

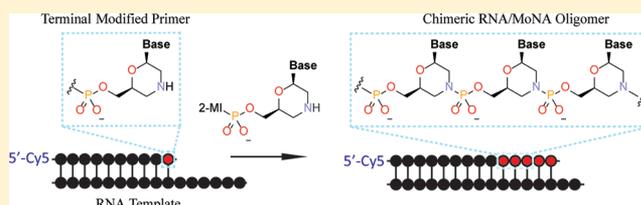
Template-Directed Nonenzymatic Primer Extension Using 2-Methylimidazole-Activated Morpholino Derivatives of Guanosine and Cytidine

Weicheng Zhang,[†] Ayan Pal,^{†,‡} Alonso Ricardo,[§] and Jack W. Szostak^{*†}

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

Supporting Information

ABSTRACT: Efforts to develop self-replicating nucleic acids have led to insights into the origin of life and have also suggested potential pathways to the design of artificial life forms based on non-natural nucleic acids. The template-directed nonenzymatic polymerization of activated ribonucleotide monomers is generally slow because of the relatively weak nucleophilicity of the primer 3'-hydroxyl. To circumvent this problem, several nucleic acids based on amino-sugar nucleotides have been studied, and as expected, the more-nucleophilic amine generally results in faster primer extension. Extending this logic, we have chosen to study morpholino nucleic acid (MoNA), because the secondary amine of the morpholino-nucleotides is expected to be highly nucleophilic. We describe the synthesis of 2-methylimidazole-activated MoNA monomers from their corresponding ribonucleoside 5'-monophosphates and the synthesis of an RNA primer with a terminal MoNA nucleotide. We show that the activated G and C MoNA monomers enable rapid and efficient extension of the morpholino-terminated primer on homopolymeric and mixed-sequenced RNA templates. Our results show that MoNA is a non-natural informational polymer that is worthy of further study as a candidate self-replicating material.



INTRODUCTION

Constructing a model protocell that is capable of Darwinian evolution is one of the fundamental challenges in understanding the transition from chemical to biological evolution in the prebiotic environment of the early Earth.¹ With this goal in mind, significant efforts have been made to develop two mutually compatible and supportive self-replicating systems: membrane compartments to provide spatial localization and a genetic polymer to serve as an informational and functional material. Both systems must replicate without the assistance of any complex protein enzymes or other evolved machinery.^{2–5} Although membranes composed of fatty acids and related amphiphiles have been shown to be a promising basis for self-sustaining and self-replicating vesicles,⁶ a self-replicating genetic polymer is yet to be established. Many lines of evidence suggest that RNA is a reasonable candidate for the primordial genetic system due to its ability to both encode genetic information and to carry out catalytic functions.^{7,8} One possibility is that an initial phase of nonenzymatic chemical replication was sufficient to enable the evolution of the first functional ribozyme, which in turn led to strong selective pressure for increased RNA replication efficiency and, thus, the gradual evolution of ribozyme-catalyzed RNA replication.⁹ Decades of research have established a working model of nonenzymatic template copying involving RNA primer extension with activated nucleoside monophosphates (NMPs).^{5,10} However, despite recent progress, this approach remains limited to the copying of very short mixed sequence

RNA templates, and improving the rate and extent of template copying is still a major challenge.¹¹ To address the question of whether this limitation of RNA is really due to the poor nucleophilicity of the 3'-hydroxyl, one strategy that has been explored is the use of nucleotide analogs that position a stronger amine nucleophile at the terminus of the growing primer chain. Such model systems may provide a means of studying aspects of nucleic acid replication that cannot currently be addressed with RNA, such as the copying of structured templates and the difficulty of achieving cycles of replication.

Orgel and colleagues first showed that replacing the 3'-hydroxyl of uridine with a more nucleophilic 3'-amine enhances, by 1–2 orders of magnitude, the nontemplated solution phase formation of dinucleotides resulting from reaction with ribonucleoside-5'-phosphorimidazolides. They also observed that 3'-amino-3'-deoxyguanosine-5'-phosphorimidazolide condenses on both poly(C) and poly(dC) templates, whereas the guanosine-5'-phosphorimidazolide does not condense efficiently under similar conditions.^{12,13} More recently, we have found that 3'-amino-2',3'-dideoxyribonucleoside-5'-phosphor-2-methylimidazolides polymerize efficiently on RNA templates with the catalytic assistance of 1-(2-hydroxyethyl)imidazole (HEI).¹⁴ However, in the presence of HEI, the rate of intramolecular monomer cyclization to

Received: June 17, 2019

Published: July 12, 2019

yield unreactive 3'-5' cyclized monomers as a side product also increases dramatically.^{14,15} Imidazole-activated 2'-amino-2',3'-dideoxy-guanosine and -cytidine are both efficiently incorporated in template-copying reactions, while intramolecular cyclization is prevented due to ring strain.¹⁶ Template-directed polymerization using an amine as the nucleophile at the growing end of the primer has also been observed for the npGNA¹⁷ and 2'-NH₂-TNA¹⁸ systems. Activated 2'-amino GNA dinucleotides can be used in nonenzymatic GNA template-directed ligation to yield npGNA oligonucleotides; however, the addition of monomers is unsuccessful due to rapid intramolecular cyclization.¹⁷

Encouraged by previous results with nucleophilic amines, we have now investigated the use of morpholino nucleic acids (MoNA) in RNA template-directed nonenzymatic primer extension (Figure 1). MoNA is an interesting candidate genetic

