Back to the future of nucleic acid self-amplification

Andrew D Ellington

The development of an autocatalytic, exponential replicator that is based solely on nucleic acids has implications for our understanding of the origins of life and potential applications in nucleic acid engineering.

It is thought that the first self-replicating systems were nucleic acid enzymes. Those replicators that were most fecund, that could exponentially amplify, likely inherited the Earth. Engineering such an ur-catalyst has long been the goal of scientists interested in understanding origins of life. However, it is difficult to imagine how a ribozyme that had to fold to catalyze replication could also itself be a template for replication. In addition, primordial template-directed, chemical polymerization of activated nucleotides may have been inhibited by nucleotides that had the wrong stereochemistry[1]. This prompted researchers to identify alternatives to polymerization as a replication mechanism, with ligation of oligonucleotides containing nonstandard nucleotides being the obvious work-around. Recently, the Joyce lab has engineered a cross-catalytic system in which one ligase ribozyme serves as a template for the other[2]. These authors have also adapted this system to biotechnology applications by appending ligand-binding domains (aptamers) to the catalytic core. The resultant 'aptazymes' participate in a ligand-dependent exponential cascade, allowing small molecules to be detected without the aid of protein catalysts or more traditional nucleic acid amplification methods such as PCR[3].

The exponential replication system that the Joyce lab has developed was the culmination of years of work. The original catalyst was a ligase, R3C, whose parent was originally selected from a random sequence population that lacked cytidine[4], although in later selections that nucleotide was let back in[5]. This ligase ribozyme was broken into two pieces such that the ligation junction was symmetric, allowing the same ribozyme to act as a template for rejoining these pieces (and thus recreating the ribozyme[6]) (**Fig. 1a**). In order to overcome product inhibition in the fully symmetric system, the two catalysts were altered so that they differed from one another, leading to a cross-catalytic system in which four substrates could act to recreate two catalysts[7].

Andrew D. Ellington is in the Department of Chemistry and Biochemistry and the Center for Systems and Synthetic Biology, University of Texas at Austin, Austin, Texas, USA.
e-mail: andy.ellington@mail.utexas.edu

