Enhanced Nonenzymatic RNA Copying with 2-Aminoimidazole Activated Nucleotides

Li Li, Noam Prywes,[†] Chun Pong Tam,[ORCID] Derek K. O'Flaherty, Victor S. Lelyveld,[ORCID] Enver Cagri Izgu,[ORCID] Ayan Pal, and Jack W. Szostak*[ORCID]

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

S *Supporting Information*

ABSTRACT: Achieving efficient nonenzymatic replication of RNA is an important step toward the synthesis of self-replicating protocells that may mimic early forms of life. Despite recent progress, the nonenzymatic copying of templates containing mixed sequences remains slow and inefficient. Here we demonstrate that activating nucleotides with 2-aminoimidazole results in superior reaction kinetics and improved yields of primer extension reaction products. This new leaving group significantly accelerates monomer addition as well as trimer-assisted RNA primer extension, allowing efficient copying of a variety of short RNA templates with mixed sequences.

Replication of the genetic information is a key chemical process in living organisms. Whereas modern cells copy their genomes using sophisticated protein enzymes, primordial cells had to rely on primitive mechanisms without access to an elaborate translation system. One such mechanism is nonenzymatic RNA replication,[1] in which chemically activated nucleotides, such as phosphoroimidazolides, spontaneously polymerize when guided by a template sequence. Its simplicity has made it an attractive approach in laboratory attempts to develop self-replicating protocells that mimic early forms of life.[2] However, a major limitation of this approach is its poor yield and slow rate, particularly when copying templates with mixed sequences.[3] Here we report the development of a superior leaving group, 2-aminoimidazole, which enables efficient copying of short RNA templates with mixed sequences.

Since the discovery in 1981 that 2-methylimidazole (2-MI) activated nucleotides show both good reactivity and the ability to direct the formation of natural 3′-5′ linkages, 2-MI has been the most widely used leaving group for studies of nonenzymatic RNA template copying by primer extension.[4] Interestingly, 2-MI activated monomers are more reactive in primer extension than the corresponding imidazole and 4-methylimidazole derivatives,[4] both of which contain better leaving groups than 2-MI as judged by their hydrolysis rates.[5] The higher reactivity of 2-MI activated nucleotides has been attributed to an interaction between leaving groups of neighboring monomers, such that the leaving group on the downstream nucleotide exhibits a catalytic effect that increases the rate of reaction between the upstream nucleotide and the adjacent primer.[6] This leaving-group interaction accelerates the primer extension reaction by 10−100 times, depending on whether the downstream nucleotide is an activated monomer[6] or an activated oligonucleotide.[7] We have recently suggested that the physical origin of this rate enhancement is due to the covalent reaction of 2-MI-activated monomers with each other to form imidazolium bridged dinucleotides, which are highly reactive intermediates in the copying of RNA templates by primer extension.[8]

In an effort parallel to the above mechanistic investigations,[5−8] we sought to develop a leaving group with enhanced catalytic activity when used to activate nucleotides in positions downstream of the reaction site from a screen of leaving groups. We measured primer extension rates on a template that encodes the product sequence -CGGG-3′. Reaction mixtures contained C monomers activated by 2-MI (2-MeImpC), which could only bind the initial +1 position, and G monomers activated by various leaving groups, which could only bind downstream (**1a−1h**, Figure 1). The results revealed two general trends. First, a larger substituent at the 2-position slows down the reaction (**1a−1d**). Second, the reaction rate decreases as the pK_a of the leaving group decreases (**1a, 1e−1h**). Replacing 2-MI on the downstream G monomers by 2-methylpyrrole (**1i**) abolishes the catalytic effect, indicating that neither pi-stacking between neighboring leaving groups, nor the NH moiety alone, which may serve as a hydrogen-bond donor, are sufficient to catalyze primer extension; the lack of a catalytic effect is consistent with primer extension proceeding through an imidazolium-bridged intermediate.

