3272–3283 *Nucleic Acids Research, 2019, Vol. 47, No. 7* *Published online 1 March 2019*
doi: 10.1093/nar/gkz029

One-step enzymatic modification of RNA 3′ termini using polymerase θ

Crystal Thomas[1], Timur Rusanov[1], Trung Hoang[1], Taurai Augustin[1], Tatiana Kent[1], Imre Gaspar[2] and Richard T. Pomerantz[1,*]

[1]Fels Institute for Cancer Research, Department of Medical Genetics and Molecular Biochemistry, Temple University Lewis Katz School of Medicine, Philadelphia, PA, 19140, USA and [2]European Molecular Biology Laboratory, Heidelberg, Meyerhofstrasse 1, 69117, Germany

Received April 10, 2018; Revised January 13, 2019; Editorial Decision January 14, 2019; Accepted February 13, 2019

ABSTRACT

Site-specific modification of synthetic and cellular RNA such as with specific nucleobases, fluorophores and attachment chemistries is important for a variety of basic and applied research applications. However, simple and efficient methods to modify RNA such as at the 3′ terminus with specific nucleobases or nucleotide analogs conjugated to various chemical moieties are lacking. Here, we develop and characterize a one-step enzymatic method to modify RNA 3′ termini using recombinant human polymerase theta (Polθ). We demonstrate that Polθ efficiently adds 30–50 2′-deoxyribonucleotides to the 3′ terminus of RNA molecules of various lengths and sequences, and extends RNA 3′ termini with an assortment of 2′-deoxy and 2′,3′-dideoxy ribonucleotide analogs containing functional chemistries, such as high affinity attachment moieties and fluorophores. In contrast to Polθ, terminal deoxynucleotidyl transferase (TdT) is unable to use RNA as a substrate altogether. Overall, Polθ shows a strong preference for adding deoxyribonucleotides to RNA, but can also add ribonucleotides with relatively high efficiency in particular sequence contexts. We anticipate that this unique activity of Polθ will become invaluable for applications requiring 3′ terminal modification of RNA and potentially enzymatic synthesis of RNA.

INTRODUCTION

DNA polymerase theta (Polθ) is essential for the double-strand break (DSB) repair pathway alternative end-joining (alt-EJ), also known as microhomology-mediated end-joining (MMEJ) (1–8). Due to the promiscuous nature of the unique A-family Polθ, alt-EJ is highly mutagenic. For example, the repair junctions formed by alt-EJ frequently harbor nucleotide insertions that are catalyzed by Polθ 3′ terminal deoxynucleotidyl transferase (TdT) activity (4–7,9). Intriguingly, Polθ has been reported to exhibit template- dependent and -independent deoxynucleotidyl transfer activities at the 3′ terminus of single-stranded DNA (ssDNA), which likely contributes to the mutagenic signature of alt-EJ (9,10) that has been observed in invertebrates and vertebrates (4,5,7,8). Notably, the apparent template-independent activity has been exclusively observed in the presence of $MnCl_2$ rather than $MgCl_2$ (9).

Aside from Polθ, the only other polymerase known to efficiently transfer several nucleotides to DNA 3′ terminal ends is the X-family polymerase member TdT (11–14). The X-family members, Polλ and Polμ, also exhibit terminal transferase activities, albeit at significantly lower efficiencies, and these activities contribute to antibody diversity generated during VDJ recombination (15–18). Similar to Polθ, TdT was selected throughout evolution to generate nucleotide insertions at DSBs, specifically during VDJ recombination (11,14). Recent biochemical studies directly compared the ssDNA terminal transferase activities of TdT and Polθ (9). Both enzymes showed a similar ability to transfer canonical deoxyribonucleotides to the 3′ terminus of ssDNA (9). Yet, Polθ demonstrated a substantially higher proficiency for transferring ribonucleotides to ssDNA, and transferred a wider variety of nucleotide analogs containing functional motifs (e.g. attachment chemistries and fluorophores) (9). Thus, overall Polθ appears to be more versatile in adding nucleotides to ssDNA 3′ termini than TdT. Interestingly, recent biochemical and structural studies demonstrate that TdT also exhibits template-dependent 3′ terminal nucleotidyl transferase activity like Polθ (12). Thus, the idea that 3′ terminal nucleotidyl transferase activity solely depends on template-independent mechanisms might need to be revisited.