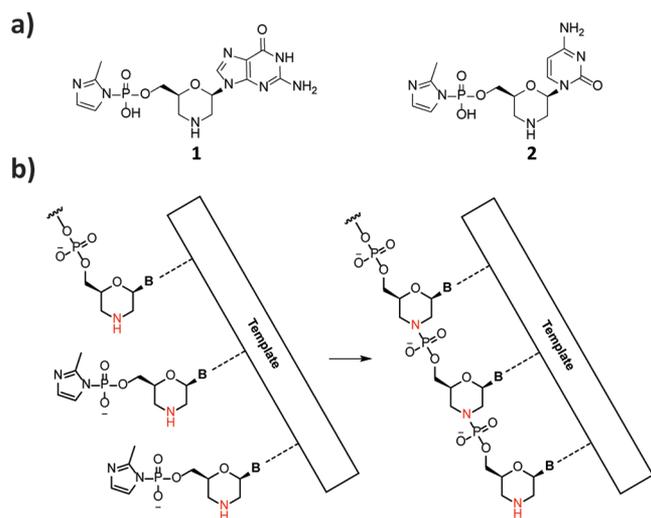


Figure 1. Morpholino nucleic acid (MoNA) system. (a) Structure of activated MoNA monomers: **1**, 5'-phosphor-2-methylimidazolide of morpholino guanosine (2-MeImpmG); **2**, 5'-phosphor-2-methylimidazolide of morpholino cytidine (2-MeImpmC). (b) Schematic representation of RNA template-directed primer extension of a morpholino-terminated primer in the presence of activated MoNA monomers. The attacking nucleophile is the morpholino amine highlighted in red.

polymer because (a) the secondary morpholino amine should serve as a stronger nucleophile than a primary amine in water,¹⁹ (b) molecular modeling suggests that morpholino oligos in which the morpholine rings are in a chair conformation should form an A-type heteroduplex with an RNA oligonucleotide that is structurally similar to an RNA–RNA duplex,²⁰ and (c) the chair conformation of the morpholine ring should prevent cyclization of the activated morpholino monomers. The chair conformation of the morpholine ring is supported by proton NMR analysis of a carbamate-linked morpholino oligonucleotide.²¹ The A-form geometry of the morpholino oligo is important because it is the preferred geometry for template-directed nonenzymatic extension by activated monomers, since the A-form conformation is suitable for an in-line attack by the nucleophile of the growing primer on the phosphate of the incoming activated monomer.

Here we present the synthesis of 2-methylimidazole-activated morpholino guanosine (2-MeImpmG) and cytidine

(2-MeImpmC) and the synthesis of RNA primers terminating in a morpholino-nucleotide. The monomers do not cyclize, but at neutral pH and high monomer concentration, nontemplated oligomerization is significant. We show that the extension of an RNA primer terminated with a morpholino-nucleotide using low concentrations (5 mM) of 2-MeImpmG or 2-MeImpmC proceeds very rapidly on RNA templates.

RESULTS

We first developed a two-step synthesis of 2-methylimidazole-activated 5'-phosphoro-MoNA monomers. We began by converting NMPs to their corresponding 2-methylimidazole-activated derivatives following a standard and well-established protocol.²² Next, the ribose sugar was converted to a morpholine ring by oxidative opening of the five-membered ribose ring with periodate. The resulting dialdehyde was treated with ammonium tetraborate to give a six-membered morpholine ring, followed by reductive removal of the original 2'- and 3'-hydroxyls with cyanoborohydride (Figure 2a).

One of the limiting factors for the extent of template copying in a closed reaction system is the longevity of chemically activated mononucleotides. 2-Methylimidazole-activated monomers are known to be prone to hydrolysis and/or cyclization under template-copying conditions. For example, in the presence of 100 mM Mg²⁺, the rate of hydrolysis of 2-MeImpmG at pH 8 is $\sim 0.016 \text{ h}^{-1}$, corresponding to a half-life of 43 h.²³ 3'-NH₂-2-MeImpmC, under its optimal primer extension conditions in the presence of 100 mM HEI, undergoes cyclization at a rate of 0.53 h^{-1} , corresponding to a half-life of only 1.3 h.¹⁴ Such undesired side reactions greatly impair the efficiency and sustainability of template copying in these systems. We therefore monitored the stability of the activated MoNA monomers at circumneutral pH by ³¹P NMR spectroscopy. Neither cyclization nor hydrolysis products were observed, even after 1 day. Instead, activated MoNA monomers were found to undergo nontemplated polymerization to yield activated oligomers (Figure 2b). The accumulation of various activated oligomers and corresponding loss of activated monomer was evident from the decay of the ³¹P signal related to the phosphoramidate linkage and the corresponding growth of the ³¹P resonances due to internal phosphodiester linkages (Figure 2c). In the case of 2-MeImpmG at pH 7.5, the rate of decline of the ³¹P phosphoramidate resonance was 0.037 h^{-1} , corresponding to a half-life of 19 h (Figure 2d). These results were confirmed by liquid chromatography–mass spectrometry (LCMS) studies. The absence of cyclized monomer is likely related to the conformational constraints imposed by the morpholine ring. In order to directly determine the conformation of the morpholine ring in morpholino G and C nucleosides, we obtained crystals of both nucleosides and determined their molecular structures by X-ray diffraction crystallography. The morpholine ring is in the expected chair conformation in both nucleosides (Figure 3b).

Because the spontaneous oligomerization of the activated MoNA monomers to varying extents would complicate the interpretation of template-directed primer extension studies, we searched for conditions in which monomers could be stored without the formation of unwanted side products. Since all the side reactions of concern, i.e., cyclization, hydrolysis, and oligomerization, require a protonated 2-methylimidazole leaving group, we reasoned that basic pH would be crucial for stability. As expected, we found that activated MoNA

