Life today is based on the "central dogma" of molecular biology, which describes how genetic information is expressed as protein sequences. DNA information is copied into RNA, and RNA information then directs the synthesis of a protein sequence. However, these processes, as well as DNA replication, all require protein enzymes, posing a chicken-and-egg problem. If protein expression requires proteins, how did this system come about before there were proteins? Based on the dual ability of RNA to fold into catalytic structures and carry genetic information (1-5), the RNA World theory suggests that life based on the central dogma is derived from a simple living system having a single process: RNA replication, catalyzed by RNA. Much attention in this field has centered around developing an RNA polymerase ribozyme that can replicate an RNA template (6). In this issue, Papastavrou et al. demonstrate a complete ribozyme life cycle for the first time (7), discovering an RNA polymerase ribozyme that catalyzes replication, and therefore evolution, of another ribozyme.

The RNA polymerase ribozyme itself has a significant and ongoing evolutionary history in the laboratory (6). In brief, its ancestor, the class I ligase, was selected from a pool of random RNAs for its ability to ligate two RNAs together (8). Evolution of an appended domain converted the ligase into a polymerase that could add a short stretch of nucleotides across from a template RNA (9). Subsequent mutation and selection, including evolution of a domain increasing processivity, led to a version that could copy RNA templates to make catalytic RNA (10-12). However, until now, even the most advanced RNA polymerase ribozyme, dubbed 52-2 (13), could only complete half of an RNA replication cycle, such as copying a template (-) strand to produce a self-cleaving "hammerhead" ribozyme (+) strand. Due to the complementary nature of nucleic acid base-pairing, a full replication cycle would also require copying the (+) strand to reproduce the original template (-) strand (**Figure 1A**). Although 52-2 was technically capable of polymerizing long lengths of RNA, it suffered from a critical problem: insufficient fidelity.

The acute importance of fidelity was formalized several decades ago by Manfred Eigen (14). A simple model of replication with an error rate (u) showed that information of the fittest sequence (having fitness f_0 and informative length L) could only be maintained if $u < (\ln f_0)/L$. Thus, this model, and more sophisticated ones (15-17), tell us that if the error rate is higher than a critical threshold value, the information of the fittest sequence disappears in an 'error catastrophe'. The transition is sharp; with an error rate below the threshold, sequence information can be maintained and propagated (**Figure 1B**), but above the threshold, future replicator generations devolve catastrophically into random sequences (**Figure 1C**). In the case of the RNA polymerase ribozyme, while 52-2 was accurate enough to synthesize reasonably functional copies of the hammerhead ribozyme (+) from the (-) strand, it lacked sufficient fidelity to beat the error threshold for an ongoing replicative cycle of copying the (+) strand into the (-) strand and the (-) strand into the (+) strand, and so on.

Maintaining a low error rate is costly. In modern biology, the fidelity of DNA replication is ensured by multiple energetically expensive mechanisms, including kinetic proofreading and mismatch repair. In RNA viruses, which have very high error rates, increased fidelity comes at a fitness cost of decreased replication (18). Consistent with Orgel's Second Rule of evolution (approximately, "you get what you select for") (19), replicative systems tend to evolve just enough fidelity to survive the selection pressure. To push the RNA polymerase ribozyme toward higher fidelity, Papastravrou et al. realized that mutations introduced while synthesizing a

ribozyme generally decrease the mutant ribozyme's activity. Indeed, replication errors by 52-2 when synthesizing the class I ligase cause a >1000-fold drop in ligase activity. This correlation creates an opportunity to select for higher fidelity by selecting for higher activity of the ribozyme products. In addition, fidelity of the RNA polymerase ribozyme is higher in lower Mg^{2+} concentrations, but at the cost of lower catalytic activity. Papastavrou et al. took advantage of both correlations to push for higher fidelity, by evolving the RNA polymerase ribozyme under decreasing Mg^{2+} concentrations and selecting for it to produce functional, and therefore accurate, copies of the class I ligase ribozyme.

The choice of the class I ligase as the product was based on its high information content. This ribozyme contains 33 highly conserved sites and 41 partially conserved sites. By assigning bits based on conservation (e.g., 2 bits to specify one out of four possible nucleotides (A, C, G, U), 1 bit to specify 1 out of 2 nucleotides, etc.), the information content of the class I ligase has been estimated to be 95 bits (20). On the other hand, the self-cleaving hammerhead ribozyme is a smaller motif, consisting of 12 invariant sites, or 24 bits (21). A complete replication cycle of the hammerhead ribozyme would require accurate copying of 24 bits per strand, or 48 bits. Thus, Papastavrou et al. reasoned that if an RNA polymerase ribozyme were evolved to accurately copy the 95 bits of the class I ligase, it would be able to complete the 48-bit replication cycle, avoiding an error catastrophe for the hammerhead ribozyme.