With the recent characterization of Polθ as a highly versatile ssDNA nucleotidyl terminal transferase in the presence of $MnCl_2$ (9,10), further studies are needed to

*To whom correspondence should be addressed. Tel: +1 215 707 8991; Fax: +1 215 707 1454; Email: richard.pomerantz@temple.edu
Publication delayed at the author's request.

investigate new applications for this enzyme in nucleic acid modification and synthetic biology. For example, there is currently a lack of enzymatic means to chemically modify RNA 3′ termini. Thus, the development of a templated or non-templated RNA terminal nucleotidyl transferase capable of utilizing a wide range of nucleotide analogs would be highly useful for biotechnologies requiring chemically modified RNA. Insofar, poly(A) polymerase, CCA-adding enzymes and terminal uridyl transferases are the only enzymes capable of efficient transfer of nucleotides to the 3′ terminus of RNA (19–21). Yet, these enzymes were selected to use particular ribonucleotides and/or RNA substrates that greatly limit their capacity for diverse biotechnology applications (19–21). Considering that RNA biological research and RNA biotechnology have demonstrated extensive growth and shown increasing importance in medicine, the development of an enzyme capable of efficient RNA 3′ terminal extension activity is warranted.

MATERIALS AND METHODS

Polθ 3′ terminal extension activity

An amount of 200 nM Polθ (or other concentrations as indicated) was incubated with 50 nM of the indicated 5′ ^{32}P-labeled RNA or ssDNA at 37° (or other temperatures as noted) in the presence of 20 mM Tris–HCl pH 8.2, 0.01% NP-40, 0.1 mg/ml bovine serum albumin (BSA) and 10% glycerol. Figure 1A and B contained either 2 mM $MnCl_2$ or 2 mM $MgCl_2$ as indicated. All other reactions contained 2 mM $MnCl_2$. Twenty units of Ambion™ RNase Inhibitor (Thermo-Scientific) was added to reactions containing RNA. An amount of 500 μM of either ribonucleoside triphosphates (NTPs) or deoxyribonucleoside triphosphates (dNTPs) (or individual dNTP or NTPs as noted) was present in reactions as indicated. Nucleotide analogs were added at 100 μM. Reactions were terminated after 60 min (unless otherwise noted) by the addition of 25 mM ethylenediaminetetraacetic acid (EDTA) and 45% formamide and were resolved by electrophoresis in 15% urea polyacrylamide gels, then visualized by phosphorimager or autoradiography. Percentage extension was determined by dividing the sum of the intensities of the extended product bands by the sum of the intensities of the extended and unextended bands. ImageJ was used to determine the intensities of individual bands.

Polθ 3′ terminal extension processivity assays

Figure 1G: Reactions were performed with 200 nM Polθ at 37° in the presence of 50 nM of the indicated 5′ ^{32}P-labeled RNA, 2 mM $MnCl_2$, 500 μM dNTPs, 20 mM Tris–HCl pH 8.2, 0.01% NP-40, 0.1 mg/ml BSA, 10% glycerol and 20 units of AmbionTM RNase Inhibitor (Thermo-Scientific). Cold trap (7.5 μM 454R) was added 30 s later. Reactions were terminated at the indicated time points by the addition of 25 mM EDTA and 45% formamide and were resolved by electrophoresis in 15% urea polyacrylamide gels, then visualized by phosphorimager or autoradiography.

Figure 1H: An amount of 2 nM Polθ was incubated with 100 nM of the indicated 5′ ^{32}P-labeled RNA at 37° in the presence of 2 mM $MnCl_2$, 20 mM Tris-HCl, pH 8.2, 0.01% NP-40, 0.1 mg/ml BSA, 10% glycerol and 20 units of AmbionTM RNase Inhibitor (Thermo-Scientific). Reactions were initiated with 500 μM dNTPs. Thirty seconds later, 10 μM cold trap (454R) was added to the reaction. Reactions were terminated at the indicated time points by the addition of 25 mM EDTA and 45% formamide and were resolved by electrophoresis in 15% urea polyacrylamide gels, then visualized by phosphorimager or autoradiography.