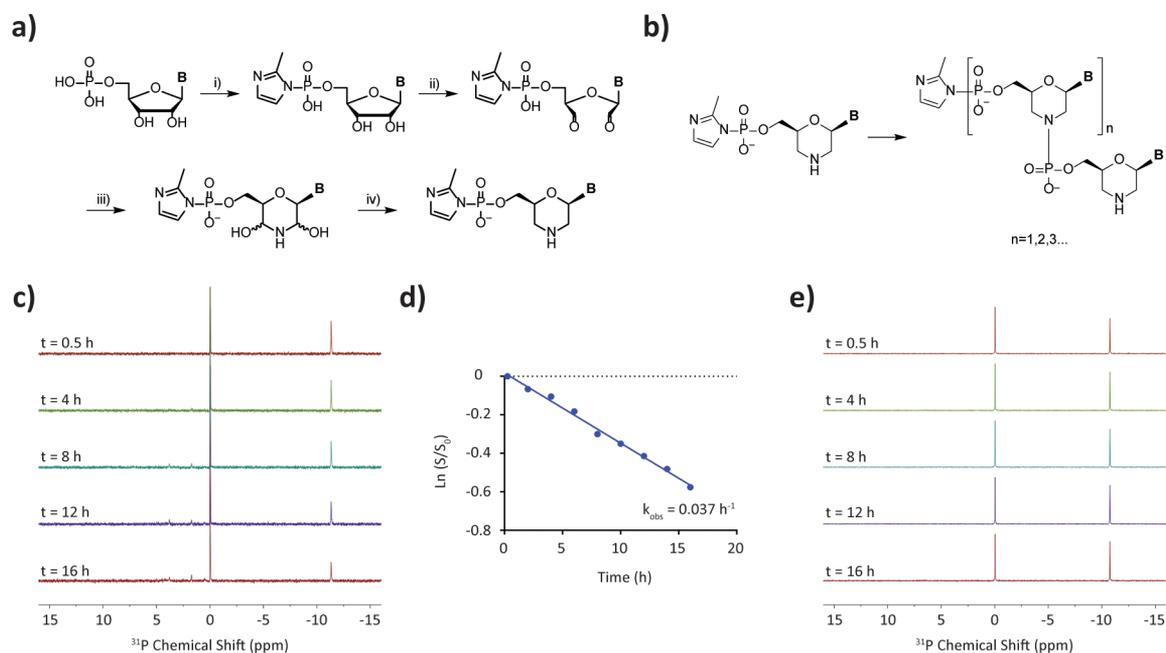


Figure 2. Preparation of activated MoNA monomers and NMR studies of the decay of 2-MeImpmG. (a) Synthesis of 2-methylimidazole-activated MoNA monomers. Reaction conditions: (i) 2-methylimidazole, DPDS, TPP, TEA, DMSO, 20 °C, 2.5 h; (ii and iii) NaIO₄, (NH₄)₂B₄O₇, MeOH, 20 °C, 0.5 h; (iv) NaBH₃CN, 20 °C, 3 h. Reactions were performed in one pot for steps ii–iv. For details see the [Supporting Information \(SI\)](#). (b) Schematic representation of the nontemplated polymerization of activated MoNA monomers to yield activated oligomers. (c) Activated 2-MeImpmG monitored at $\delta = -11.3$ ppm by ³¹P NMR spectroscopy at different time points. Phosphorus signals corresponding to internal phosphodiester linkages were observed between $\delta = 0$ and 5 ppm. Phosphate buffer ($\delta = 0$ ppm) was used as an external reference. Reactions contained 10 mM 2-MeImpmG in 200 mM HEPES pH 7.5 and 15% D₂O. (d) Rate of decline of the ³¹P phosphoramidate resonance from a plot of $\ln(S/S_0)$ vs time. Solid line is linear fit with $R^2 > 0.99$ and $k_{\text{obs}} = 0.037 \text{ h}^{-1}$. (e) Stability of 2-MeImpmG at pH 12. Activated 2-MeImpmG monitored at $\delta = -10.8$ ppm by ³¹P NMR spectroscopy at different time points. Phosphate buffer ($\delta = 0$ ppm) was used as an external reference. Solution contained 15 mM 2-MeImpmG in H₂O with 15% D₂O.

monomers are stable in the pH range from 10 to 12. For example, no change was observed in the ³¹P NMR spectrum of a 10 mM solution of 2-MeImpmG at pH 12 after 16 h ([Figure 2e](#)).

The morpholino-terminated primer was prepared following a similar strategy to that previously described by Eisenhuth and Richert for the synthesis of a 3'-amino-terminated DNA primer.²⁴ A protected 5'-DMTr-morpholino residue was incubated with long-chain-alkylamine controlled-pore glass (LCAA-CPG) previously treated with hexafluoroglutaric anhydride to prepare an appropriate solid support, which was subsequently used for standard solid-phase oligonucleotide synthesis to yield fluorescently labeled chimeric RNA/MoNA oligonucleotides [[Figures 3a](#) and [S1 \(SI\)](#)]. Those oligonucleotides were used as primers for subsequent template-directed primer extension reactions with activated morpholino-nucleotides and ribonucleotides.

We initially compared primer extension reactions using primers ending in either a morpholino-nucleotide or standard ribonucleotide, using the chimeric RNA/MoNA primer **P1** (5'-Cy5-GACUGACUGmG-3') and the corresponding all RNA primer **P2** (5'-Cy5-GACUGACUGG-3'). Reactions were carried out on the C₄ RNA template **T1** in the presence of guanosine 5'-phosphoro-2-methylimidazolide (2-MeImpG). In the absence of divalent cations, primer **P1** yielded >95% of the *N* + 1 product in 5 min ([Figure S2b](#), [SI](#)), while when the same reaction was performed with **P2**, >98% of the primer remained unreacted, even after 24 h. In the presence of 100 mM Mg²⁺, ca. 81% of **P2** extended to various lengths after 1 h ([Figure S2d](#), [SI](#)). The morpholino-terminated primer is clearly

much more reactive than the all-RNA primer, even in the presence of Mg²⁺, presumably due to the greater nucleophilicity of the morpholino amine than the 3'-hydroxyl of ribose. To compare the reactivity of the morpholino secondary amine with that of a primary amine, we used two primers, one ending in a morpholino-nucleotide and the other ending in a 3'-amino-2',3'-dideoxyribonucleoside (3'-NP-DNA). For this comparison, we used the chimeric RNA/MoNA primer **P3** (5'-Cy3-AGUGAGUAACGmC-3') and the 3'-amino-2',3'-dideoxy C-terminated primer **P4** (5'-Cy3-AGUGAGUAACGC_{NH₂}-3') on the C₄ RNA template **T2** in the presence of 2-MeImpG. Primer extension of primer **P3** yielded 93% of the *N* + 1 product in 5 min, while primer **P4** yielded only 42% of the *N* + 1 product in the same time ([Figure 3d](#)). Kinetic analysis of the two reactions showed that extension of the morpholino-terminated primer **P3** was ca. 4-fold faster than for the 3'-amino-terminated primer **P4** ([Figure 3e](#)). Similar results were observed when both primers were tested on a G₄ RNA template in the presence of cytidine 5'-phosphoro-2-methylimidazolide (2-MeImpC) ([Figure S3](#), [SI](#)). These results are consistent with the hypothesis that the morpholino amine of MoNA nucleotides is a better nucleophile than both the secondary alcohol of RNA nucleotides and the primary amine of 3'-NP-DNA nucleotides, in the context of primer extension reactions.