Guided by these two empirical rules, we reasoned that 2-aminoimidazole (2-AI, **1j**), which has a small substituent on the 2-position and a higher pK_a than 2-MI,[9] might serve as a better catalyst of primer extension with 2-MeImpC. Indeed, 2-AI activated pG accelerated 2-MeImpC addition by 7 times compared to its 2-MI activated counterpart. A plausible explanation for the enhanced reactivity of 2-AI-activated pG is that 2-AI accelerates the formation of the 2-amino-imidazolium bridged dinucleotide intermediate due to the greater nucleophilicity of its imidazole N3. The larger 2-methylaminoimidazole leaving group (**1k**) strongly reduced the rate despite its high pK_a. Although the poor reactivity of **1k**-activated pG can be attributed in part to steric hindrance by the bulkier 2-methylamino substituent, the magnitude of the effect is surprising and is the subject of ongoing investigation.
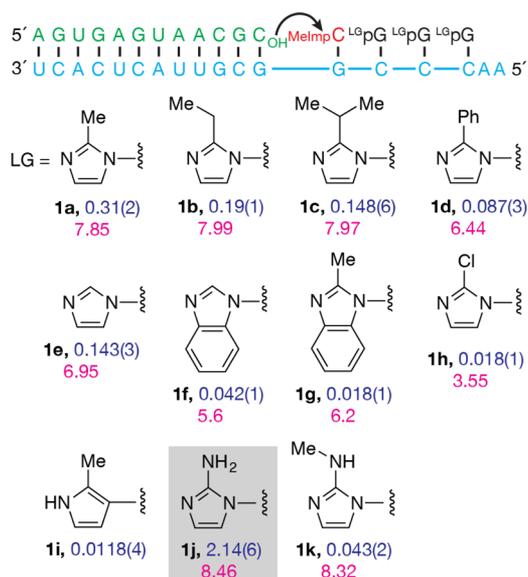
Figure 1. Development of 2-aminoimidazole as a superior leaving group for nonenzymatic primer extension. Top row: the template (cyan)-dependent reaction of 2-MI activated C monomers (red) with the primer (green) was used to probe the effect of adjacent leaving groups (LG, black) on the reaction rate. Curved arrow indicates the site of reaction. Varying the leaving group (**1a**–**1k**) influenced the rates by up to 200 times, with 2-aminoimidazole (**1j**) giving the fastest rate. Reaction conditions: 1.5 μM primer, 5 μM template, 200 mM HEPES pH 8.0, 50 mM $MgCl_2$, and 20 mM monomers. All experiments were performed in triplicate. Blue: mean observed rate in h^{-1}; standard error of the mean (s.e.m.) in the last digit is shown in parentheses. Magenta: pK_a of the leaving group.

In addition to their superior catalytic effect, monomers activated by 2-AI are also better substrates at the +1 site than those activated by 2-MI (Figure S1), an observation consistent with their faster hydrolysis rates in a solution mimicking the primer-extension condition (Figure S2). As a result, the phosphoroimidazolide **1j-pG** extends the primer far more efficiently than **1a-pG** on a template that encodes the product sequence -GGGG-3′ (Figure 2). This motivated us to try using 2-AI activated monomers in both the upstream and downstream positions in primer extension reactions. For various
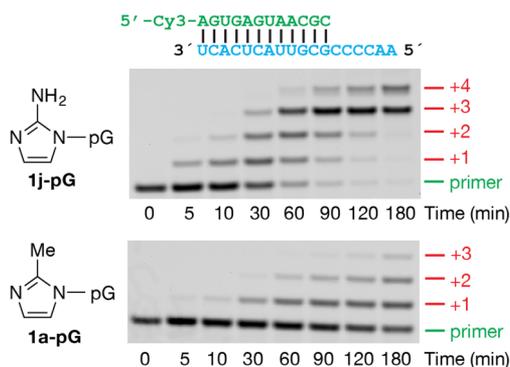
Figure 2. Efficient nonenzymatic RNA copying using phosphoroimidazolides with 2-AI leaving group. The time course of primer extension reaction using phosphoroimidazolides **1j-pG** and **1a-pG** were assayed using polyacrylamide gel electrophoresis (PAGE). Positions of primer and +1 to +4 products are indicated. Reaction conditions as described in Figure 1.