TdT terminal transferase activity

An amount of 200 nM of TdT was incubated with 50 nM of the indicated 5′ ^{32}P-labeled RNA or ssDNA in conditions recommended by New England Biolabs (50 mM potassium acetate, 20 mM Tris acetate, 10 mM magnesium acetate, pH 7.9, with 0.25 mM cobalt) at 37°C. Twenty units of Ambion™ RNase Inhibitor (Thermo-Scientific) was added to reactions containing RNA. NTPs or dNTPs were present at 500 μM and reactions were incubated for 1 h. Reactions were terminated and processed as above.

Proteins

Polθ was purified as described using a newly generated SUMOstar (Life Sensors) expression vector and SUMOstar protease I (Life Sensors) (9). TdT was obtained from New England Biolabs (NEB).

- **Nucleic acids** (5′-3′) (Integrated DNA Technologies)
- DNA:
- RP334cy3 /5Cy3/TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
- RP360 CCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
- 469D CTGTCCTGCATGATG
- 470D CATCATGCAGGACAGTCGGATCGCAGTCAG
- RNA:
- 454R GUCCAGCCCAAACU
- 456R GGAGAGAAAAGG
- 457R ACAGUUUUUUUUGCCCUUUUCUCUCC
- 458R GUCCAGCCCAAACUCU
- 465R GCUCACCUGACUUUUUUUUGACCUUUUCUCUCC
- 467R AGAACCUGUUGAACAAAAGC
- 468R ACAGUUUUUUUUGCUUUUGUUCAACAGGUUCU
- 470R CAUCAUGCAGGACAGUCGGAUCGCAGUCAG
- 477R CAUCAUGCAGGACAGUUUUUUUUUUUUUUGCUUUUGUUCAACAGGUUCU
- 478R UGAAGUACUUCGGUACU
- 493R UUUUUUUCGCGCUGCGACGUCG
- 494R UUUUCUGCGCGCUGCGACGUCG
- 495R CCCCCCGGAGAGAAAAGG
- 496R UUUUCCCCGGAGAGAAAAGG
- SM44R GCUCUGAUGCCGCAUAGUUAAGCCAGCCCCGACACCCG
- SM103R GUACACGGCCGCAUAAUCAGAGUUUGGGCUGGAC

