Executive Summary

In this proposal, I detail four presentations relating to my research on the chemical origins of life. The first, my planned lecture to the general public, will entail an overview of the history of the study of the chemical origins of life, then a high-level discussion of an aspect of my research and that of others in my lab of high accessibility and interest to the public: the evolution of the earliest cellular compartments. My planned colloquia relate to two aspects of my postdoctoral research, and one project I was involved in as a senior graduate student. In the first, I will discuss roles short oligonucleotides could have played in the emergence of life. In the second, I will discuss work relating to backbone heterogeneity in RNA in the origins of life. In the third, I will discuss a water-free solvent we discovered can support RNA and DNA folding, and its potential relevance to the emergence of life.

Public Lecture: From soap to life: How did the first cells form?

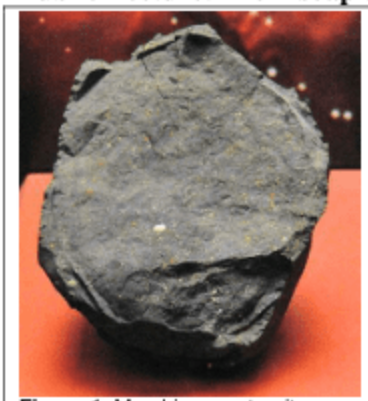

Figure 1. Murchison meteorite fragment held in National Museum of Natural History (Washington, DC). Image from Wikipedia user Basilofresco. License: CC BY-SA 3.0

Since scientists began discovering the molecular basis of life today, they have investigated how those molecules might have assembled before life existed. Researchers got some clues as to how this might have happened in 1952, when two researchers, Stanley Miller and Harold Urey, showed that the a remarkable number of molecules found in life today, including components of proteins, DNA, RNA, and sugars, could be made easily on earth. By using a sterile glass apparatus containing a mixture of gases thought to be representative of the early earth's atmosphere and subjecting it to heat and simulated lightning discharges, these investigators discovered that many of the same compounds found in life could be synthesized without the help of living things. Not long after, in 1969, we got our first indication that these molecules could have been made in outer space as well, when a 200 pound-plus meteorite crashed into a small town called Murchison in Victoria, Australia (Figure 1). Analysis of the meteorite showed that it was a special type called a carbonaceous chondrite, which contained organic compounds. This meteorite contained many of the same compounds found in life as those that Miller and Urey had prepared. Nonetheless, it was clear that these compounds did not come from living things. First, dating by observing how much naturally-occuring radioactive elements in the meteorite had decayed suggested that it dated back to the beginnings of the universe – well before life evolved. Second, the molecules present in the meteorite occurred in two forms – mirror images, like a right- and left-handed glove. When molecules that can have a mirror image occur in life, only one of these two forms typically is found. Since the Murchison meteorite contained many of the molecules found in life today, but in both mirror-image forms, this indicated that the molecules in the meteorite must have come from sources other than of living organisms. These discoveries prompted a wide range of studies as to how life might have first assembled from simple chemical precursors.

Today, researchers examining how life began have made great strides towards synthesizing "life in the lab" from simple chemical precursors. One aspect of our research towards this aim relates to the goal of preparing a "protocell" capable of primitive forms of metabolism, growth, and division. In this work, we use fatty acids – long chains of carbon and hydrogen atoms with oxygen and hydrogen atoms at their tails. These fatty acids are the same ones found in a bar of soap and as components of modern cell membranes. Fatty acids tend to fold up into compact structures when dissolved in water, with the long carbon-hydrogen tails facing each other and their other ends facing the water. Under some conditions, these compounds self-assemble into "vesicles," or cell-like structures, with a semipermeable membrane enclosing a compartment. In our research, we have found that these vesicles can exhibit remarkably cell-like behaviors, with simple physical processes inducing them to grow and divide (Figure 2).

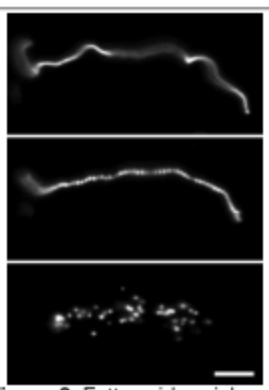

Figure 2. Fatty acid vesicles containing just a few simple chemicals, when irradiated with intense light, spontaneously "pearl" and divide. From Szostak et al. *PNAS* (2012) 9828-9832.

In this lecture, I will give a brief overview of the history of studies of the chemical origins of life. I also will discuss in detail one aspect of early life we study in our work: the development of a "protocell" from simple chemical precursors.

Colloquium #1: Low-efficiency nonenzymatic polymerization: a problem for the RNA world?

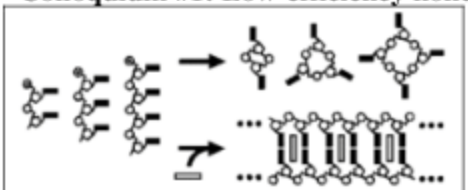

Figure 1. Short oligonucleotides (left), when not base-paired, cyclize (top right). When base pairing is induced by introduction of ethidium (grey rectangle), polymerization occurs (bottom right). From reference 1.

