JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
pubs.acs.org/JACS

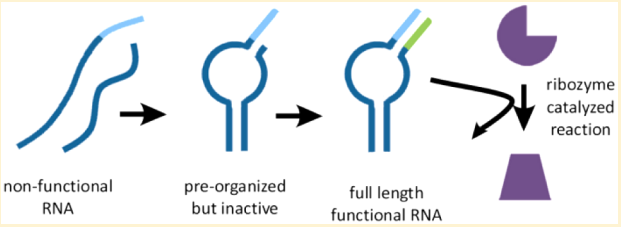
Generation of Functional RNAs from Inactive Oligonucleotide Complexes by Non-enzymatic Primer Extension

Katarzyna Adamala,† Aaron E. Engelhart, and Jack W. Szostak*

Howard Hughes Medical Institute and Department of Molecular Biology and Center for Computational and Integrative Biology, Massachusetts General Hospital, 185 Cambridge Street, Boston, Massachusetts 02114, United States

Supporting Information

ABSTRACT: The earliest genomic RNAs had to be short enough for efficient replication, while simultaneously serving as unfolded templates and effective ribozymes. A partial solution to this paradox may lie in the fact that many functional RNAs can self-assemble from multiple fragments. Therefore, in early evolution, genomic RNA fragments could have been significantly shorter than unimolecular functional RNAs. Here, we show that unstable, nonfunctional complexes assembled from even shorter 3′-truncated oligonucleotides can be stabilized and gain function via non-enzymatic primer extension. Such short RNAs could act as good templates due to their minimal length and complex-forming capacity, while their minimal length would facilitate replication by relatively inefficient polymerization reactions. These RNAs could also assemble into nascent functional RNAs and undergo conversion to catalytically active forms, by the same polymerization chemistry used for replication that generated the original short RNAs. Such phenomena could have substantially relaxed requirements for copying efficiency in early nonenzymatic replication systems.



INTRODUCTION

Prior to the evolution of protein-based enzymes, primitive cells are thought to have relied on RNA enzymes for catalysis.[1−3] The hypothesis of such an early, RNA-dominated stage in the evolution of life is supported by the continued existence of ribozymes in modern biology,[4,5] and especially by the observation that protein synthesis is catalyzed by the RNA component of the large ribosomal subunit.[6] In addition to this evidence, ribozymes obtained by *in vitro* directed evolution can catalyze a wide variety of reactions, encompassing all of the basic classes of transformations necessary for protocell metabolism and reproduction.[7,8] However, the origins of RNA-based protocells remain unclear, and numerous problems must be solved to establish an understanding of how prebiotic chemical systems gave rise to protocells containing both replicating genomic RNAs and functional ribozymes derived from those genomic RNAs.

One of the central problems faced by the earliest cells is likely to have been the slow and inefficient nature of primordial RNA replication. Indeed, prior to the evolution of a sophisticated RNA replicase ribozyme, protocells may have had to rely on the ambient chemical environment to drive non-enzymatic RNA replication. Efforts to demonstrate prebiotically plausible RNA replication chemistry date back some four decades,[9] but despite recent progress,[10−13] several obstacles remain.[14] At present, non-enzymatic RNA polymerization allows for the partial copying of short RNA templates. In model studies, the growing RNA primer strand is elongated by addition of ribonucleotides activated with a good leaving group (e.g., 2-methylimidazole or hydroxyazabenzotriazole).[15−17]

In this paper, we assume that the unsolved obstacles to non-enzymatic RNA replication reflect our current lack of understanding, and are not insurmountable hurdles. However, we also expect that primitive, chemically driven RNA replication was only marginally sufficient for the emergence of protocells bearing small replicating fragments of RNA. We then ask: given such a situation, how could protocells bearing selectively advantageous, heritable ribozyme activities have emerged?

Ribozymes with a k_{cat} of greater than 1 min^{-1} are generally at least 30 nucleotides in length and often 50−100 nucleotides long; the few known ribozymes with a k_{cat} of greater than 1 s^{-1} are over 100 nucleotides in length.[18−22] Generating RNAs of 50−100 nucleotides in length by non-enzymatic template-directed polymerization seems like a severe challenge. Fortunately, several previous reports have shown that such ribozymes can be reconstituted from shorter oligonucleotides by their assembly into a noncovalent complex that retains good catalytic activity.[23−27] The minimal length of such oligonucleotides reflects the need for stable base-paired stems as the driving force for complex assembly. If each oligonucleotide is anchored in the complex by two or three stems of 5−6 nucleotides, with interspersed sequences of similar length forming binding or catalytic sites, then a minimal fragment length of 15−25 nucleotides would be reasonable. We, therefore, hypothesize that primitive cells contained a set of genomic oligonucleotides in this size range, all being replicated by non-enzymatic chemistry, with some of these oligonucleotides assembling into complexes

Received: November 11, 2014
Published: December 18, 2014

DOI: 10.1021/ja511564d
J. Am. Chem. Soc. 2015, 137, 483−489

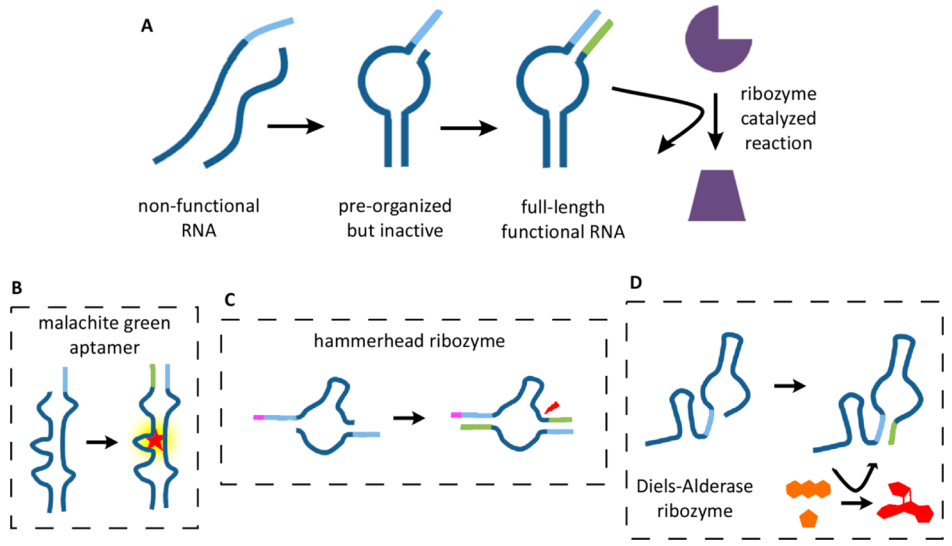


Figure 1. General scheme for activation of functional RNAs by nonenzymatic RNA synthesis. (A) 3′-Truncated oligonucleotides assemble into an unstable, inactive structure. After extending one or more 3′-ends on a template derived from another region of the truncated functional RNA, the full-length functional RNA folds into a stable and active aptamer or catalyst. (B–D) Schematic diagrams of the aptamer and two ribozyme systems used in this study. Template region, light blue; extended primer region, green.