short templates, including those directing the incorporation of A, U, and 2-thiouridine (s^2U) monomers, 2-AI activated monomers react from 6 to 100 times faster than their 2-MI activated counterparts (Table 1). The rate enhancement of s^2U

Table 1. Reaction Rates of Nonenzymatic Primer Extension Measured by Primer Consumption

	reaction rates (h^{-1})[a]	
product sequence	2-AI leaving group	2-MI leaving group
-CCC-3′	6.7 ± 0.1	1.11 ± 0.03
-GGG-3′	4.92 ± 0.09	0.61 ± 0.01
-CGG-3′	7.5 ± 0.1	0.31 ± 0.02
-GCC-3′	4.5 ± 0.2	0.090 ± 0.003
-AGG-3′	2.81 ± 0.03	0.026 ± 0.003
-UGG-3′	0.62 ± 0.01	n.d.[b]
-s^2UGG-3′	1.4 ± 0.3	0.035 ± 0.002

[a]Reaction conditions as described in Figure 1. Results are given as the mean ± s.e.m. of triplicate experiments. [b]Too slow to be measured.

over U is consistent with our previous studies using the 2-MI leaving group.[10] Collectively, these results establish that monomers activated by 2-AI are superior substrates for nonenzymatic RNA primer extension.

Encouraged by the superior performance of 2-AI activated monomers, we next asked if these monomers can copy templates containing mixed sequences with the assistance of downstream activated "helper" trinucleotides (trimers).[7,11] To copy a template that encodes the product sequence -UGAC-3′, four 2-AI activated trimers (GAC, ACG, CGG, and GGU, where the underlined nucleotides have a 5′-phosphoro-(2-aminoimidazole) moiety) were mixed with 2-AI activated A, C, G, and s^2U monomers. The reaction time course (Figure 3a) revealed +1, + 2, and +3 extension intermediates, indicating that most primers extended via addition of monomers rather than ligation of trimers. After 1 day, 95% of the primer had been converted to +4 products, some of which had then further reacted to form +5 to +7 products.

We confirmed the identities of the primer extension products by liquid chromatography—mass spectrometry (Figure 3b). To enhance the chromatographic separation, we used unlabeled primers, and templates that contained five-nucleotide 3′-overhangs. An exhaustive mass-based search revealed peaks corresponding to the correct +1 to +7 products including a major peak of the +4 full-length product and only one minor peak that is consistent with a product having misincorporated s^2U in place of A at the +3 position, likely due to an s^2U:U mismatch.[12] Our analysis did not reveal any misincorporated +4 products, indicating that such a mismatch slows down further primer extension[13] even in the presence of "helper" trimers. In addition, we did not observe products containing misincorporations caused by G-U wobble pairing, confirming that 2-thiolation effectively suppresses G-U mismatch formation.[10]