The discovery of catalytic RNAs by Cech and Altman in the 1980s has since led to a proliferation of interest in an "RNA world," wherein the dual catalytic and self-templating capabilities of RNA enabled it to play a role as the first biopolymer. Considerable progress has been made in the following 35 years, including the discovery of ribozyme polymerases, increasingly efficient nonenzymatic polymerization reactions, and, recently, the replication of RNA within a lipid vesicle – thought to be a precursor to extant modern cells. Nonetheless, efficient copying of nucleic acids without aid of the protein-based polymerases found in modern life remains an unsolved problem. Even those reactions catalyzed by the aforementioned ribozyme polymerases produce significant amounts of truncation products, and nonenzymatic copying reactions are still less efficient. In this colloquium, I will present two systems from our recent work that afford potential solutions to these problems.

In the first system, we have shown that short oligonucleotides (4 nt) capable of base pairing, but with very low affinity, due to their length (Figure 1), cyclize when treated with a condensing reagent (*N*-cyanoimidazole). However, when an intercalating small molecule is introduced (ethidium), the binding energy associated with ethidium-duplex interactions drives duplex formation (reference 1). When the oligonucleotide-ethidium complex is treated with the condensing agent, long polymers of >100 nt are formed. Furthermore, we show that Watson-Crick specificity is maintained, and that even the presence of a large excess (>10,000-fold) of a competing oligonucleotide of non-complementary sequence does not suppress polymerization. Thus, we have demonstrated that even short oligonucleotides, as expected to be produced in low-efficiency coping reactions, could have been elaborated into long polymers, given the correct condensing agent.

Figure 2. Very short oligonucleotides (first and second panel), when subjected to nonenzymatic primer extension conditions, can be elongated and become functional ribozymes and aptamers. From reference 2.

In the second system, we have shown a potential means of addressing the problem of inefficient copying of functional RNAs (reference 2). In this system, we employed truncated forms of ribozymes and aptamers, as one would expect to be produced in inefficient copying reactions. Due to their length and inability to form a critical structural stem, all these molecules are inactive as ribozymes and aptamers (Figure 2). In this system, we observed that even the nonfunctional complexes formed by these truncated strands are competent to undergo nonenzymatic primer extension. When primer extension is performed on these structures, full-length functional RNAs are reconstituted, and activity is restored. Thus, even a pool of nonfunctional partial copies of a functional RNA could have been reconstituted into a functional RNA, potentially *via* the same copying chemistry that generated the original RNA fragment. Such phenomena could have reduced the required efficiency for prebiotic nonenzymatic RNA polymerization systems considerably.

Bibliography: 1) Horowitz, EH; Engelhart, AE; Quarles, KA; Smith, MW; Chen, MC; Lynn, DG. *Proc. Nat. Acad. Sci. USA* (2010) 107:5288-5293. 2) Adamala, K; Engelhart, AE; Szostak, JW. *J. Am. Chem. Soc.* (2015) 137:483-489.

Colloquium #2: Backbone heterogeneity in the RNA world

A central "chicken and the egg" problem in putative RNA world scenarios, wherein RNA played both a catalytic and information inheritance role in the emergence of life, relates to the difficulty associated with obtaining the first catalytic RNAs. Prior to the emergence of polymerases, RNAs must necessarily have replicated by nonenzymatic processes. However, all known methods of nonenzymatic RNA replication result in a mixture of products containing the canonical 3′-5′ linkage, as well as the nonstandard 2′-5′ linkage (Figure 1). 2′-5′ linkages are known to have a number of deleterious effects on RNA, including diminution of double helical stability and diminished

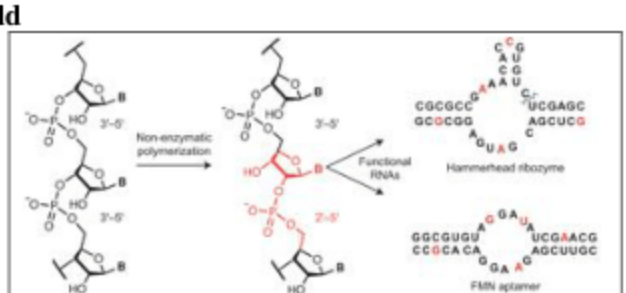

Figure 1. RNAs prepared by nonenzymatic polymerization contain a mixture of the canonical 3′-5′ linkages found in life (black) and nonstandard 2′-5′ linkages (red). Despite this, such RNAs can fold into functional structures, including an aptamer and ribozyme. From reference 1.

chemical stability, with helical 2′-5′ linkages being optimally positioned for transesterification, resulting in strand degradation. In this colloquium, I will present results from our recent work relating to 2′-5′ substitution in primitive RNAs that suggest it may not have been as deleterious to an RNA world as it might appear at first sight.