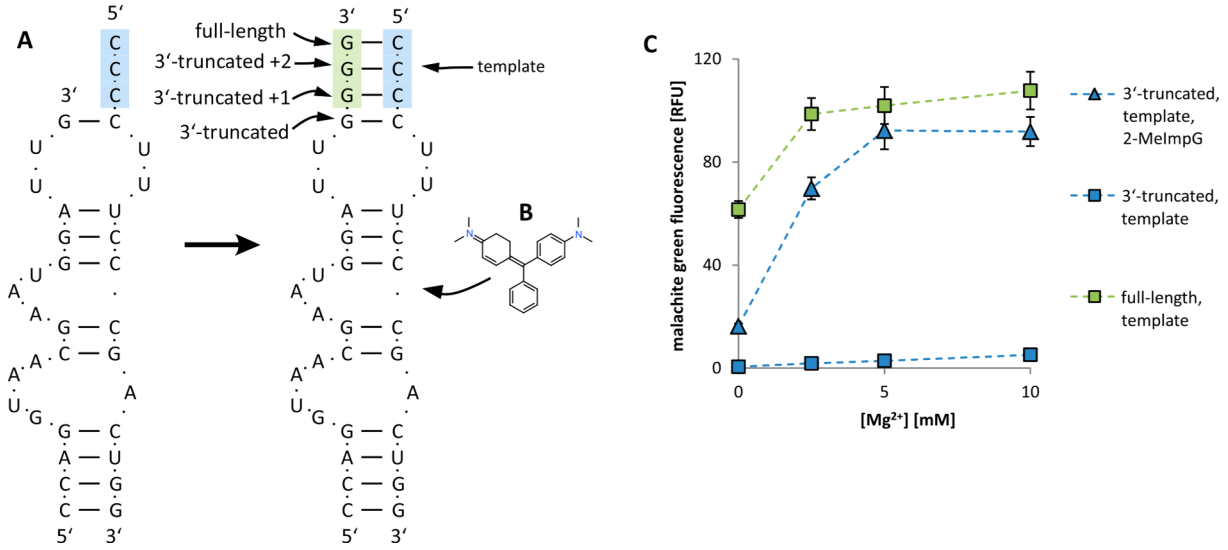


Figure 2. Malachite green aptamer. (A) Structure and sequence of the malachite green aptamer before and after primer extension. Blue bases, template region; green bases, primer extension region. (B) One molecule of malachite green binds to each fully assembled aptamer. (C) Fluorescence of malachite green in the presence of aptamer RNA at different concentrations of Mg^{2+}. Blue squares, unextended 3′-truncated strand plus template strand; blue triangles, 3′-truncated strand plus template after 24 h of RNA primer extension with 2-MeImpG; green squares, 3′-truncated strand after 24 h extension with 2-MeImpG and template strand. Each sample contained 0.25 M Tris-HCl pH 8.0, 0.15 M NaCl, 2 μM malachite green, 1 μM each strand of the aptamer. Error bars indicate SEM, $N = 3$.

with useful catalytic activities. If it were possible for even shorter genomic oligonucleotides to generate functional ribozyme complexes, the constraints on RNA replication chemistry would be correspondingly diminished.

Here we show that functional ribozymes can be reconstituted from oligonucleotides that have been further shortened by 3′-truncation. In this scenario, the 3′-truncated oligonucleotides assemble into unstable and inactive precursor complexes that are held together by very short unstable terminal stems. Non-enzymatic template-directed extension of the oligonucleotide 3′-termini, which act as primers, lengthens the initially short and unstable base-paired stems, generating stable complexes with good catalytic activity. Such extended oligonucleotides could be removed from the pool of replicating RNA fragments, as they would be locked into stable catalytic complexes. Non-extended genomic fragments would be easier to replicate, owing to both the comparative ease of copying short RNA sequences and the fact that they were not sequestered in stable complexes. This simple mechanism substantially relaxes the requirements for RNA replication systems, and it could have facilitated the transition from prebiotic chemical systems to evolving living cells.

RESULTS

We studied an aptamer and a *cis*-acting ribozyme that could each be reconstituted by the assembly of two oligonucleotides into a

DOI: 10.1021/ja511564d
J. Am. Chem. Soc. 2015, 137, 483–489

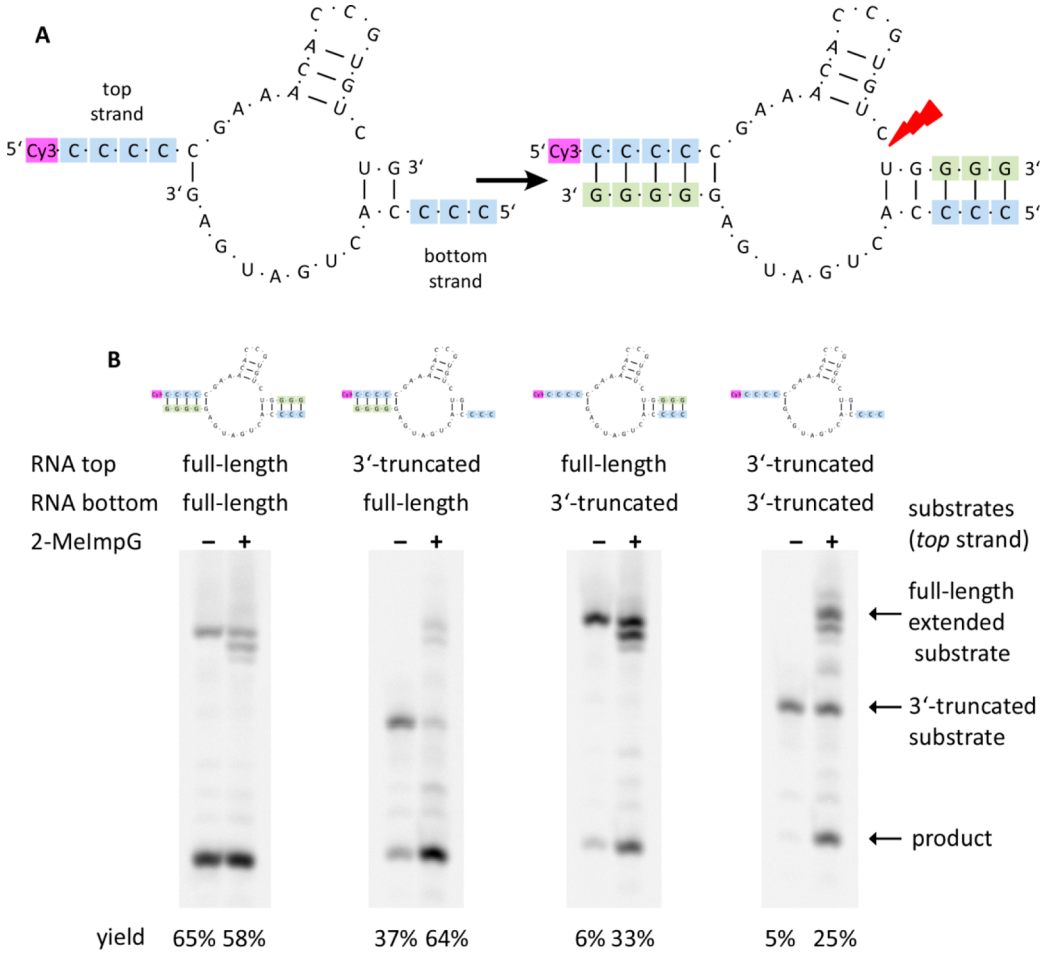


Figure 3. Hammerhead ribozyme. (A) Structure and sequence of the hammerhead ribozyme before and after primer extension. The ribozyme is assembled from top and bottom strands; each of the strands contains a template region (blue) and a primer extension region (green). The red arrow indicates the cleavage site. The 5′-end of the top strand is fluorescently labeled with Cy3 for PAGE analysis. (B) PAGE analysis of hammerhead ribozyme self-cleavage products. Each panel shows two parallel reactions: samples incubated for 24 h with and without 2-MeImpG (i.e., with or without primer extension). The sequences above the gel panels show starting material in each sample: two hammerhead ribozyme strands, each as either unextended 3′-truncated or full-length strand. Cleavage product yields are given below each lane. Each sample contained 0.25 M Tris-HCl pH 8, 0.15 M NaCl, 2.5 μM each RNA strand, 50 mM 2-MeImpG, 50 mM $MgCl_2$. 2-MeImpG-containing lanes exhibit several minor bands, which are believed to be derived from 2-MeIm-induced RNA degradation.[33]

functional complex, held together in part by terminal base-paired stems. We also studied a second, *trans*-acting ribozyme consisting of a single oligonucleotide, the active form of which required the formation of a base-paired stem between its 3′-terminus and an internal sequence. We then modified the structure of each of these oligonucleotides by shortening the 3′-end by 3 or 4 nucleotides; as a result, the 3′-truncated oligonucleotides were unable to assemble into functionally active complexes. However, these shortened oligonucleotides could assemble into transiently stable complexes, in which the oligonucleotide 3′-ends formed one or two base-pairs in the nascent stem. Such assembly resulted in a "primer–template" configuration, in which the 3′-end could be extended by non-enzymatic polymerization, thus generating an extended and, therefore, more stable base-paired stem. In each case, we were able to observe the recovery of ligand binding or catalytic activity following non-enzymatic template-directed primer extension (Figure 1).