Encouraged by the rapid and complete template-directed reaction of a primer ending in a morpholino-nucleotide with an activated ribonucleotide, we proceeded to examine primer extension using activated morpholino-nucleotides. We first considered the extension of primer **P1** on the homopolymeric

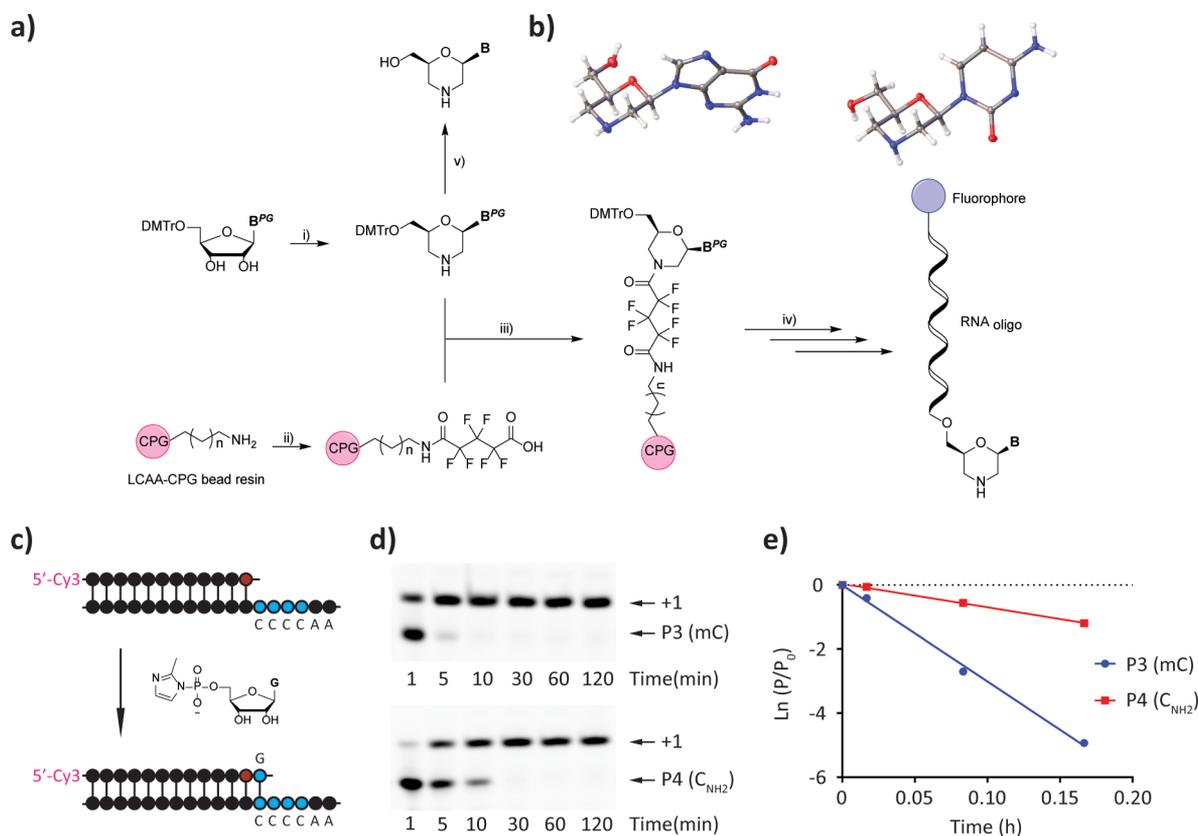


Figure 3. Preparation of fluorophore-labeled 3'-morpholino-terminated RNA oligonucleotide and nonenzymatic primer-extension reactions with 3'-terminal modified primers. (a) Preparation of fluorophore-labeled 3'-morpholino-terminated RNA oligonucleotide. Reaction conditions: (i) (A) $(\text{NH}_4)_2\text{B}_4\text{O}_7$, NaIO_4 , MeOH, 20 °C, 0.5 h, (B) NaBH_3CN , 20 °C, 2 h; (ii) hexafluoroglutaric anhydride, DIEA, DCM, 20 °C, 16 h; (iii) PPh_3 , DIEA, I_2 , DMF, 20 °C, 16 h; (iv) standard RNA synthesis followed by standard deprotection procedure; (v) (A) 50% CH_3COOH , 20 °C, 2 h, (B) $\text{NH}_4\text{OH}/\text{CH}_3\text{NH}_2$, 20 °C, 16 h. For details see the [Supporting Information](#) (SI). (b) Crystal structures of morpholino nucleosides with a guanine (left) and cytosine (right) nucleobase. (c) Schematic representation of morpholino C-terminated primer P3 or the 3'-amino-2',3'-dideoxy C-terminated primer P4 annealed to a C₄ RNA template (T2) in the presence of 2-MeImpG. In the absence of Mg^{2+} , the monomers participate in primer extension on the corresponding template region, resulting in predominantly $N + 1$ product. (d) High-resolution PAGE analysis of the primer-extension products at different reaction times. Reaction conditions: 1 μM primer P3 (top) or P4 (bottom), 5 μM C₄ template T2, 300 mM HEPES pH 7.5, 5 mM 2-MeImpG. (e) Determination of the rates of the primer extension from the morpholino-terminated primer P3 and 3'-amino-terminated primer P4 via the plot of $\ln(P/P_0)$ vs time. Solid line is a linear fit with $R^2 > 0.99$ and k_{obs} of 30 h^{-1} with P3 and 7.3 h^{-1} with P4.