Compared with the 2-MI leaving group, the new 2-AI leaving group significantly enhances the efficiency of trimer-assisted template copying. Until now, an AA dinucleotide in the template has been a significant roadblock for primer extension.[3,14] Even with the assistance of trimers activated by 2-MI, the yield of +2 products is only 10% and 50% for U and s^2U monomers, respectively.[7] With 2-AI as a leaving group, the yield increased to 92% for U monomers (Figure 4a) and 98% for s^2U monomers (Figure 4b), respectively. As such, the nonenzymatic copying of an AA dinucleotide, a reaction long
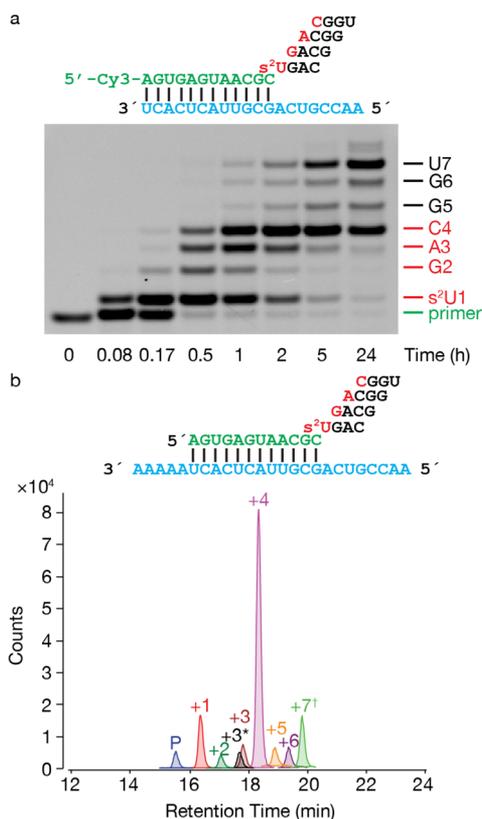
Figure 3. Efficient one-pot extension of primers by four nucleotides. (a) Electropheretogram of extension products after different reaction times, as assayed by PAGE. Positions of primer and +1 to +7 products are indicated. Reaction conditions: 1.2 μM primer (green), 1.5 μM template (cyan), 200 mM HEPES pH 8.0, 50 mM MgCl$_2$, 0.5 mM trimer (black), and 10 mM monomer (red). (b) Overlay of extracted ion chromatograms of residual primer (P) and extension products after 1 day (+1 to +7) observed by LC-MS. Observed ions for the peak labeled +3* are consistent with a product having misincorporated s^{2}U in place of A at the +3 position. Observed ions for the peak +7† are consistent with either the correct +7 product or an alternative double mutant product (see Table S1).

considered to be impossible under normal reaction conditions, has now been achieved by combining the 2-thiolation-induced preorganization of monomers in the reactive C$_{3'}$-*endo* conformation,[10,15] the catalytic effect of helper trimers,[7] and the superior reactivity of the 2-AI leaving group.

To test the generality of this method, we synthesized two additional mixed-sequence templates containing AU and UA sequences that are also considered to be challenging for nonenzymatic RNA copying.[3] Satisfyingly, both templates can be copied with overall yields higher than 85% in just 1 day with unmodified A, C, G, and U monomers (Figure 4c,d). The average single-nucleotide extension yield is 97–98%. We are currently engaged in a systematic exploration of the sequence space available to nonenzymatic RNA copying using the 2-AI leaving group.

The development of the 2-aminoimidazole leaving group is a significant advance in nonenzymatic RNA copying. 2-AI imparts higher reactivity to phosphoroimidazolides, significantly increasing both the rate and the yield of primer extension. The improvement in leaving-group reactivity also synergizes with 2-thiolation of uridine and the use of catalytic trimers, both of which are prebiotically plausible,[7,16] enabling
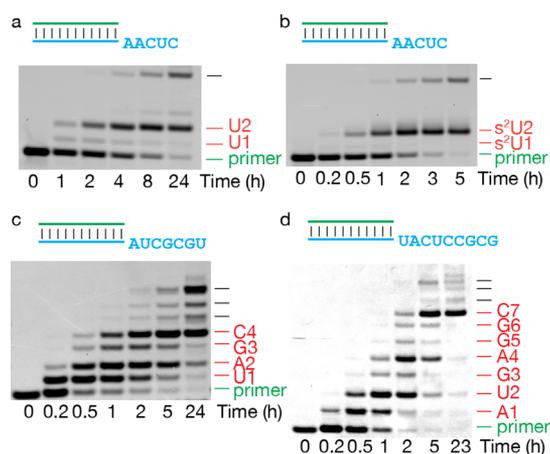
Figure 4. Efficient one-pot nonenzymatic copying of various short RNA templates. (a) Copying AA dinucleotide templates using 2-AI activated U monomers and trimers. (b) Copying AA dinucleotide templates using 2-AI activated s^{2}U monomers and trimers. (c) Extending primers by UAGC using 2-AI activated A, C, G, and U monomers and trimers. (d) Extending primers by the sequence AUGAGGC using 2-AI activated A, C, G, and U monomers and trimers. The extension products were assayed by PAGE. Positions of primers and extended products are indicated. Reaction conditions as described in Figure 3.

efficient nonenzymatic copying of various short RNA templates with mixed sequences. Notably, 2-AI is also related to 2-aminooxazole, a key intermediate in the prebiotic synthesis of the pyrimidine ribonucleotides.[17] Further optimization of the primer-extension reaction may reveal a prebiotically plausible pathway to the efficient copying of RNAs that are long enough to encode important biological functions such as ribozyme activity.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.6b13148.