Malachite Green Aptamer. An RNA aptamer that binds malachite green has been previously reported.[28] We chose to work with this aptamer for an initial proof-of-principle of the hypothesis that 3′-truncated oligonucleotides can self-assemble into intermediate structures that can serve as templates for the reconstitution of functional RNAs. The free malachite green ligand has very low fluorescence, but its fluorescence increases by several orders of magnitude upon binding to the RNA aptamer.[29] This is a convenient signal for the formation of a functional aptamer. In addition, the aptamer and its mode of ligand binding have been studied extensively.[30] NMR studies indicate that the terminal base-paired stems form in the absence of ligand, but the central region of the aptamer folds around its ligand, and it becomes ordered only upon ligand binding. Thus, it seemed likely that the ability of the aptamer to fold and bind its ligand could be readily controlled by altering the length of its terminal stems. Additionally, this aptamer has been previously deconstructed into two oligonucleotides that can assemble into a functional complex.[29]

To demonstrate aptamer stabilization by non-enzymatic template-directed RNA synthesis, we modified the previously reported sequences of the two-fragment malachite green aptamer.[29] We first changed the original design of the binary aptamer so that one of the base-paired stems is a four nucleotide GGGG/CCCC stem (Figure 2); this aptamer retains full

DOI: 10.1021/ja511564d
J. Am. Chem. Soc. 2015, 137, 483−489

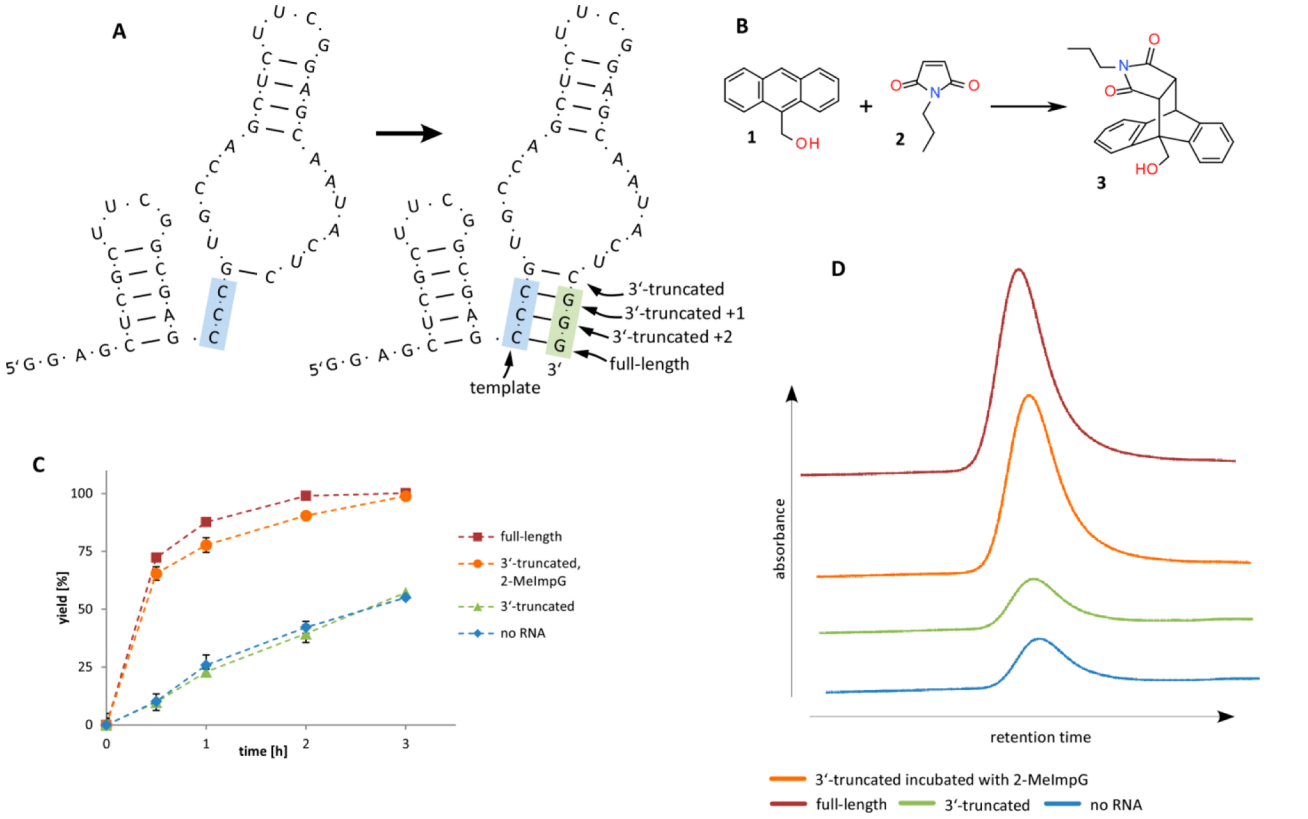


Figure 4. Diels–Alderase ribozyme. (A) Structure and sequence of the Diels–Alderase ribozyme before and after primer extension. Blue bases, template region; green bases, 3′-truncated region. (B) The reaction between *N*-propylmaleimide and 9-hydroxymethylanthracene, catalyzed by the Diels–Alder ribozyme. (C) Time course of the Diels–Alder reaction with and without the ribozyme. Red squares, full-length ribozyme; orange circles, 3′-truncated ribozyme after 24 h extension with 2-MeImpG; green triangles, nonextended 3′-truncated ribozyme; blue diamonds, uncatalyzed reaction. Reaction conditions: 100 mM Tris-HCl pH 7.5, 100 mM $MgCl_2$, 300 mM NaCl, 200 μM 9-hydroxymethylanthracene, 1000 μM *N*-propylmaleimide, 10 μM RNA. Error bars indicate SEM, $N = 3$. (D) Product region of HPLC traces of samples stopped at time 0.5 h, corresponding to the reactions on graph in panel C. Detection wavelength, 260 nm; product peak retention time, 12.5 min.

malachite green binding activity. However, if the 3′-terminal GGGG sequence is shortened to a single G, there is no detectable binding of malachite green. If the 3′-terminal GGGG sequence is shortened to two or three guanosine nucleotides, malachite green binding is reduced, but not eliminated (Figure S1). In the presence of the activated monomer guanosine 5′-monophosphate 2-methylimidazolide (2-MeImpG), non-enzymatic extension of the single 3′-terminal G on the CCCC template regenerates the GGGG/CCCC stem and results in a stable, functional aptamer (Figure 3 and Figure S1). After a 24 h room temperature incubation of 3′-truncated and template strands with 2-MeImpG, the measured malachite green binding was comparable to that of the aptamer assembled from two full-length oligonucleotides (Figure 2C). This experiment shows that inactive functional RNAs, made from oligonucleotides that are truncated at the 3′-end, can be converted to a functional form by non-enzymatic template-directed RNA synthesis.

Hammerhead Ribozyme. Having established that a functional RNA aptamer can be generated by non-enzymatic extension of the 3′-end of one oligonucleotide on a template region of a second oligonucleotide, we investigated the possibility of generating an active ribozyme from two truncated oligonucleotides, each of which could be extended on the other. We chose the hammerhead ribozyme, which cleaves a substrate strand at a defined cleavage site,[31] for these experiments. We modified a previously reported minimized consensus sequence so that the ribozyme could be reconstituted from two strands, referred to as "top" and "bottom".[32] These strands form a complex held together by two stems flanking the catalytic loop region (Figure 3).

The top strand, which contains the cleavage site, was labeled at its 5′-end with Cy3. Ribozyme activity was measured by fluorescence-detected polyacrylamide gel electrophoresis (PAGE) gel separation of the full-length fluorescently labeled top strand from the shorter, cleaved product. When we mixed full-length top and bottom strands in equimolar proportions and incubated the resulting complex at room temperature, ribozyme activity was evident from the appearance of cleaved product (65% product in the reaction incubated with 2-MeImpG, Figure 3, and Experimental Section). We then prepared 3′-truncated versions of both the top and bottom strands. Full-length top strand mixed with truncated bottom strand, and truncated top strand mixed with full-length bottom strand, resulted in weak ribozyme activity in both cases. Incubation of the 3′-truncated top and bottom strands resulted in no ribozyme activity (Figure 3). We then examined the effect of incubating the above oligonucleotide mixtures with the activated monomer 2-MeImpG. In the complexes assembled with 3′-truncated oligonucleotides, both strands can act simultaneously as primer and template: the 3′-end of the top strand can be extended by up to 3 guanosine residues (employing the 5′-end of the bottom strand as a template), while the 3′-end of the bottom strand can be extended by up to 4 guanosine residues (employing the 5′-end of the top strand as a template). In all cases, incubation with 2-

DOI: 10.1021/ja511564d
J. Am. Chem. Soc. 2015, 137, 483–489

MeImpG resulted in extension of the 3′-ends of the truncated oligonucleotides and reconstitution of ribozyme activity. In primer extension reactions, if the top 3′-truncated strand was incubated with the bottom full-length strand, 64% cleavage yield was obtained; the top full-length strand incubated with the bottom 3′-truncated strand resulted in 33% product, and if both top and bottom strands were 3′-truncated, 25% product was observed (compared to only 5% background cutting activity by mixing both 3′-truncated strands without 2-MeImpG) (Figure 3). 3′-Truncation of the bottom strand is generally more deleterious than that of the top, likely due to the additional base pair present in the complex of the 3′-truncated top strand with either bottom strand.