C₄ and G₄RNA templates (T1 and T3, respectively) (Figure 4). Polyacrylamide gel electrophoresis analysis of the reaction of the P1/T1 complex in the presence of 5 mM 2-MeImpmG showed that >99% of the primer was consumed within 1 min, resulting in mostly full-length $N + 4$ product with some $N + 3$ product (Figure 4b). A minor amount of overhanging $N + 5$ product accumulated over 1 h. Similar results were obtained when 5 mM 2-MeImpmC was added to the P1/T3 complex (Figure 4e). After 1 min, the reaction was essentially complete with >98% of the primer converted to full length products ($\geq +4$ nts). The identities of the extended products were confirmed by LCMS [Figures 4c,f and S4 (SI)]. When primer extension reactions were performed in the absence of template or in the presence of noncomplementary templates, only traces of $N + 1$ product formed after 1 h. Thus, nontemplated primer extension is ca. 1000-fold slower than template-directed copying (Figure S5, SI). The rate of template copying did not vary noticeably between the primer terminated with mG (P1) and the primer terminated with mC (P3). Remarkably, the latter reaction was at least 1 order of magnitude faster than primer extension of the corresponding 3'-amino-2',3'-dideoxy C-terminated primer P4 on the G₄ RNA template T4 in the presence of 3'-NH₂-2-MeImpddC (Figure S6, SI).

We examined primer extension reactions with activated morpholino monomers under a range of pH and metal ion conditions, to assess the potential compatibility of this system with model protocell membranes. Primer extension reactions with primer P1 on templates T1 and T3 in the presence of 2-MeImpmG and 2-MeImpmC, respectively, were indistinguishably rapid in the pH range from 6 to 8.5. At pH >8.5, the reaction slowed considerably (Figure S7, SI). As previously observed with amino-sugar nucleotides, the presence of divalent cations such as Mg^{2+} and Mn^{2+} had minimal effects on the kinetics of the primer extension reactions. This observation suggests that template copying in the morpholino system should be compatible with primitive fatty acid-based vesicles, which are destabilized by high concentrations of divalent cations.¹¹

The pattern of primer extension products obtained with the 3'-morpholino-terminated primer and morpholino-nucleotide monomers is strikingly different from that typically observed for an all-RNA system. The template-directed addition of nucleotides to an all-RNA primer occurs in a stepwise manner, with a distribution of +1 to +3 products observed at early time points and longer addition products accumulating at later times. In contrast, the addition of morpholino-nucleotides to

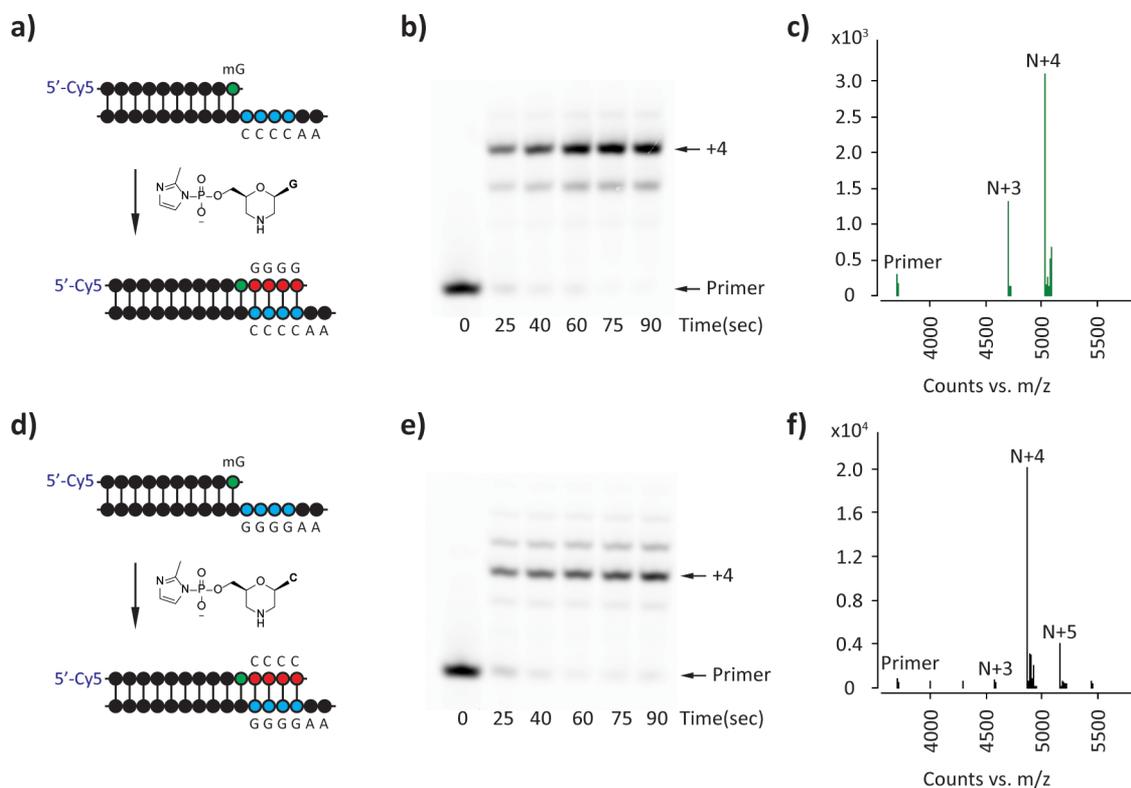


Figure 4. Nonenzymatic primer extension reactions on homopolymeric RNA templates using morpholino monomers 2-MeImpmG and 2-MeImpmC. (a) Schematic representation of primer P1 annealed to a C₄ template (T1) in the presence of 2-MeImpmG. Monomers participate in primer extension on the template, resulting in the synthesis of chimeric RNA/MoNA oligomers. (b) High-resolution PAGE analysis of the primer extension products at increasing reaction times. Reaction conditions: 1 μ M primer P1, 5 μ M C₄ template T1, 300 mM HEPES pH 7.5, 5 mM 2-MeImpmG. (c) High-resolution MS analysis of the primer extension products formed after a 1 day incubation. Reaction conditions: 2.75 μ M primer P1, 10 μ M template T1, 200 mM HEPES pH 7.5, 5 mM 2-MeImpmG. (d) Schematic representation of primer P1 annealed to a G₄ template (T3) in the presence of 2-MeImpmC. (e) High-resolution PAGE analysis of the primer extension products at increasing reaction times. Reaction conditions: 1 μ M primer P1, 5 μ M G₄ template T3, 300 mM HEPES pH 7.5, 5 mM 2-MeImpmC. (f) High-resolution MS analysis of the primer extension products formed after a 1 day incubation. Reaction conditions: 2.75 μ M primer P1, 10 μ M template T3, 200 mM HEPES pH 7.5, 5 mM 2-MeImpmC.