Using ingenious biochemical techniques to attach each molecule of RNA polymerase ribozyme to the class I ligase it produced and then to the product of the ligase, RNA polymerase fitness was linked to ligase activity, and therefore synthesis fidelity. After 18 rounds of selection, winning RNA polymerase ribozyme 71-89 was isolated and compared with 52-2. Ribozyme 71-89 contained 10 mutations compared to 52-2 and needed much less Mg^{2+}, allowing low Mg^{2+} conditions promoting higher fidelity. During a complete cycle of hammerhead ribozyme replication, 52-2 exhibited an error rate of 18.6% per nucleotide incorporation (fidelity of 81.4%), and 71-89 had an error rate of only 10.9% (fidelity of 89.1%). Thus 71-89 showed a notable quantitative decrease in the error rate. However, because of the sharp threshold behavior of the error catastrophe, the qualitative question is much more important: is a 10.9% error rate low enough to propagate a ribozyme for many replication cycles, or not?

Papastavrou et al. tested experimentally whether 71-89 could replicate and evolve the hammerhead ribozyme for multiple rounds of in vitro evolution. The RNA polymerase ribozyme was used to copy the hammerhead ribozyme sequence back and forth into the (+) and (-) strands, for 8 complete replication cycles. At each cycle, the hammerhead ribozymes also underwent selection for self-cleaving activity. The evolutionary trajectory was followed by high-throughput sequencing.

When 52-2 was used as the RNA polymerase ribozyme in this regimen, the sequence of the hammerhead ribozyme disappeared as its information was dissipated in a few rounds, demonstrating a clear error catastrophe. In striking contrast, when 71-89 was used as the RNA polymerase ribozyme, the hammerhead ribozyme sequence survived throughout the rounds and indeed underwent noticeable evolution toward increased fitness, primarily through improved templating.

After observing the success of 71-89, one might reflect on whether 52-2 might have been able to overcome the error catastrophe. Since error threshold is proportional to $\ln f_0$, in theory, strong selection conditions could amplify fitness differences to the point that replication would tolerate the high error rate of 52-2. In practice, however, various realities may thwart this approach, as stringent conditions would also lower absolute activity, possibly to the point that degradation might outpace synthesis (22), such that solving one problem might cause new ones. Thus ribozyme 71-89 represents a form in which the 'rubber' of replication theory successfully met the 'road' of biochemical reality.

The origin of life is marked by transitions, often small quantitative changes that cause large qualitative jumps when understood in evolutionary terms. Here, a <2-fold change in the error rate of the RNA polymerase ribozyme led to the first RNA-catalyzed ribozyme replication and evolution cycle. This new functionality is heading toward something profound: RNA self-replication. If the RNA polymerase ribozyme could complete a full replication cycle of not only the hammerhead ribozyme, but of *its own* sequence, the system will meet a widely accepted definition of life as a "self-sustaining chemical system capable of Darwinian evolution" (23). Looking forward, perhaps small improvements in the quantitative traits of 71-89 – a bit more fidelity, a bit more activity – could bridge the existing gap. To paraphrase legendary coach John Wooden (24), little changes to an RNA polymerase ribozyme could make big things happen in an RNA world.

Acknowledgements
The author's research is supported by NASA (80NSSC21K0595), the Simons Foundation (290356FY18), the Sloan (AWD103574) and Moore (AWD103653) Foundations, NSF (2318736), and NIH (R35GM148249).

Figure 1. Ribozyme replication and the error threshold. (A) A complete replication cycle for the hammerhead ribozyme consists of copying the template (-) strand to produce the catalytic (+) strand (top arrow), and copying the (+) to produce the (-) strand (bottom arrow). While earlier RNA polymerase ribozymes could perform the top reaction to some extent, ribozyme 71-89 performs the full replication cycle with high fidelity. **(B and C)** Starting from the hammerhead ribozyme sequence, cycles of full replication ("R") and selection for catalytic activity of the (+) strand ("S") cause evolution. In this depiction, colored dots represent different sequences, with blue representing the wild-type hammerhead sequence and other colors representing mutants (with lower activity) generated during replication. If the mutation rate u is lower than the critical error threshold value, selection can corral the population in sequence space, maintaining the wild-type information despite replication errors (**B**). However, if u is too high, selection cannot beat the mutational overload and information is lost to an error catastrophe, and sequences diverge (**C**).