Diels−Alderase Ribozyme. Encouraged by the above results, we sought to examine whether a *trans*-acting ribozyme (i.e., a ribozyme that catalyzes a reaction between two substrates not covalently bound to the ribozyme) could be activated in a similar manner. We started with the minimized sequence of a Diels−Alderase ribozyme,[34] which catalyzes a Diels−Alder cycloaddition reaction between a broad range of substrates[35] with very good stereoselectivity.[36] We redesigned the ribozyme, in this case composed of a single oligonucleotide, to end with the sequence 5′-GGG-3′, which could pair with the internal sequence 5′-CCC-3′ (Figure 4A). We then constructed a set of truncated versions ending in one or two 3′-G residues, which we expected to be less active than the full-length ribozyme.

We followed the reaction between the substrates *N*-propylmaleimide and 9-hydroxymethylanthracene by HPLC. The background (i.e., uncatalyzed) reaction, in the absence of the catalyst, resulted in conversion of 11% of the anthracene substrate to product in 30 min, and 26% in 1 h. The truncated ribozyme ending in a single 3′-G did not result in any significant increase in product formation, while the full-length ribozyme yielded 72% product in 30 min and 88% in 1 h. A synthetic ribozyme variant ending in two 3′-G residues had intermediate activity (Figure S2).

In order to assess the ability of non-enzymatic primer-extension to convert the inactive truncated ribozyme ending in a single 3′-G residue to an active ribozyme, we incubated the truncated ribozyme with the activated monomer 2-MeImpG for 24 h. The resulting extended ribozyme was nearly as effective a catalyst as the full-length ribozyme, converting 65% of substrates in 30 min and 78% in 1 h (Figure 4).

DISCUSSION

Our observations indicate that functional ribozymes can be generated from non-functional sets of short RNA oligonucleotides by using non-enzymatic template-directed RNA synthesis to convert unstable and inactive oligonucleotide complexes into stable active assemblies. We have used three separate model systems to show that inactive sets of 3′-truncated oligonucleotides can be converted to active aptamers or ribozymes following incubation with the activated monomer 2-MeImpG. In each case, the 3′-end of one or more oligonucleotides was extended by non-enzymatic template-directed primer extension, converting a very weakly base-paired stem (generally a single base-pair) into a stable base-paired stem consisting of four G:C pairs. As a result, the functionally relevant structure was stabilized, and binding or catalytic activity appeared.

We suggest that the first ribozymes may have originated within protocells in a similar three-step process. First, multiple short RNA fragments replicated within protocells by non-enzymatic template-directed RNA synthesis. These RNAs would have been generated in higher-yield than full-length RNAs, owing both to their lowered propensity to fold and the higher efficiency of copying short RNAs non-enzymatically. Second, subsets of these short RNA oligonucleotides assembled into inactive ribozyme precursors. Third, non-enzymatic RNA copying chemistry (potentially the same chemistry that generated the initial fragments) extended the 3′-ends of the ribozyme precursor oligonucleotides, using other regions of RNA in the transient complex as templates, yielding fully functional stable ribozyme complexes. As a result of such a scenario, short, nonfunctional RNAs could afford a protocell vesicle fitness, given their ability to assemble into transient complexes and undergo subsequent primer extension to elaborate these complexes into full-length functional RNAs. As we have recently demonstrated non-enzymatic primer extension inside fatty acid vesicles, such a scenario seems highly plausible.[13] Additional support for our hypothesis is provided by recent results demonstrating that Fe^{2+} can substitute for Mg^{2+} in functional RNAs, suggesting that the low solubility of Mg^{2+} in the presence of certain counterions, such as phosphate, may not have been as problematic as anticipated, given the high availability of soluble iron prior to the Great Oxidation Event.[37] While our system incorporates only 2-MeImpG, the presence of the other three nucleotide monomers 2-MeImpA, 2-MeImpC, and 2-MeImpU does not impair aptamer or ribozyme reconstitution (Figures S3 and S4), demonstrating that even prebiotically plausible levels of nonenzymatic replication fidelity allow for functional RNA reconstitution via this mechanism.

Given the enhanced thermal stability of the products of primer extension (Figure S5) and increases in aptamer and ribozyme activity with increasing stem length (Figures S1 and S2), it is likely that these RNAs develop function as a result of stabilization of critical structural stems. It is also possible that some of the enhanced function observed in the RNAs occurs as a result of these lengthened homopolymeric stems directing the RNA to assume the correct fold. Both such results would occur as a direct result of the extension of truncated stems.

Our observations have significant implications with regard to two aspects of the evolution of functional RNAs in early cells. First, our results suggest a natural way in which genomic RNA fragments, which must be replicated to enable heritability, can be shorter than the RNAs that actually assemble into functional complexes. Multiple previous examples have shown that aptamers and ribozymes can often be reconstituted from relatively short oligonucleotides.[23,24,38,39] This alone shows that primitive RNA replication systems, whether chemical or ribozyme-catalyzed, need only be sufficient to copy fairly short pieces of RNA. However, since primordial non-enzymatic replication systems are expected to be slow and inefficient, and, in fact, no robust copying system has yet been demonstrated experimentally, it remains useful to consider scenarios in which the length of genomic RNAs can be further minimized. Our results suggest that genomic RNA fragments could be at least 3−4 nucleotides shorter than functional RNA fragments, further ameliorating the difficulty associated with primitive RNA replication. Thus, the process we describe for synthesis of functional RNAs relaxes the requirements for prebiotic RNA copying efficiency.

The second aspect of primitive RNA replication that is facilitated by the scenario outlined above has to do with the conflict between the requirements for a good template as opposed to those of a good catalyst. Optimal templates should be completely unfolded, with no internal structure to impede the

DOI: 10.1021/ja511564d
J. Am. Chem. Soc. 2015, 137, 483−489

synthesis of a complementary strand. Similarly, a good template should not be sequestered within a stable complex with other oligonucleotides. In contrast, a good ribozyme must have a stable folded structure, and if the ribozyme is generated through the association of two or more oligonucleotides, the resulting complex should also be a stable structure. This conflict is partially resolved through the strategy of assembling functional complexes from sets of relatively unstructured oligonucleotides; however, in order to allow for uncomplexed templates and functional complexes to co-exist, the affinity of the oligonucleotides for each other must be quite weak. Such a loose association could make it difficult to form a structurally well-defined complex, although this deserves closer experimental attention. Nonetheless, our observations suggest that nonenzymatic primer-extension may provide a natural way to allow for the co-existence of short unstructured templates, and slightly longer derivative oligonucleotides that form stable, functional complexes. We have previously suggested that randomly positioned 2′−5′ linkages, generated as a consequence of non-enzymatic RNA replication, could also contribute to the distinction between good templates and good ribozymes.[40] We suggest that the effects of 2′−5′ linkages could act synergistically with the effects of non-enzymatic primer-extension to produce distinct sets of genomic and functional oligonucleotides in primitive cells. The reaction conditions we have employed generate predominantly 3′−5′ linkages, but the small proportion of 2′−5′ linkages produced could have been beneficial in this regard.[41]

EXPERIMENTAL SECTION

Oligonucleotides. All oligonucleotides were obtained from IDT with HPLC purification.