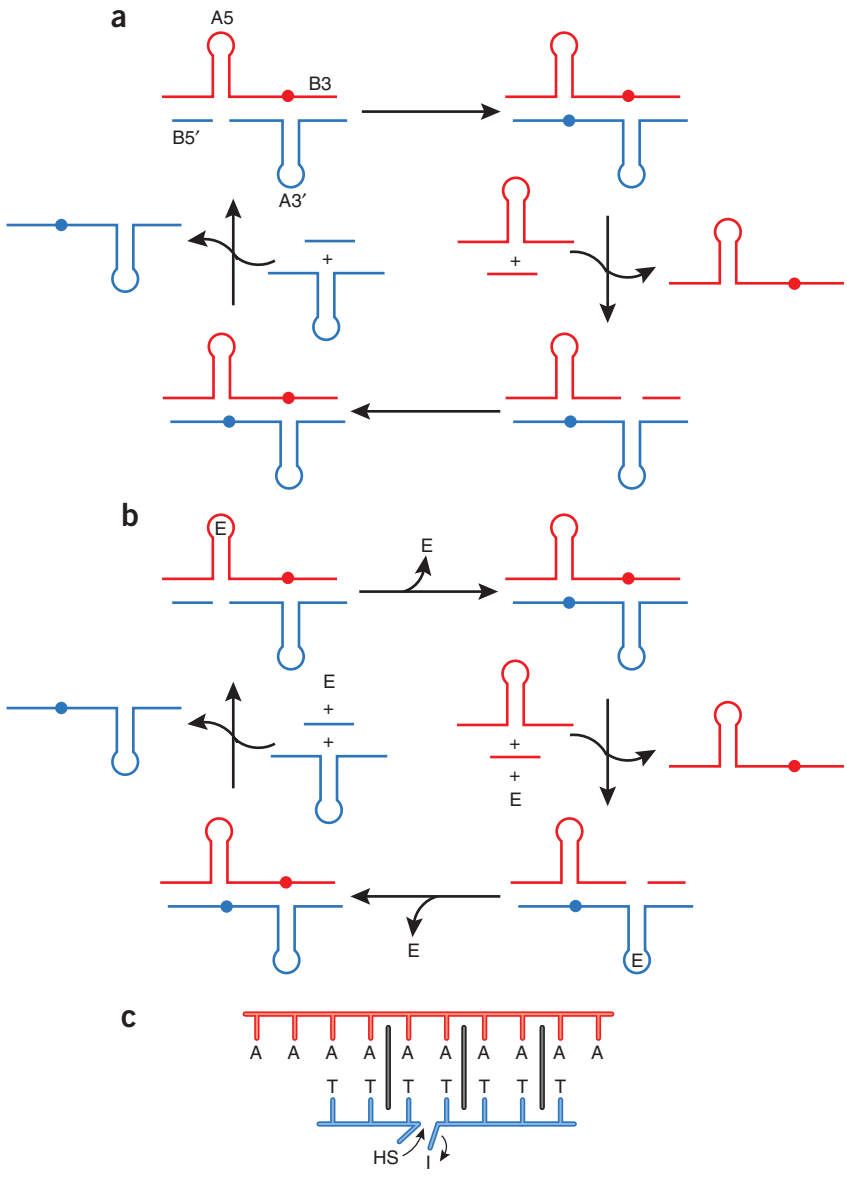

Figure 1 Autocatalytic ligase cycles. (**a**) The autocatalytic ligase cycle developed by Lincoln and Joyce[2]. Half-ligases are aligned by a complementary ligase and can be released following ligation. Dots represent ligated junctions. Red and blue lines represent ligases with complementary but different templates. The recombinant pair A5B3 is shown cross-replicating B5′A3′. (**b**) The adaptation of the autocatalytic ligase cycle to the detection of an effector (E)[3]. The effector accelerates the rate of ligation, allowing the autocatalytic cycle to proceed more quickly. Although addition and release of the effector are shown to occur at specific steps, these could occur at any point in the cycle. (**c**) Mechanisms for the adoption of prebiotic compounds by early replicating systems. The 'molecular midwife' hypothesis of Hud and Anet[11] suggests that intercalators may have served to partially template base-stacking and possibly polymerization of early nucleotides. Intercalators assisting with chemical ligation are shown. SH indicates a 3′ phosphorothioate, and I indicates a 5′ iodide moiety. Nucleophilic attack leads to the formation of a phosphorothioate linkage.

In the manuscript published in *Science*[2], the cross-catalytic system was improved by directed evolution. One of the half-ribozyme substrates was connected to its complementary catalyst via a hairpin stem, and both this construct and the other half-ribozyme substrate were mutagenized. Selection for ligation with incubation times as short as 10 ms led to improved catalytic activity. The resultant ribozyme pair was found to work extremely well in the cross-catalytic reaction (**Fig. 1a**); the pair had an exponential growth rate of ~1 per hour, and the two ribozymes could continuously amplify one another upon serial transfer.

To gain potential insights into how self-replicators could further evolve, the ribozyme cascade was converted to a replicating ecosystem by altering the half-ribozymes so that there were several unique, cross-catalytic pairs. The catalysts were also mutated to contain a separate (nonlethal) sequence substitution in the catalytic domain. Thus, the efficiency of replication for each pair should be based both upon the kinetics of base pairing and product release, and the catalytic activities of the core ribozymes.

The various pairs were competed head-to-head in a serial transfer experiment. After 20 serial transfers and an overall amplification of greater than 10^{25}-fold, there was surprisingly no clear 'winner'; however, the limited number of ribozymes remaining in the population did point to preferential propagation of some RNA species. The authors observed that in addition to self-amplifying their cognate partner, the ligases could also potentially ligate noncognate partners, producing new, unanticipated combinations that could themselves carry out additional ligation reactions. The most efficient replicator was a recombinant, A5B3 (half-ribozyme A5 with half-ribozyme B3, rather than with their original partners B5 and A3, respectively; **Fig. 1a**). The complementary ribozyme of the pair, B5′A3′, was also well represented, as might have been expected. In addition, though, there were a number of other recombinants linked to either half-ribozyme A5 or half-ribozyme B5′ that were well represented in the selected population. In the end, the success of the dominant recombinant pair was based in part on its ability to partially pair with and cross-replicate numerous other substrates. In other words, it was the very messiness of the system that led to the success of A5B3 and its partner.

In short order Lam and Joyce[3] adapted this exponential amplification system to function as a biosensor platform (**Fig. 1b**). Both antitheophylline and antiflavin aptamer domains were added to the ligase in a manner that rendered them ligand dependent (creating so-called aptazymes). When the corresponding exponential cascade was generated, no amplification was seen in the absence of ligand, but when ligand and both half-aptazyme pairs were available, exponential amplification took place, resulting in nearly complete accumulation of products within about 12 h. A dual sensor system, in which theophylline-dependent enzymes amplified flavin-dependent catalysts, and vice versa, was also created. A pyrophosphate-dependent luciferase could be used to monitor pyrophosphate release during ligation.