the 3'-morpholino-terminated primer appears to jump rapidly to full length products, with almost no +1 and +2 products observed, even at early times when only a fraction of the primer has been converted to +4 full length product. One hypothesis that could explain this observation is that the addition of the first (and possibly second) morpholino-nucleotides to a morpholino-terminated primer is significantly slower than subsequent additions. This could be caused by a transition in the duplex geometry between that of the RNA:RNA homoduplex to that of the MoNA:RNA heteroduplex. To test this hypothesis, we purified an RNA primer with three terminal morpholino-nucleotides P5 (5'-Cy5-GACUGACUGmGmGmG-3') from a primer extension reaction of P1 with 2-MeImpmG. Primer extension using P5 as the primer on a C₄ template with morpholino-G monomers showed a more typical stepwise addition pattern (Figure S8, SI), consistent with the above hypothesis.

Although homopolymeric templates have been traditionally used to investigate the efficiency of individual activated nucleotides, primer extension reactions on mixed sequence templates are more relevant to the ultimate goal of replicating functional sequences. We therefore prepared four mixed sequence C+G templates with template regions—3'-GCC-5', 3'-CGGG-5', 3'-GGGC-5', and 3'-CCCG-5' (T5, T6, T7, and T8 respectively)—and studied the appearance of primer

extension products in the presence of 5 mM each of 2-MeImpmG and 2-MeImpmC ([Figures 5 and S9–S11 (SI)]. PAGE analysis showed that in all the cases the primers were consumed within 30 min, yielding predominantly N + 3 and N + 4 products. Primer extension on the templates that contained a different nucleotide at the last position of the templating region (T7 and T8) took longer to complete the addition of the last nucleotide. The products of these reactions were further analyzed by LCMS [Figures 5c and S9c, S10c, and S11c (SI)]. The extracted ion chromatograms confirmed that in all cases, the major products were the full-length correct sequences. The presence of the correct +1 to +3 products in the extracted ion chromatograms is consistent with the sequential addition of monomers to the growing primer. Only traces of sequences with a single incorrect base at the terminus were seen, suggesting not only that misincorporations were rare but that continued primer extension following a mismatch was very slow, as previously observed for a DNA system with 3'-amino-terminated DNA primers.²⁵ Surprisingly, the rates of primer extension on the mixed sequence templates T5 and T6 were slower than that on homopolymeric templates T1 and T3. The reasons for the slow copying of these mixed-sequence templates are unclear and are the subject of ongoing study.

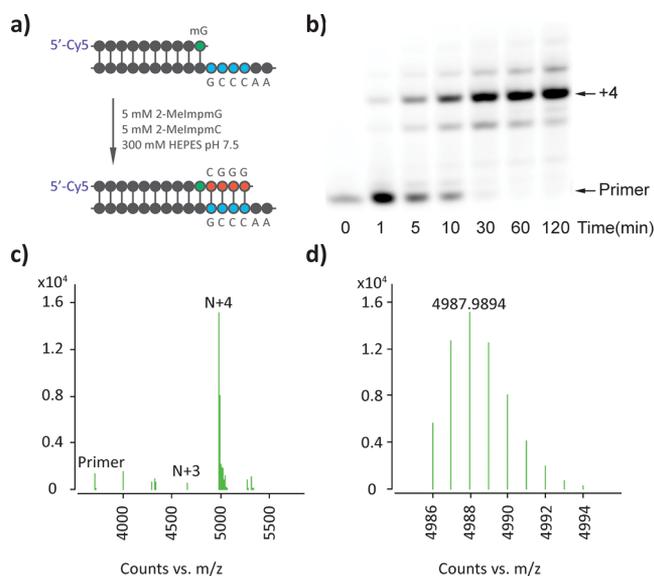


Figure 5. Nonenzymatic primer-extension reactions on a mixed-sequence RNA template. (a) Schematic representation of primer P1 annealed to template (T5) in the presence of 2-MeImpmG and 2-MeImpmC, followed by primer extension. (b) High-resolution PAGE analysis of the primer-extension products at increasing reaction times. Reaction conditions: 1 μ M primer P1, 5 μ M template T5 (template region 3'-GCCC-5'), 300 mM HEPES pH 7.5, 5 mM 2-MeImpmG, 5 mM 2-MeImpmC. (c) High-resolution MS analysis of the primer-extension products after 1 day of reaction. Reaction conditions: 2.75 μ M primer P1, 10 μ M template T5, 200 mM HEPES pH 7.5, 3 mM 2-MeImpmG, 3 mM 2-MeImpmC. (d) Monoisotopic mass for the full-length primer +4 product: calculated mass, 4986.0102 Da; observed mass, 4985.9961 Da; error, 2.8 ppm.

Template-directed RNA polymerization with 2-methyl- or 2-aminoimidazole-activated ribonucleotide monomers proceeds through the formation of a highly reactive imidazolium-bridged dinucleotide intermediate.²⁶ The intermediate binds to the template, and in the bound state, its conformation is preorganized so as to favor reaction with the 3'-hydroxyl of the primer. In principle, two possible mechanisms could account for primer extension with a morpholino-terminated primer. The high nucleophilicity of the morpholino secondary amine could potentially allow for direct monomer addition, or alternatively, primer extension might still proceed largely or entirely through formation of an imidazolium-bridged dinucleotide intermediate, which then reacts rapidly with the morpholino-terminated primer. To distinguish between these two mechanisms, we used the mixed sequence template with template region 3'-CGGG-5' (T6) and studied primer extension using primer P1 in the presence of 2-MeImpmG and different downstream monomers (Figure 6). PAGE analysis suggested that the rates of primer extension with activated mG were similarly slow in the absence of any downstream monomer or when unactivated morpholino-cytidine monophosphate (mCMP) or unactivated ribocytidine monophosphate (rCMP) was present. In all three cases, only a small fraction of primer was converted to +1 products in 10 min, and >20% of primer remained unreacted even after 24 h. In contrast, the addition of 2-MeImpmC resulted in a dramatic increase in the rate of primer extension, such that >90% of the primer was converted to full length products (\geq +4 nts) in 10 min. This observation strongly suggests that template copying in the MoNA system proceeds via the formation of an