Synthesis of 2-MeImpN. Nucleotide 2-methylimidazolides (2-MeImpN, where N = G, A, C, U) were synthesized as previously described.[42]

Synthesis of the Product of the Diels−Alder Reaction. The product of the Diels−Alder reaction used in our studies (**3**) was synthesized, for use as a standard, by refluxing 0.625 g (3 mmol) of 9-hydroxymethylanthracene (**1**) and 0.390 g (2.8 mmol) of *N*-propylmaleimide (**2**) in 15 mL of dry toluene for 14 h. The product was purified by silica gel chromatography with 1:1 hexane:ethyl acetate, and it was obtained as a yellow solid in 71% yield. ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.27 (d, 2H), 7.36 (d, 2H), 7.24 (m, 4H), 5.29 (s, 1H), 5.11 (m, 2H), 4.75 (s, 1H), 3.32 (m, 2H), 3.08 (m, 2H), 1.63 (s, 1H), 0.86 (m, 2H), 0.495 (t, 3H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 177.34, 177.06, 142.42, 142.40, 139.57, 139.24, 127.15, 127.11, 126.85, 126.76, 125.50, 124.27, 123.35, 122.74, 60.90, 49.56, 48.01, 46.55, 45.86, 40.36, 20.65, 11.15.

General Procedure for Non-enzymatic RNA Polymerization. Non-enzymatic RNA polymerization was initiated by addition of activated RNA monomer(s) to the mixture of 3′-truncated and template oligonucleotides. Specific reaction conditions are given below for each functional RNA system. All reaction mixtures were incubated at room temperature for specified times, desalted using ZipTip C_{18} pipet tips (Millipore), and analyzed by TBE−urea 20% PAGE.

Malachite Green Aptamer. The binding of malachite green to the aptamer was measured with a SpectraMAX GeminiEM fluorescence plate reader (Molecular Devices) with excitation wavelength 610 nm and emission wavelength 670 nm. The malachite green stock solution was mixed with RNA top and bottom strands to the final concentration of 2 μM malachite green and 1 μM of each strand of the aptamer in 0.25 M Tris-HCl pH 8.0, 0.15 M NaCl.

In RNA primer extension reactions, the RNA strands were mixed with 2-MeImpG to final concentrations of 50 mM 2-MeImpG and 5 μM each RNA strand (i.e., 5 μM of the RNA complex) in 0.25 M Tris-HCl pH 8.0, 0.15 M NaCl, 50 mM $MgCl_2$. After the reaction, samples were desalted using ZipTip C_{18} pipet tips (Millipore). Samples were eluted from the tip with 1:1 acetonitrile:water, solvent was removed on a SpeedVac vacuum concentrator, and the RNA was dissolved in water. The concentration of the sample was measured using a NanoDrop spectrophotometer, adjusted to 2 μM of the RNA complex, and analyzed as shown above.

Hammerhead Ribozyme. Hammerhead ribozyme reactions were initiated by mixing both RNA strands to a final concentration of 2.5 μM each strand, in 0.25 M Tris-HCl pH 8.0, 0.15 M NaCl, 50 mM $MgCl_2$. Reactions where primer extension was performed contained 50 mM 2-MeImpG. Reactions were incubated at room temperature for 24 h. After the reaction, samples were desalted using ZipTip C_{18} pipet tips (Millipore) and analyzed by TBE−urea 20% PAGE. The reaction yield was calculated as the percentage of the product band intensity relative to the total intensity of all bands on the gel.

Diels−Alderase Catalyzed Reaction. The Diels−Alderase reaction was initiated by mixing 9-hydroxymethylanthracene and *N*-propylmaleimide to final concentrations of 200 μM 9-hydroxymethylanthracene and 1000 μM *N*-propylmaleimide, in100 mM Tris-HCl pH 7.5, 100 mM $MgCl_2$, and 300 mM NaCl. The Diels−Alderase ribozyme-catalyzed reactions were performed in the presence of 10 μM RNA. In RNA primer extension reactions, the RNA strands were mixed with 2-MeImpG to final concentrations of 50 mM 2-MeImpG and 5 μM each RNA strand in 0.25 M Tris-HCl pH 8.0, 0.15 M NaCl, 50 mM $MgCl_2$. Reactions were incubated at room temperature for 24 h. After the reaction, samples were desalted using ZipTip C_{18} pipet tips (Millipore). Samples were eluted from the tip with 1:1 acetonitrile:water, solvent was removed on a SpeedVac vacuum concentrator, and the RNA was dissolved in water. The concentration of the sample was measured using a NanoDrop spectrophotometer, the concentration was adjusted to 10 μM of the RNA complex, and this solution was subjected to the above ribozyme reaction conditions.

The reactions were incubated at room temperature for the reported amount of time and analyzed on an Agilent 1100 Analytical HPLC with a Varian Microsorb-mv 100-8 C_{18} 250 × 4.6 mm column with a gradient of solvent A: 0.1% TFA in H_2O, and solvent: B 0.1% TFA in acetonitrile, with UV detection at 260 nm.

ASSOCIATED CONTENT

Supporting Information

Supplementary discussion and Figures S1−S5. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

szostak@molbio.mgh.harvard.edu

Present Address

[†]K.A.: MIT Media Lab and McGovern Institute, Departments of Biological Engineering and Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors thank Dr. Neha Kamat for helpful discussions and supplying activated nucleotides and members of the Szostak lab for helpful comments on the manuscript. This research was supported in part by a grant from the Simons Collaboration on the Origin of Life to J.W.S. A.E.E. was supported by an appointment to the NASA Postdoctoral Program through NASA Astrobiology, administered by Oak Ridge Associated Universities through a contract with NASA, and a Tosteson Fellowship from the Massachusetts General Hospital Executive Committee on Research. J.W.S. is an Investigator of the Howard Hughes Medical Institute.

DOI: 10.1021/ja511564d
J. Am. Chem. Soc. 2015, 137, 483−489

REFERENCES

(1) Gilbert, W. *Nature* **1986**, *319*, 618.
(2) Orgel, L. E. *Crit. Rev. Biochem. Mol. Biol.* **2004**, *39*, 99.
(3) Robertson, M. P.; Joyce, G. F. *Cold Spring Harb. Perspect. Biol.* **2012**, *4*, No. a003608.
(4) Guerrier-Takada, C.; Gardiner, K.; Marsh, T.; Pace, N.; Altman, S. *Cell* **1983**, *35*, 849.
(5) Zaug, A. J.; Cech, T. R. *Science* **1986**, *231*, 470.
(6) Ban, N.; Nissen, P.; Hansen, J.; Moore, P. B.; Steitz, T. A. *Science* **2000**, *289*, 905.
(7) Wilson, D. S.; Szostak, J. W. *Annu. Rev. Biochem.* **1999**, *68*, 611.
(8) Talini, G.; Gallori, E.; Maurel, M. C. *Res. Microbiol.* **2009**, *160*, 457.
(9) Weimann, B. J.; Lohrmann, R.; Orgel, L. E.; Schneider-Bernloehr, H.; Sulston, J. E. *Science* **1968**, *161*, 387.
(10) Vogel, S. R.; Richert, C. *Chem. Commun. (Camb.)* **2007**, 1896.
(11) Mansy, S. S.; Schrum, J. P.; Tobe, S.; Krishnamurthy, M.; Treco, D. A.; Szostak, J. W. *Nature* **2008**, 122.
(12) Deck, C.; Jauker, M.; Richert, C. *Nat. Chem.* **2011**, *3*, 603.
(13) Adamala, K.; Szostak, J. W. *Science* **2013**, *342*, 1098.
(14) Szostak, J. W. *J. Syst. Chem.* **2012**, *3*, 2.
(15) Hey, M.; Gobel, M. *Methods Mol. Biol.* **2005**, *288*, 305.
(16) Vogel, S. R.; Deck, C.; Richert, C. *Chem. Commun. (Camb.)* **2005**, 4922.
(17) Inoue, T.; Orgel, L. E. *J. Am. Chem. Soc.* **1981**, *103*, 7666.
(18) Green, R.; Szostak, J. W. *Science* **1992**, *258*, 1910.
(19) Zamel, R.; Poon, A.; Jaikaran, D.; Andersen, A.; Olive, J.; De Abreu, D.; Collins, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 1467.
(20) Canny, M. D.; Jucker, F. M.; Kellogg, E.; Khvorova, A.; Jayasena, S. D.; Pardi, A. *J. Am. Chem. Soc.* **2004**, *126*, 10848.
(21) Chadalavada, D. M.; Gratton, E. A.; Bevilacqua, P. C. *Biochemistry* **2010**, *49*, 5321.
(22) Doudna, J. A.; Couture, S.; Szostak, J. W. *Science* **1991**, *251*, 1605.
(23) Hayden, E. J.; Lehman, N. *Chem. Biol.* **2006**, *13*, 909.
(24) Butcher, S. E.; Heckman, J. E.; Burke, J. M. *J. Biol. Chem.* **1995**, *270*, 29648.
(25) Guerrier-Takada, C.; Altman, S. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 1266.
(26) Galloway Salvo, J. L.; Coetzee, T.; Belfort, M. *J. Mol. Biol.* **1990**, *211*, 537.
(27) Jarrell, K. A.; Dietrich, R. C.; Perlman, P. S. *Mol. Cell. Biol.* **1988**, *8*, 2361.
(28) Grate, D.; Wilson, C. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 6131.
(29) Kolpashchikov, D. M. *J. Am. Chem. Soc.* **2005**, *127*, 12442.
(30) Flinders, J.; DeFina, S. C.; Brackett, D. M.; Baugh, C.; Wilson, C.; Dieckmann, T. *ChemBioChem* **2004**, *5*, 62.
(31) Birikh, K. R.; Heaton, P. A.; Eckstein, F. *Eur. J. Biochem.* **1997**, *245*, 1.
(32) Uhlenbeck, O. C. *Nature.* **1987**, *328*, 596.
(33) Breslow, R.; Huang, D. L.; Anslyn, E. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 1746.
(34) Seelig, B.; Jaschke, A. *Chem. Biol.* **1999**, *6*, 167.
(35) Stuhlmann, F.; Jaschke, A. *J. Am. Chem. Soc.* **2002**, *124*, 3238.
(36) Schlatterer, J. C.; Stuhlmann, F.; Jaschke, A. *ChemBioChem* **2003**, *4*, 1089.
(37) Athavale, S. S.; Petrov, A. S.; Hsiao, C.; Watkins, D.; Prickett, C. D.; Gossett, J. J.; Lie, L.; Bowman, J. C.; O'Neill, E.; Bernier, C. R.; Hud, N. V.; Wartell, R. M.; Harvey, S. C.; Williams, L. D. *PloS One* **2012**, *7*, No. e38024.
(38) Isomoto, N.; Maeda, Y.; Tanaka, T.; Furuta, H.; Ikawa, Y. *J. Mol. Evol.* **2013**, *76*, 48.
(39) Whoriskey, S. K.; Usman, N.; Szostak, J. W. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 2465.
(40) Engelhart, A. E.; Powner, M. W.; Szostak, J. W. *Nat. Chem.* **2013**, *5*, 390.
(41) Inoue, T.; Orgel, L. E. *J. Mol. Biol.* **1982**, *162*, 201.
(42) Joyce, G. F.; Inoue, T.; Orgel, L. E. *J. Mol. Biol.* **1984**, *176*, 279.