By examining exponential amplification rates as a function of analyte concentrations, it was determined that the autocatalytic ribozyme biosensors had apparent K_d values for theophylline and flavin of 51 μM and 68 μM, respectively, and can likely be used to quantitate analytes into the low micromolar range. The aptamers used to create the aptazymes had K_d values that were 10 to 50 times lower, but there is no free lunch when using conformational switching as a strategy to create biosensors: some binding energy must be transduced into the ligand-dependent conformational change[8,9]. This limitation could be overcome by energetically poising the aptazymes to more readily access the ligand-dependent conformation, but that would also lead to greater background ligation. To distinguish analyte-dependent ligation from background ligation, the rate of amplification (rather than its extent) could be monitored. By comparing the 'breakthrough' of amplification (similar to a Ct value in real-time PCR) in the absence and presence of analytes, it might be possible to further extend the sensitivity of the autocatalytic ligase biosensors.

One of the more interesting aspects of these remarkable autocatalytic reaction cycles is that they emphasize the similarity between the challenges in understanding the origins of life and the challenges in developing useful biotechnologies. One of the problems that confronted the earliest replicators, as well as the Joyce lab, was that a ligase almost of necessity creates a product that is better able to base-pair with its template, and thus is less likely to be released. Ligation inherently yields product inhibition. Because of this, cross-catalytic ligases are confronted with the prospect of limiting their own replication, a phenomenon that is termed 'parabolic amplification'. Whereas exponential replicators can readily increase their numbers at the expense of other competing replicators, parabolic replicators cannot outcompete their peers, and as Guenter von Kiedrowski[10] has pointed out, this leads to the "survival of everyone." Until the careful engineering of the Joyce lab showed that exponential replication by a tetrapartite ligase system was possible, product inhibition was considered to be a (possibly fundamental) limitation on the ability of early nucleic acid enzymes to replicate themselves. Now that autocatalytic oligonucleotide ligation is a possibility, it is not unreasonable to expect that elaborating the basic chemistries of the oligonucleotide substrates to include nonstandard nucleotides (and thus making them even more 'messy') might also potentiate robust self-selection.

Finally, these recent papers highlight one of several ways in which early replicators may have taken advantage of other compounds in the prebiotic world. It is possible that intercalators in the environment may have inadvertently assisted with the initial replication of nucleic acids—the so-called molecular midwife hypothesis[11] (**Fig. 1c**). Once early replicators depleted their environment of nucleotide substrates, they may have evolved additional functionalities to scavenge other available precursors. For example, Unrau and Bartel have used directed evolution to generate a ribozyme that can synthesize a glycosidic bond, and doppelgangers of this ribozyme might have once helped form nucleotides from nucleobases and sugars[12]. And now Joyce and co-workers show that interactions between autocatalytic cycles and other molecules (to wit, theophylline versus flavin) can lead to the preferential amplification of some replicators over others. This would seem to imply that in addition to developing a nascent metabolism by expanding the range of substrates, an autocatalytic cycle could also just acquire any molecule that led to faster ligation rates. Though this is a contrived system, it is nonetheless an excellent stepping stone toward the cross-catalytic hypercycles between RNA and other compounds (peptides) proposed long ago by Eigen and Schuster[13]. Given the proximity of the Joyce and Ghadiri labs at the Scripps Research Institute, perhaps such cross-replicating nucleic acid and peptide hypercycles[14] are already being born, either planned or unplanned.

1. Joyce, G.F. *et al. Nature* **310**, 602–604 (1984).
2. Lincoln, T.A. & Joyce, G.F. *Science* **323**, 1229–1232 (2009).
3. Lam, B.J. & Joyce, G.F. *Nat. Biotechnol.* advance online publication, doi: 10.1038/nbt.1528 (22 February 2009).
4. Rogers, J. & Joyce, G.F. *Nature* **402**, 323–325 (1999).
5. Rogers, J. & Joyce, G.F. *RNA* **7**, 395–404 (2001).
6. Paul, N. & Joyce, G.F. *Proc. Natl. Acad. Sci. USA* **99**, 12733–12740 (2002).
7. Kim, D.E. & Joyce, G.F. *Chem. Biol.* **11**, 1505–1512 (2004).
8. Soukup, G.A. & Breaker, R.R. *Proc. Natl. Acad. Sci. USA* **96**, 3584–3589 (1999).
9. Robertson, M.P. & Ellington, A.D. *Nucleic Acids Res.* **28**, 1751–1759 (2000).
10. von Kiedrowski, G. *Bioorg. Chem. Frontiers* **3**, 115–146 (1993).
11. Hud, N.V. & Anet, F.A. *J. Theor. Biol.* **205**, 543–562 (2000).
12. Unrau, P.J. & Bartel, D.P. *Nature* **395**, 260–263 (1998).
13. Eigen, M. & Schuster, P. *Naturwissenschaften* **64**, 541–565 (1977).
14. Lee, D.H. *et al. Nature* **390**, 591–594 (1997).