch of 4 cytosine residues, and the corresponding strand (lacking the C residues) served as a primer for spontaneous base addition of primitive guanosine derivative, 2MeImpG. The primer strand also contains a Cy3-fluorophore at the 5’end, which allows the authors to visualize the length of the primer strand over time by gel analysis. At different time points, aliquots of the reaction were taken. RNA was isolated by ethanol precipitation, and analyzed by gel electrophoresis. Gel bands were visualized by fluorescence imaging. The results would be similar to the gel in Figure 2C, which shows the amount of primer extension (all bands except the bottom row) and unused primer template (bands in the bottom-most row of the gel) with 50mM Mg^{2+} and 200mM citrate over time.

Gel band intensities were then quantified by a specific analysis program. The quantitation is useful to determine total primer intensity (P_0 in Figure 2A; all bands per one lane in 2C) and unextended primer intensities (P in Figure 2A; bottom-most band per one lane in 2C). Graphing the ratio of unextended primer/total primer (y-axis Figure 2A) over time (x-axis Figure 2A) gives the rate of RNA synthesis (essentially, depletion

of unextended primers over time). The slope of the graph represents the rate of RNA synthesis. The slope of the line for each condition in Figure 2A is plotted in Figure 2B.

Looking at Figures 2A and 2B, we see the 50mM Mg^{2+}/no chelator condition gives the smallest amount of unextended primer, and the maximal rate of RNA synthesis at 1.4 $hour^{-1}$. However, this concentration of Mg^{2+} is incompatible with membrane vesicles (see question 1 above or authors' supplementary text). Of the two chelators tested with 50mM Mg^{2+}, citrate allows the highest rate of RNA synthesis at 0.67 $hour^{-1}$. The rate of RNA catalysis with Mg^{2+} and EDTA is close to 0 $hour^{-1}$, showing EDTA is too efficient of a chelator (it immobilizes too many Mg^{2+} ions thereby depleting the stock available for RNA catalysis, resulting in large amounts of unextended primer over time).

The main conclusion of Figure 2 is that the membrane-protective chelator, citrate, allows a reasonable rate of non-enzymatic RNA template copying in 50mM Mg^{2+}, both in solution and with oleate vesicles.

One potential bias of the primer-extension assay is the choice of nucleotide base. The authors measured synthesis rates with only one base, the guanosine derivative 2MeImpG. However, primitive adenine and uracil bases are known to have slower synthesis rates in non-enzymatic RNA synthesis chemistry. It is unknown if citrate could be an effective chelator for all bases (A, U, C and G), or if the timing would be too slow with other bases (A, U) for sufficient RNA synthesis in protocells.

In model protocells, chelators would perform two roles: (1) protection against fatty acid membrane dissolution by high concentrations of Mg^{2+} ions, and (2) protection against Mg^{2+}-induced RNA degradation.

Other potential answer: Chelators must prevent increases in the melting temperature (Tm) of a RNA duplex. If the melting temperature is high, the duplex is stable. Spontaneous "unwinding" of the duplex to allow single-stranded template copying will occur less frequently. Therefore, with high melting temperatures, the protocell would not be able to quickly replicate its RNA.

3. Talk through Figure 3. What did they test, and how? What is the "liposome dialyzer", and what is the method's significance for the paper (focus on 3E)? Any critiques for this figure?

In Figure 3, the authors measured the amounts of primer extension (RNA synthesis) in different membrane vesicle compositions over time.

RNA synthesis was tested by primer extension assay, with a template region of 7 C residues (Figure 3AD) or template region of mixed G and C residues (Figure 3E). Primer strands are complementary to template strands but lack the specified template regions, promoting non-enzymatic RNA synthesis off of the 3' end of the primer. The authors formed vesicles with different fatty acid compositions (Figure 3B-D), and each vesicle encapsulated template-primer duplex RNA, Mg^{2+} ions and citrate. Vesicles were

incubated with 2MeImpG nucleotide base (Figure 3A-D), or a mixture of prebiotic G- and C-derivative bases (Figure 3E). At indicated times, the authors collected aliquots of vesicles, extracted and purified the RNA, and analyzed primer RNA length by polyacrylamide gel electrophoresis. Primer extension was visualized by fluorescence imaging of the 5' Cy3 fluorophore signal on the primer. Bands in the bottom row of the gel correspond to unextended primer, and bands at increasing lengths represent one base extended (per interval) on the primer strand.

As shown in Figures 3A-D, most of the primer is extended by 24-72 hours after one-time addition of excess nucleotide base. Figure 3A serves as a control reaction (in solution, no vesicles), while Figures 3B-D show non-enzymatic RNA synthesis in three different model protocell membrane compositions. 3D shows the most prebiotically plausible membrane composition, as it consists of shorter chain lipids that are naturally permeable to reagents like prebiotic nucleotide triphosphates.

Figure 3E analyzes two important conditions: the efficiency of synthesizing across a mixed G/C-base template region, and the efficiency of RNA catalysis with fresh monomer addition. Fresh monomers are added by incubating vesicles in a liposome dialyzer over the course of the reaction. In the lipid dialyzer, vesicles are constantly exposed to fresh sources of the prebiotic nucleotide triphosphates. This is significant because the lipid dialyzer more accurately mimics the environmental conditions that prebiotic protocells would have experienced. With fresh monomers, primer extension occurs as efficiently for vesicle-encapsulated RNA primers (lanes 1 and 2, Figure 3E) as for RNA primers in solution (no vesicles, lane 3, Figure 3E). Addition of fresh nucleotide bases results in a more intense fully-extended primer band (lane 3, Figure 3E) than one-time addition of nucleotide bases (lane 4, Figure 3E), implying fresh addition of nucleotide base can increase the amount of extended primer at a given time point.

One critique for this figure is that "mixed bases" only refers to G/C-base template. There is still no data offered on A and U base synthesis.

Other avenues of discussion: Is citrate a viable prebiotic molecule? Would it have a more likely prebiotic form (for example, the “prebiotic form” of guanosine nucleotide triphosphate is 2MeImpG), and would this prebiotic form act differently?

4. Comment on the time scale for RNA replication vs. RNA degradation in a fatty acid vesicle containing Mg^{2+} and citrate. Does this seem plausible for early life?

RNA replication with 50mM Mg^{2+} and 200mM citrate in solution or in oleate vesicles was ~0.7 $hour^{-1}$, while RNA degradation with 50mM Mg^{2+} and 200mM citrate in solution was ~0.004 $hour^{-1}$. With these numbers, spontaneous RNA synthesis occurs ~175 times faster than degradation when 2MeImpG nucleotide base is in excess. The time scale is certainly plausible for a protocell to non-enzymatically replicate its RNA.

5. Recall the model for the protocell life cycle (Jack Szostak iBio seminars). To what extent does this paper piece together critical elements of the model? What aspects of the model remain to be tested?

The model for the protocell life cycle involves spontaneous protocell vesicle membrane growth/division and spontaneous, non-enzymatic replication of vesicle-encapsulated RNA. This paper focuses on the latter, in particular resolving the incompatibility between high Mg^{2+} concentrations required for RNA replication chemistry and Mg^{2+}-catalyzed destabilization of the vesicle membrane. The Szostak lab has previously shown that vesicles can spontaneously grow and divide through a multilamellar vesicle intermediate; this paper shows RNA molecules can self-replicate in fatty acid vesicles with Mg^{2+} and a chelator. What remains to be tested is combining the two facets together: combining the replication chemistry with conditions that promote self-replicating vesicles to mimic the complete protocell life cycle.