DOI: 10.1021/ja511564d
J. Am. Chem. Soc. 2015, 137, 483−489

Supplementary Information

for

Generation of functional RNAs from inactive oligonucleotide complexes by non-enzymatic primer extension

Katarzyna Adamala, Aaron E. Engelhart and Jack W. Szostak

Supplementary Discussion

Nonenzymatic primer extension in presence of mixed base monomers

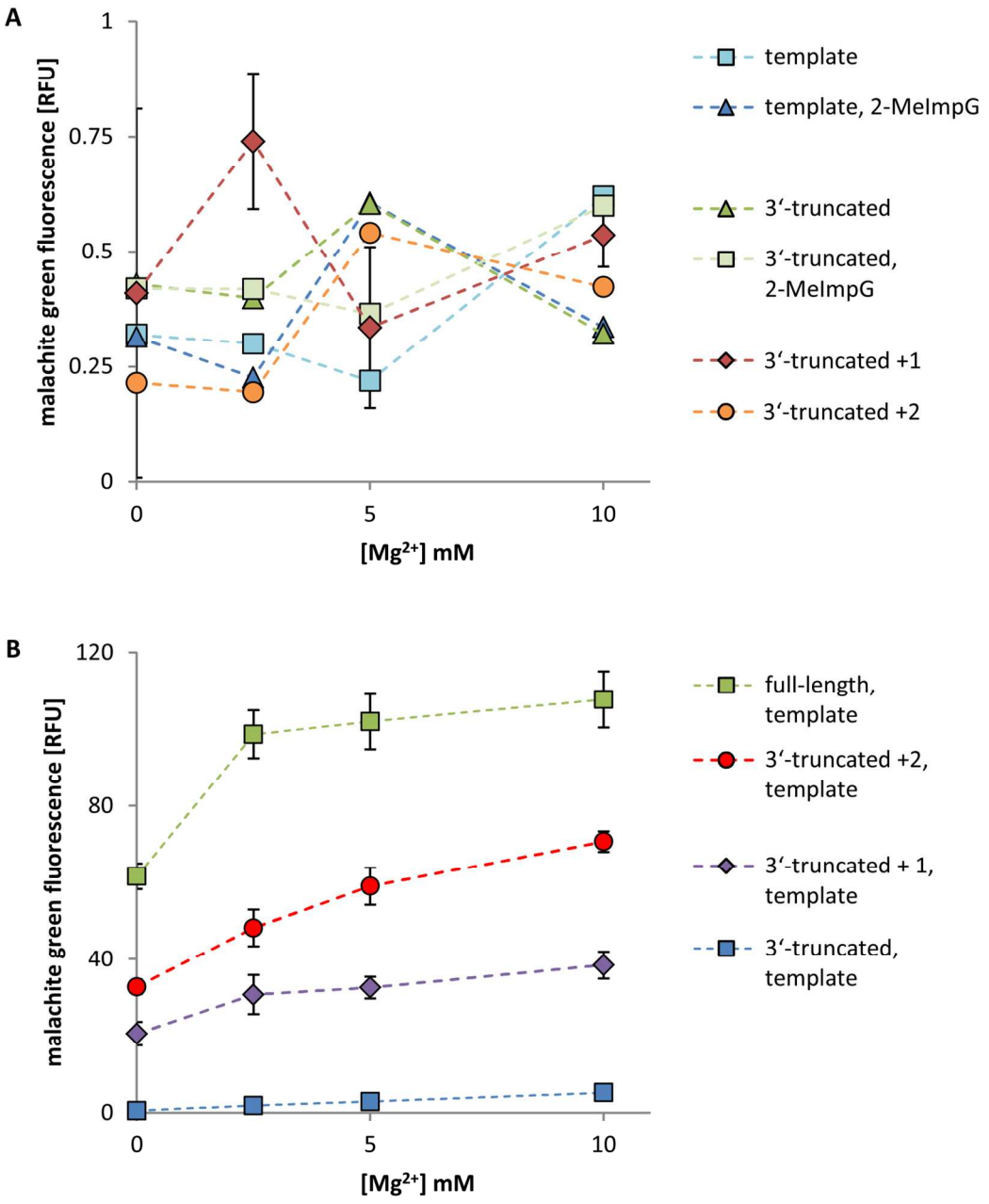
As a means of probing the sensitivity of primer-extension based reconstitution of functional RNA activity from partial fragments to the presence of mismatched bases, we examined the malachite green aptamer and hammerhead ribozyme in the presence of all four activated nucleotides: 2MeImpG, 2MeImpC, 2MeImpU and 2MeImpA. Even in the presence of equimolar amounts of each base, the nonfunctional malachite green aptamer containing a truncated primer-length strand yields active aptamer after a 24 hour primer extension reaction (Figure S3). Similarly, when all four monomers are present, hammerhead ribozyme fragments yield active ribozymes after primer extension, in all permutations of primer and template strands (Figure S4).

RNA stability

We examined the UV-monitored thermal stability of the functional RNAs employed (in both full-length and primer, 3‘-truncated forms) in divalent-ion-free conditions, under which observed changes in absorbance would be principally due to secondary (and not tertiary) structure, and the hammerhead ribozyme would not self-cleave. The thermal denaturation traces of the functional RNAs employed exhibit clear stabilization of secondary structure in their full-length forms, relative to their 3‘-truncated primer forms.

The malachite green aptamer, in its primer form, exhibits two transitions, with midpoints of 21 °C and 45 °C (Fig. S5, open triangles). When the full-length stems expected to be formed by primer extension are present, the aptamer exhibits a single, sharp melting transition, with a midpoint at 50 °C (Fig. S5, closed triangles). The hammerhead ribozyme, in its 3‘-truncated primer form (Fig. S5, open diamonds), is largely unstructured, with a broad melting transition centered at 29 °C. In its full-length form, which is capable of forming two additional stems, this transition is shifted upward by ca 40 °C (Fig. S5, closed diamonds). The Diels-Alderase is known to have a high degree of intramolecular structure and exhibits two melting transitions, with the second incomplete even at 95 °C in either form of this highly structured RNA (Fig. S5, circles). The first melting transition, with an apparent midpoint of 18 °C in the 3‘-truncated form of this RNA (Fig. S5, open circles), increases to 38 °C in the full-length form of this RNA (Fig. S5, closed circles); the second partial melting transition also occurs at higher temperature.