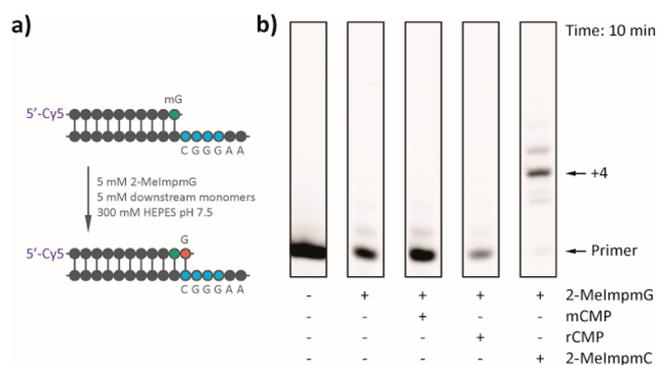


Figure 6. Fast primer extension with a morpholino monomer requires an activated downstream monomer. (a) Schematic representation of primer P1 annealed to template (T6) in the presence of 2-MeImpmG and various downstream monomers. (b) High-resolution PAGE analysis of the primer-extension products when different downstream monomers are used. Left to right: lane 1, primer only; lane 2, activated mG only, no downstream monomer; lane 3, activated mG and unactivated mC; lane 4, activated mG and unactivated rC; lane 5, activated mG and activated mC. The nucleotides added to each reaction are indicated below each lane. Reaction conditions: 1 μ M primer P1, 5 μ M template T6, 300 mM HEPES pH 7.5, 5 mM 2-MeImpmG, and 5 mM of downstream nucleotides. All reactions were quenched with urea at 10 min.

imidazolium-bridged reactive intermediate, as previously observed for the all-RNA system.

DISCUSSION

Morpholino-nucleotides are readily accessible through oxidation of the canonical ribonucleotides followed by reductive amination. Surprisingly, the corresponding morpholino oligonucleotides have not been characterized, although a neutral-backbone diamidate variant that is the basis of a recently approved therapeutic oligonucleotide has been studied in detail.^{27–29} Our work provides a novel synthetic route to the charged backbone morpholino oligonucleotides, through the nonenzymatic copying of RNA templates with activated morpholino-nucleotide monomers. In the course of this work, we have found that the 2-methylimidazole-activated morpholino G and C nucleotides add very rapidly to a morpholino-terminated primer across from both homopolymeric and heteropolymeric RNA templates. Side-by-side comparisons reveal that the morpholino copying reactions are considerably faster than the corresponding reactions with 3'-amino-2',3'-dideoxy-nucleotide-terminated primer and monomers, which in turn are faster than an otherwise identical but all-RNA system. These results are consistent with the expected greater nucleophilicity in water of the morpholino secondary amine vs the primary amine of 3'-amino-2',3'-dideoxy nucleotides and the 3'-hydroxyl of ribonucleotides. However, there are clearly significant structural differences between the relatively rigid chair conformation of the morpholine ring, compared with the flexible five-membered sugar ring of the 3'-amino or ribonucleotide sugars, and these structural differences are also likely to affect the observed rates of primer extension reactions. Nevertheless, an important implication of our observations is that a means of increasing the nucleophilicity of the ribose 3'-hydroxyl, for example by correctly positioning a nearby metal ion, might lead to enhanced rates of nonenzymatic primer extension in an all-RNA system.

The rapid rates of template copying, together with the very slow rates of hydrolysis and cyclization of the activated morpholino monomers, suggest that morpholino nucleic acids might serve as the basis for a nonenzymatically replicating genetic system. Furthermore, the fact that the template-copying reaction proceeds without the assistance of divalent cations and over a wide range of pH values suggests that this genetic polymer should be compatible with replicating compartment systems, such as those formed by fatty acid vesicles. To develop the MoNA system as a general genetic polymer, it will be necessary to demonstrate the copying of templates that contain all four nucleotides and a means of copying the copies so that indefinite cycles of replication can be attained. If a replicating MoNA system can be developed, it may serve as the basis for an alternative protocell design that is not derived from either biology or prebiotic chemistry.

■ ASSOCIATED CONTENT

■ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.9b06453.

X-ray crystallographic files in CIF format for morpholino guanosine (CIF)

X-ray crystallographic files in CIF format for morpholino cytidine (CIF)

All experimental materials, methods, supplementary figures S1–S13, and additional references (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*szostak@molbio.mgh.harvard.edu

ORCID

Weicheng Zhang: 0000-0003-3900-102X

Jack W. Szostak: 0000-0003-4131-1203

Present Addresses

[‡]A.P.: Dana-Farber Cancer Institute, 450 Brookline Avenue, Boston, Massachusetts 02215, United States

[§]A.R.: Ra Pharmaceuticals, 87 Cambridge Park Drive, Cambridge, MA 02140, United States

Author Contributions

[†]W.Z. and A.P. contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Dr. Shao-Liang Zheng for his help with the X-ray data collection and structure determination. We thank Dr. Aaron Larsen for assistance with the synthesis of the morpholino-nucleotide-derivatized solid support. We thank Dr. Anders Björkbo for help with mass spectrometry, and Drs. Derek O'Flaherty and Lijun Zhou for helpful discussions and comments on the manuscript. J.W.S. is an Investigator of the Howard Hughes Medical Institute. This work was supported in part by grants from the National Science Foundation (CHE-1607034) and the Simons Foundation (290363) to J.W.S.

■ REFERENCES

(1) Szostak, J. W. The Narrow Road to the Deep Past: In Search of the Chemistry of the Origin of Life. *Angew. Chem., Int. Ed.* **2017**, *56* (37), 11037–11043.

(2) Szostak, J. W.; Bartel, D. P.; Luisi, P. L. Synthesizing life. *Nature* **2001**, *409* (6818), 387–90.