References

1. F. H. Crick, The Origin of the Genetic Code. *J. Mol. Biol.* **38**, 367-379 (1968).
2. L. E. Orgel, Evolution of the genetic apparatus. *J. Mol. Biol.* **38**, 381-393 (1968).
3. C. R. Woese, D. H. Dugre, S. A. Dugre, M. Kondo, W. C. Saxinger, On the fundamental nature and evolution of the genetic code. *Cold Spring Harb Symp Quant Biol* **31**, 723-736 (1966).
4. T. R. Cech, Structural biology. The ribosome is a ribozyme. *Science* **289**, 878-879 (2000).
5. A. Pressman, C. Blanco, I. A. Chen, The RNA World as a Model System to Study the Origin of Life. *Curr Biol* **25**, R953-963 (2015).
6. G. F. Joyce, J. W. Szostak, Protocells and RNA Self-Replication. *Cold Spring Harb Perspect Biol* **10** (2018).
7. N. Papastavrou, D. P. Horning, G. F. Joyce, RNA-catalyzed evolution of catalytic RNA. *Proc Natl Acad Sci U S A* (2024).
8. D. P. Bartel, J. W. Szostak, Isolation of New Ribozymes from a Large Pool of Random Sequences. *Science* **261**, 1411-1418 (1993).
9. W. K. Johnston, P. J. Unrau, M. S. Lawrence, M. E. Glasner, D. P. Bartel, RNA-catalyzed RNA polymerization: Accurate and general RNA-templated primer extension. *Science* **292**, 1319-1325 (2001).
10. A. Wochner, J. Attwater, A. Coulson, P. Holliger, Ribozyme-catalyzed transcription of an active ribozyme. *Science* **332**, 209-212 (2011).
11. H. S. Zaher, P. J. Unrau, Selection of an improved RNA polymerase ribozyme with superior extension and fidelity. *RNA* **13**, 1017-1026 (2007).
12. K. F. Tjhung, M. N. Shokhirev, D. P. Horning, G. F. Joyce, An RNA polymerase ribozyme that synthesizes its own ancestor. *Proc Natl Acad Sci U S A* **117**, 2906-2913 (2020).
13. X. Portillo, Y. T. Huang, R. R. Breaker, D. P. Horning, G. F. Joyce, Witnessing the structural evolution of an RNA enzyme. *Elife* **10** (2021).
14. M. Eigen, Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften* **58**, 465-523 (1971).
15. V. Shah, J. de Bouter, Q. Pauli, A. S. Tupper, P. G. Higgs, Survival of RNA Replicators is much Easier in Protocells than in Surface-Based, Spatial Systems. *Life (Basel)* **9** (2019).
16. A. S. Tupper, P. G. Higgs, Error thresholds for RNA replication in the presence of both point mutations and premature termination errors. *J Theor Biol* **428**, 34-42 (2017).
17. P. Schuster, Mathematical modeling of evolution. Solved and open problems. *Theory Biosci* **130**, 71-89 (2011).
18. V. Furio, A. Moya, R. Sanjuan, The cost of replication fidelity in human immunodeficiency virus type 1. *Proc Biol Sci* **274**, 225-230 (2007).
19. J. D. Dunitz, G. F. Joyce, Leslie E. Orgel (1927-2007). *Biographical Memoirs*, https://www.nasonline.org/publications/biographical-memoirs/memoir-pdfs/orgel-leslie.pdf (2013).
20. J. M. Carothers, S. C. Oestreich, J. H. Davis, J. W. Szostak, Informational complexity and functional activity of RNA structures. *J Am Chem Soc* **126**, 5130-5137 (2004).
21. D. E. Ruffner, G. D. Stormo, O. C. Uhlenbeck, Sequence requirements of the hammerhead RNA self-cleavage reaction. *Biochemistry* **29**, 10695-10702 (1990).
22. C. Jeancolas, C. Malaterre, P. Nghe, Thresholds in Origin of Life Scenarios. *iScience* **23**, 101756 (2020).

23. G. F. Joyce, "Forward" in Origins of life: the central concepts, D. W. Deamer, G. R. Fleischaker, Eds. (Jones & Bartlett, Boston, 1994).
24. J. Wooden, S. Jamison, *Wooden on Leadership* (McGraw-Hill, New York, ed. 1st, 2005), pp. xvi, 302 p.