Supplementary Figure S1

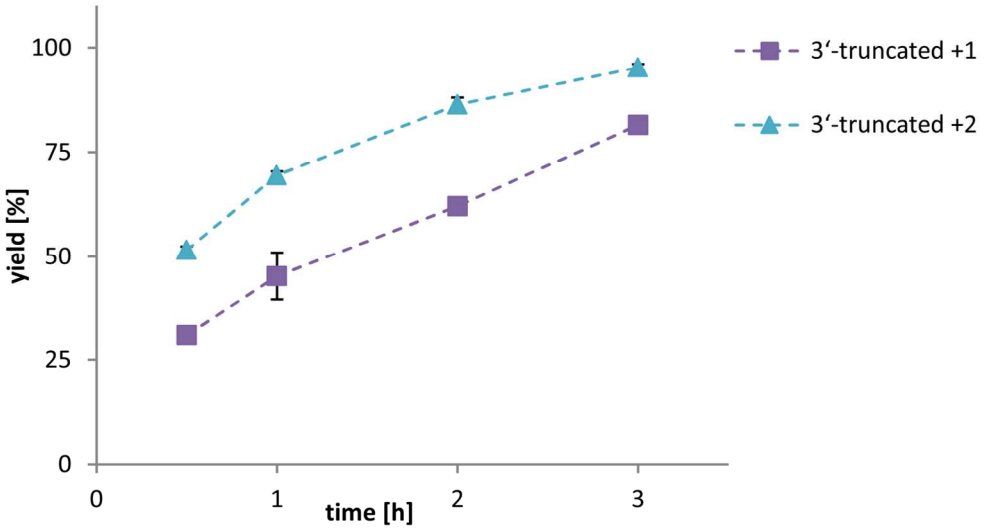


Supplementary Figure S1. Malachite green aptamer.

Both template and full-length 3′-truncated strands are required for maximum malachite green fluorescence, and each successive addition of guanosine residues to the 3′-end of the 3′-truncated strand increases malachite green fluorescence.

A: In the presence of either one of the aptamer strands alone, malachite green fluorescence does not increase, even after incubation with 2-MeImpG. Both primer and template strands are necessary for aptamer reconstitution.
B: In the presence of both 3′-truncated and template strands, malachite green fluorescence increases as the 3′-truncated primer strand is extended. The higher increase of malachite green fluorescence with non-full length 3′-truncated strands (3′-truncated +1 and 3′-truncated +2) at higher magnesium concentrations indicates the importance of duplex stabilization to the activity of the aptamer. Each sample contained 0.25 M Tris-HCl pH 8.0, 0.15 M NaCl, 2 μM malachite green, 1 μM each strand of the aptamer.

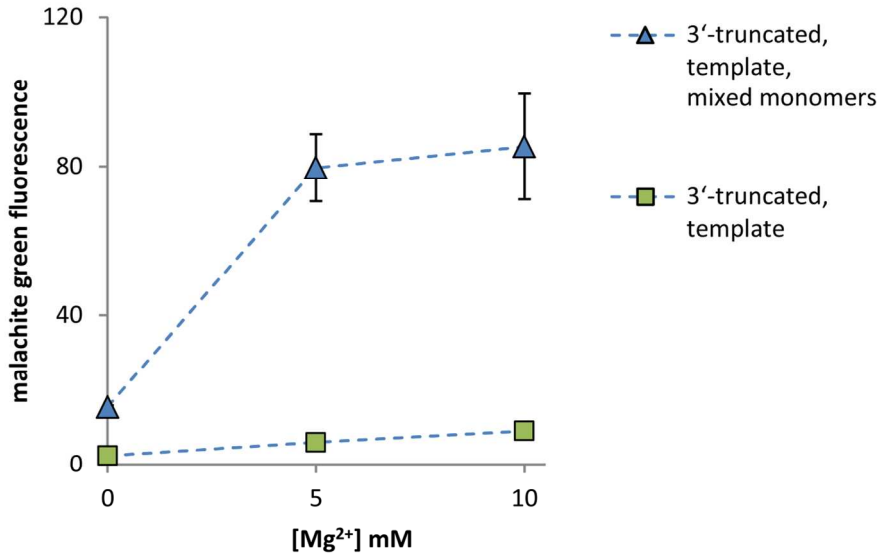
Supplementary Figure S2



Supplementary Figure S2. Yield of Diels-Alder reactions catalyzed by Diels-Alderase ribozymes with differing numbers of G residues at the 3′-end.

Violet squares: Ribozyme RNA corresponding to the addition of one G to the 3′-truncated ribozyme; blue triangles: Ribozyme RNA corresponding to the addition of 2 bases to the 3′-truncated ribozyme. Reaction conditions: 100 mM Tris-HCl pH 7.5, 100 mM $MgCl_2$, 300 mM NaCl, 200 µM 9-hydroxymethylanthracene, 1000 µM N-propylmaleimide, 10 µM RNA. Each experiment was done in duplicate; error bars indicate extreme values.

Supplementary Figure S3



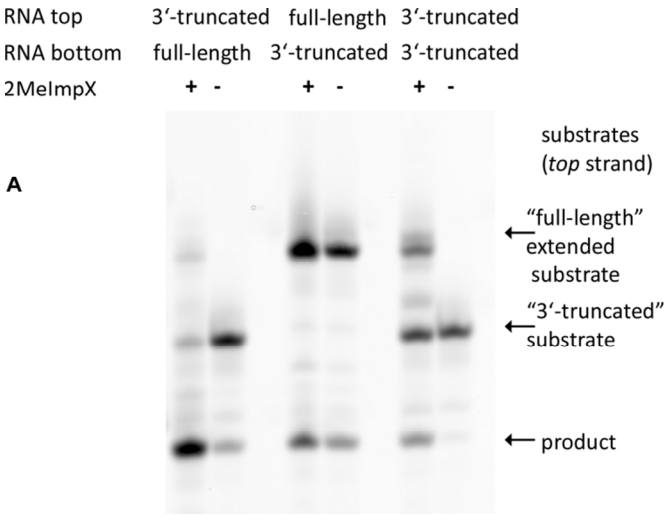
Supplementary Figure S3. The malachite green aptamer with mixed activated RNA monomers

The malachite green 3'-truncated and template strands were subjected to non-enzymatic primer extension in the presence of all four activated nucleotides.

The presence of all four monomers results in an aptamer with comparable activity to that produced in primer extension reactions containing only 2MeImpG, demonstrating that mismatch incorporation is sufficiently low as to not impair the efficacy of this system.

Primer extension using mixed monomers resulted in a 11-fold fluorescence increase of malachite green at 10mM Mg^{2+} (85 RFU after primer extension vs. 8 RFU before), which compares favorably to reactions employing only 2MeImpG, which gave a 13-fold enhancement (107 RFU vs. 8 RFU).

Each primer extension reaction contained 0.25 M Tris-HCl pH 8, 0.15 M NaCl, 5 µM RNA total (2.5 µM each strand), 12.5 mM each of the four activated nucleotides, and 50 mM $MgCl_2$. Primer extension was performed for 24 H. Malachite green fluorescence was examined in reactions containing 0.25 M Tris-HCl pH 8.0, 0.15 M NaCl, 2 µM malachite green, 1 µM each strand of desalted aptamer.