> Figures S1–S2, Tables S1–S3, experimental procedures and analytical data (^{1}H and ^{13}C NMR, HRMS) for all new compounds (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*szostak@molbio.mgh.harvard.edu

ORCID Ⓘ

Chun Pong Tam: 0000-0001-6381-9011
Victor S. Lelyveld: 0000-0002-3890-0288
Enver Cagri Izgu: 0000-0001-6673-3635
Jack W. Szostak: 0000-0003-4131-1203

Present Address

†Department of Plant and Environmental Sciences, The Weizmann Institute of Science, P.O. Box 26, Rehovot 76100, Israel

■ REFERENCES

(1) Sulston, J.; Lohrmann, R.; Orgel, L. E.; Miles, H. T. *Proc. Natl. Acad. Sci. U. S. A.* **1968**, *59*, 726−733.

(2) Blain, J. C.; Szostak, J. W. *Annu. Rev. Biochem.* **2014**, *83*, 615−640.

(3) Wu, T.; Orgel, L. E. *J. Am. Chem. Soc.* **1992**, *114*, 7963−7969.

(4) Inoue, T.; Orgel, L. E. *J. Am. Chem. Soc.* **1981**, *103*, 7666−7667.

(5) Li, L.; Lelyveld, V. S.; Prywes, N.; Szostak, J. W. *J. Am. Chem. Soc.* **2016**, *138*, 3986−3989.

(6) Wu, T.; Orgel, L. E. *J. Am. Chem. Soc.* **1992**, *114*, 5496−5501.

(7) Prywes, N.; Blain, J. C.; Del Frate, F.; Szostak, J. W. *eLife* **2016**, *5*, e17756.

(8) Walton, T.; Szostak, J. W. *J. Am. Chem. Soc.* **2016**, *138*, 11996−12002.

(9) Storey, B. T.; Sullivan, W. W.; Moyer, C. L. *J. Org. Chem.* **1964**, *29*, 3118−3120.

(10) Heuberger, B. D.; Pal, A.; Del Frate, F.; Topkar, V. V.; Szostak, J. W. *J. Am. Chem. Soc.* **2015**, *137*, 2769−2775.

(11) Deck, C.; Jauker, M.; Richert, C. *Nat. Chem.* **2011**, *3*, 603−608.

(12) Sheng, J.; Larsen, A.; Heuberger, B. D.; Blain, J. C.; Szostak, J. W. *J. Am. Chem. Soc.* **2014**, *136*, 13916−13924.

(13) Leu, K.; Kervio, E.; Obermayer, B.; Turk-MacLeod, R.; Yuan, C.; Luevano, J.-M.; Chen, E.; Gerland, U.; Richert, C.; Chen, I. *J. Am. Chem. Soc.* **2013**, *135*, 354−366.

(14) Löffler, P. M. G.; Groen, J.; Dörr, M.; Monnard, P.-A. *PLoS One* **2013**, *8*, e75617.

(15) Zhang, N.; Zhang, S.; Szostak, J. W. *J. Am. Chem. Soc.* **2012**, *134*, 3691−3694.

(16) Xu, J.; Tsanakopoulou, M.; Magnani, C. J.; Szabla, R.; Šponer, J. E.; Šponer, J.; Góra, R. W.; Sutherland, J. D. *Nat. Chem.* **2016**, DOI: 10.1038/nchem.2664.

(17) Powner, M. W.; Gerland, B.; Sutherland, J. D. *Nature* **2009**, *459*, 239−242.