(3) Chen, I. A.; Roberts, R. W.; Szostak, J. W. The emergence of competition between model protocells. *Science* **2004**, *305* (5689), 1474–6.

(4) Ganti, T. Organization of chemical reactions into dividing and metabolizing units: the chemotons. *BioSystems* **1975**, *7* (1), 15–21.

(5) Orgel, L. E. Molecular replication. *Nature* **1992**, *358* (6383), 203–9.

(6) Gebicki, J. M.; Hicks, M. Ufasomes are stable particles surrounded by unsaturated fatty acid membranes. *Nature* **1973**, *243* (5404), 232–4.

(7) Gilbert, W. Origin of life: The RNA world. *Nature* **1986**, *319*, 618.

(8) Joyce, G. F. RNA evolution and the origins of life. *Nature* **1989**, *338* (6212), 217–24.

(9) Ma, W.; Yu, C.; Zhang, W.; Hu, J. Nucleotide synthetase ribozymes may have emerged first in the RNA world. *RNA* **2007**, *13* (11), 2012–2019.

(10) Weimann, B. J.; Lohrmann, R.; Orgel, L. E.; Schneider-Bernloehr, H.; Sulston, J. E. Template-directed synthesis with adenosine-5'-phosphorimidazole. *Science* **1968**, *161* (3839), 387.

(11) Szostak, J. W. The eightfold path to non-enzymatic RNA replication. *J. Syst. Chem.* **2012**, *3* (1), 2.

(12) Zielinski, W. S.; Orgel, L. E. Oligomerization of activated derivatives of 3'-amino-3'-deoxyguanosine on poly(C) and poly(dC) templates. *Nucleic Acids Res.* **1985**, *13* (7), 2469–84.

(13) Tohidi, M.; Zielinski, W. S.; Chen, C. H.; Orgel, L. E. Oligomerization of 3'-amino-3'-deoxyguanosine-5'-phosphorimidazolate on a d(CpCpCpCpC) template. *J. Mol. Evol.* **1987**, *25*, 97–9.

(14) Zhang, S.; Zhang, N.; Blain, J. C.; Szostak, J. W. Synthesis of N3'-P5'-linked phosphoramidate DNA by nonenzymatic template-directed primer extension. *J. Am. Chem. Soc.* **2013**, *135* (2), 924–32.

(15) Zhang, S.; Blain, J. C.; Zielinska, D.; Gryaznov, S. M.; Szostak, J. W. Fast and accurate nonenzymatic copying of an RNA-like synthetic genetic polymer. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (44), 17732–7.

(16) Schrum, J. P.; Ricardo, A.; Krishnamurthy, M.; Blain, J. C.; Szostak, J. W. Efficient and rapid template-directed nucleic acid copying using 2'-amino-2',3'-dideoxyribonucleoside-5'-phosphorimidazole monomers. *J. Am. Chem. Soc.* **2009**, *131* (40), 14560–70.

(17) Chen, J. J.; Cai, X.; Szostak, J. W. N2'→p3' phosphoramidate glycerol nucleic acid as a potential alternative genetic system. *J. Am. Chem. Soc.* **2009**, *131* (6), 2119–21.

(18) Blain, J. C.; Ricardo, A.; Szostak, J. W. Synthesis and nonenzymatic template-directed polymerization of 2'-amino-2'-deoxythreose nucleotides. *J. Am. Chem. Soc.* **2014**, *136* (5), 2033–9.

(19) Brotzel, F.; Chu, Y. C.; Mayr, H. Nucleophilicities of primary and secondary amines in water. *J. Org. Chem.* **2007**, *72* (10), 3679–88.

(20) Moulton, J. D.; Yan, Y.-L., Using Morpholinos to Control Gene Expression. In *Current Protocols in Molecular Biology*; John Wiley & Sons, Inc., 2008.

(21) Stirchak, E. P.; Summerton, J. E.; Weller, D. D. Uncharged stereoregular nucleic acid analogs: 2. Morpholino nucleoside oligomers with carbamate internucleoside linkages. *Nucleic Acids Res.* **1989**, *17* (15), 6129–41.

(22) Joyce, G. F.; Inoue, T.; Orgel, L. E. Non-enzymatic template-directed synthesis on RNA random copolymers. Poly(C, U) templates. *J. Mol. Biol.* **1984**, *176* (2), 279–306.

(23) Kanavarioti, A.; Bernasconi, C. F.; Doodokyan, D. L.; Alberas, D. J. Magnesium ion catalyzed P-N bond hydrolysis in imidazole-activated nucleotides. Relevance to template-directed synthesis of polynucleotides. *J. Am. Chem. Soc.* **1989**, *111* (18), 7247–57.

(24) Eisenhuth, R.; Richert, C. Convenient syntheses of 3'-amino-2',3'-dideoxynucleosides, their 5'-monophosphates, and 3'-amino-terminal oligodeoxynucleotide primers. *J. Org. Chem.* **2009**, *74* (1), 26–37.

(25) Rajamani, S.; Ichida, J. K.; Antal, T.; Treco, D. A.; Leu, K.; Nowak, M. A.; Szostak, J. W.; Chen, I. A. Effect of stalling after mismatches on the error catastrophe in nonenzymatic nucleic acid replication. *J. Am. Chem. Soc.* **2010**, *132* (16), 5880–5.

(26) Walton, T.; Szostak, J. W. A Highly Reactive Imidazolium-Bridged Dinucleotide Intermediate in Nonenzymatic RNA Primer Extension. *J. Am. Chem. Soc.* **2016**, *138* (36), 11996–2002.

(27) Nan, Y.; Zhang, Y.-J. Antisense Phosphorodiamidate Morpholino Oligomers as Novel Antiviral Compounds. *Front. Microbiol.* **2018**, *9*, 750.

(28) Summerton, J.; Weller, D. Morpholino antisense oligomers: design, preparation, and properties. *Antisense Nucleic Acid Drug Dev.* **1997**, *7* (3), 187–95.

(29) Moulton, H. M.; Moulton, J. D. *Morpholino Oligomers: Methods and Protocols*; Humana Press: New York, 2017; 284 pp.