We have examined functional RNAs (the hammerhead ribozyme and a flavin mononucleotide aptamer obtained previously by *in vitro* selection) containing randomly dispersed 2′-5′ linkages (reference 1). When these linkages are present at the moderate levels (10-25%) expected in prebiotic copying reaction, the functionality of these RNAs is retained, albeit at reduced levels. Furthermore, we observed that a dsRNA of a length typical of a ribozyme (30 nt) containing these levels of 2′-5′ linkages could be separated by thermal denaturation, even under the conditions required for nonenzymatic copying (moderate Mg^{2+} concentrations). The same dsRNA containing only 3′-5′ linkages could not fully denatured by heat, even at 95°C. Thus, we have shown that a moderate level of 2′-5′ substitution does not preclude functional RNA behaviors, and it could even have enabled phenomena enabling multiple-turnover replication that otherwise would not have been available in early life.

Additionally, we have obtained high-resolution crystal structures of RNA containing 2′-5′ linkages (reference 2). These structures show that, while 2′-5′ linkages reduce some duplex-stabilizing interactions, they do not produce global structural perturbations. Additionally, our results provide the first direct structural evidence confirming the in-line geometry associated with helical 2′-5′ linkages responsible for the well-known propensity of these linkages to degrade when in helical form. Nonetheless, our results demonstrate that even helical 2′-5′ linkages can be stable for at least the duration (1 wk) required to grow and analyze diffraction-quality crystals, indicating that the relative chemical instability of these linkages would not have been fatal in an RNA world.

Bibliography: 1) Engelhart, AE; Powner, MW; and Szostak, JW. *Nat. Chem.* (2013) 5:390-394. 2) Sheng, J; Li, L; Engelhart, AE; Gan, J; Wang, J; Szostak, JW. *Proc. Nat. Acad. Sci. USA* (2014) 111:3050-3055.

Colloquium #3: The RNA Weird: folding of nucleic acids in a nonaqueous solvent.

An aqueous environment is a universal feature of contemporary life. Hydrophobic effects are critical to protein folding, membrane formation, and countless other processes associated with all known forms of cellular life. Nonetheless, the presence of 55 M water presents a problem for the emergence of life. The principal biopolymers found in life – proteins, nucleic acids, and polysaccharides – are all the result of what are formally dehydration-condensation reactions, which are disfavored in aqueous solution. Without the sophisticated protein enzyme catalysts and highly optimized leaving groups found in modern life, such dehydration reactions appear difficult, if not impossible, in aqueous solution. This paradox has led some to suggest alternative, nonaqueous solvents for the emergence of life, due to the relative facility of dehydration reactions in such media. However, these solvents come at the cost of not being able to dissolve highly charged polymers, such as RNA and DNA, as well as often being highly denaturing to the secondary structures formed by these polymers in water.

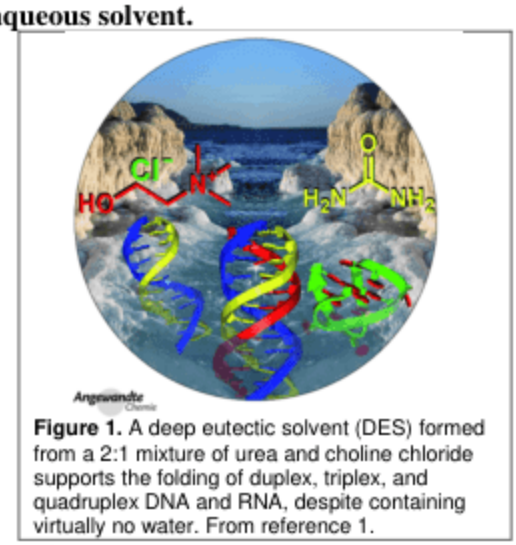

Figure 1. A deep eutectic solvent (DES) formed from a 2:1 mixture of urea and choline chloride supports the folding of duplex, triplex, and quadruplex DNA and RNA, despite containing virtually no water. From reference 1.

In this colloquium, I will present some of our recent work (reference 1) that presents a potential solution to this problem. A so-called “deep eutectic solvent” (DES) (reference 2), formed from a 2:1 mixture of urea and choline chloride, is liquid at room temperature. We have found that this DES can not only dissolve DNA and RNA, but that these polymers also form the same duplex, triplex, and quadruplex secondary structures they do in aqueous solution. Furthermore, recent work from others has demonstrated that phosphate ester formation by dehydration is possible in this solvent (reference 3). Taken together, these results suggest solvents other than water could have played a key role in the emergence of early life.

Bibliography: 1) Mamajanov, I; Engelhart, AE; Bean, HD; Hud, NV. *Angew. Chem. Intl. Ed. Engl.* (2010) 122:6454-6458. 2) Abbott, AP; et al. *Chem. Commun.* (2003) 70-71. 3) Gull, M; Zhou, M; Fernández, FM; Pasek, MA. *J. Mol. Evol.* (2014) 78:109-117.