B

Comparison of cleavage yields between reactions containing pure 2MeImpG and those containing a mixture of all 4 RNA monomers (2MeImpX)

RNA top	3′-truncated		full-length		3′-truncated	
RNA bottom	full-length		3′-truncated		3′-truncated	
monomer	+	-	+	-	+	-
2MeImpX	63	20	31	16	23	3
2MeImpG	64	37	33	6	25	5

Supplementary Figure S4. Hammerhead ribozyme with mixed activated RNA monomers

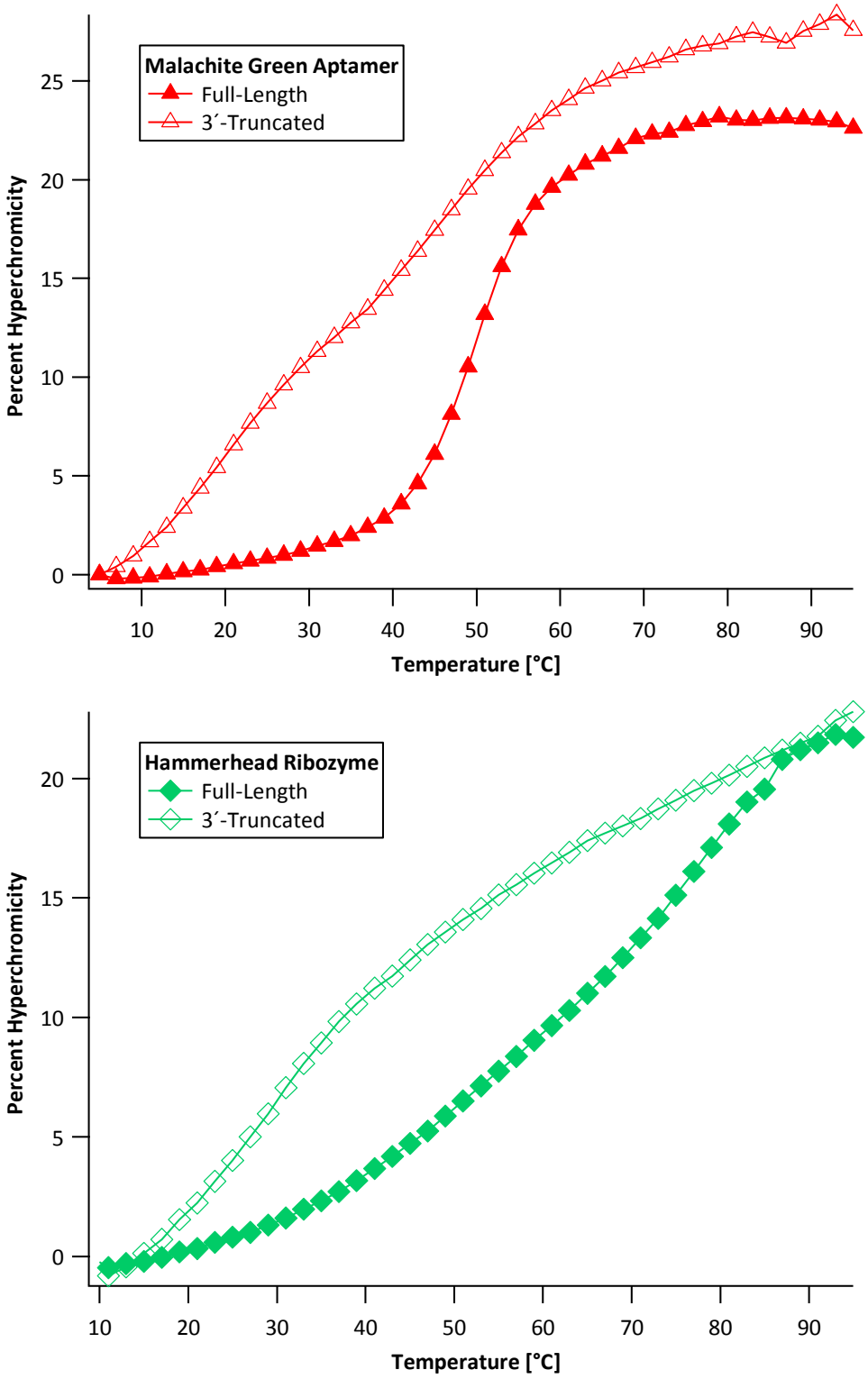
Hammerhead ribozyme 3′-truncated strands were subjected to nonenzymatic primer extension in the presence of all four activated nucleotides.

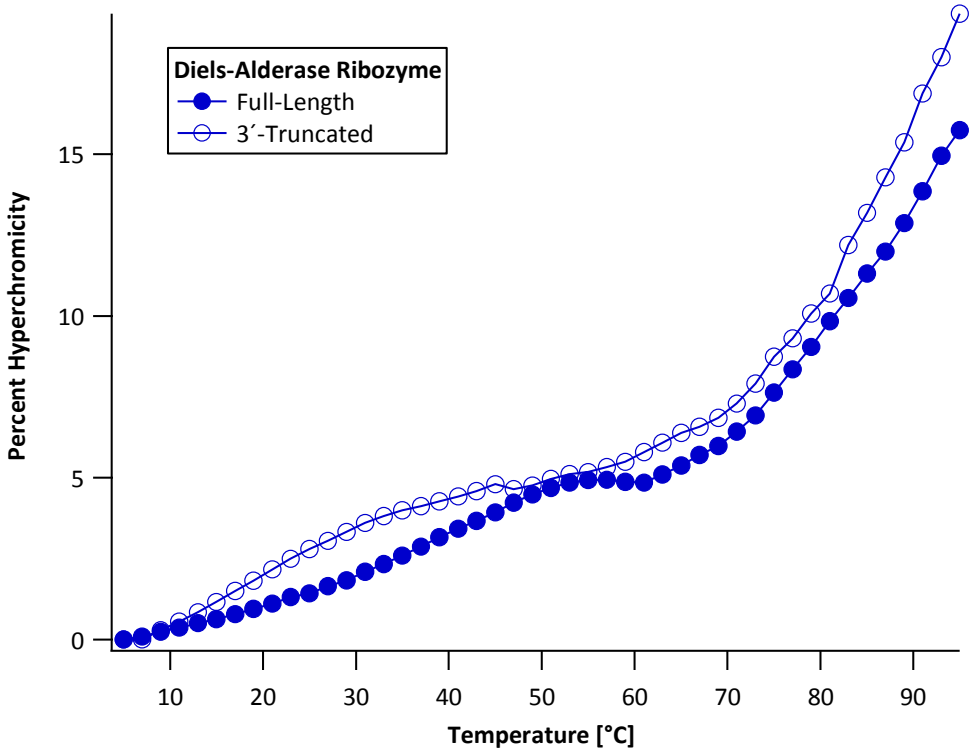
The presence of all four monomers does not inhibit nonenzymatic primer extension or ribozyme activity, demonstrating that mismatch incorporation is sufficiently low as to not impair the efficacy of this system.

Employing both 3′-truncated top and bottom strands (with no activated monomers) resulted in 3% hammerhead cleavage yield after 24 hrs. The addition of all four 2-methylimidazolides (2MeImpX) resulted in 23% cleavage yield, comparable to that achieved in the presence of only 2MeImpG, which gave 29% cleavage yield (Figure 3). Other configurations of strands gave similar yield enhancements. Cleavage yields for reactions with 2MeImpX and those with 2MeImpG (i.e., data from Figure 3) are given in table in panel B.

Each primer extension reaction contained 0.25 M Tris-HCl pH 8, 0.15 M NaCl, 2.5 μM each RNA strand, 12.5 mM each of the four activated nucleotides, and 50 mM $MgCl_2$. Primer extension was performed for 24 hrs.

Supplementary Figure S5





Supplementary Figure S5. Thermal denaturation of full-length and 3′-truncated forms of functional RNAs

All three RNAs studied: malachite green aptamer (red triangles), hammerhead ribozyme (green diamonds), and the Diels-Alderase ribozyme (blue circles) exhibit stabilization of secondary structure in their full-length (filled symbols) forms relative to their primer (3′-truncated) forms (open symbols).

Melt conditions: Malachite green aptamer: Full-length: 5 µM each template and full-length strands, 3′-truncated: 5 µM each template and 3′-truncated strands. Hammerhead: Full-length: 3.75 µM each full-length strand, 3′-truncated: 3.75 µM each 3′-truncated strand. Diels-Alderase: Full-length: 5 µM full-length strand, 3′-truncated: 5 µM 3′-truncated strand. Buffer conditions: 250 mM tris-HCl pH 8, 100 mM NaCl, 1 mM Na-EDTA pH 8. Melts were performed in a 1 mm quartz cuvette (Hellma) using an Agilent Cary 60 interfaced to a Quantum Northwest LC 600. The temperature reported is the block temperature. The background-corrected 260 nm absorbance is reported as percent hyperchromicity at this wavelength relative to the lowest temperature point on the graph. Traces shown are the first cooling trace after a slow heating ramp. All heating and cooling was performed at ≤1°C/min.