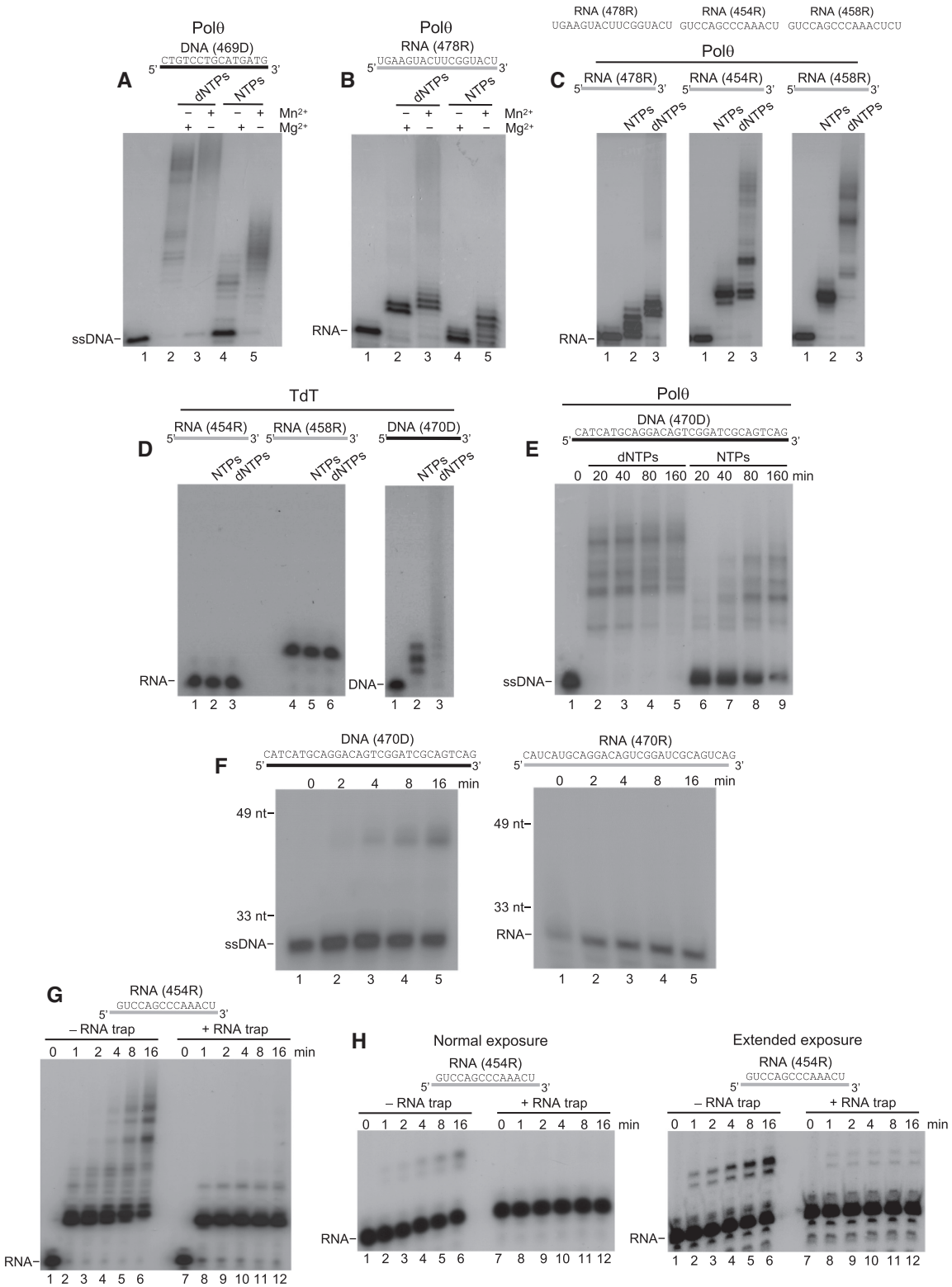

Figure 1. Comparison of Polθ and TdT RNA 3′ terminal extension activities. (**A** and **B**) Denaturing gels showing Polθ extension activity on DNA (A) and RNA (B) in the presence of the indicated dNTPs, NTPs and divalent cations. (**C**) Denaturing gels showing Polθ RNA 3′ terminal extension activity on the indicated substrates in the presence of Mn^{2+} and dNTPs or NTPs. (**D**) Denaturing gel showing lack of TdT RNA terminal nucleotidyl transferase activity in the presence of Co^{2+} and the indicated substrates (left). Denaturing gel showing TdT DNA terminal nucleotidyl transferase activity in the presence of Co^{2+} and the indicated substrate (right). (**E**) Denaturing gels showing a time course of Polθ extension activity on the indicated DNA substrate in the presence of dNTPs (lanes 2–5) and NTPs (lanes 6–9). (**F**) Denaturing gels showing a time course of Polθ extension activity on the indicated DNA and RNA substrates in the presence of dNTPs. (**G** and **H**) Denaturing gels showing a time course of Polθ extension activity on the indicated RNA substrate in the presence of dNTPs with and without excess unlabeled RNA trap added 30 s after reactions were initiated.

- SM271R GAGACUCGGCGCUGCGACGUCG
- SM275R GAGACUGCGCGCUGCGACGUCG

RNA and DNA were ^{32}P 5′-labeled using ^{32}P-γ-adenosine triphosphate (ATP) (Perkin Elmer) and bacteriophage T4 polynucleotide kinase (New England Biolabs).

RESULTS

In previous biochemical studies, the polymerase domain of DNA polymerase θ, herein referred to as Polθ, was shown to exhibit efficient 3′ terminal ssDNA extension activity in the presence of deoxyribonucleotides and Mg^{2+} (10). This 3′ terminal ssDNA extension activity was shown to be greatly stimulated by Mn^{2+}, which increases the promiscuity of polymerases (9). Kent, *et al* also showed the ability of Polθ to transfer 2′-deoxyribonucleotides to a RNA substrate (9). However, the full potential for Polθ to efficiently add canonical and modified nucleotides to the 3′ terminus of RNA remains to be elucidated. For example, whether Polθ can efficiently add 2′-deoxyribonucleotides to the 3′ terminus of RNA of various lengths and sequence remains unknown. Furthermore, whether Polθ can efficiently add to RNA 2′-hydroxy ribonucleotides and a variety of 2′-hydroxy and 2′-deoxy ribonucleotide analogs which would be beneficial for RNA biotechnology remains to be elucidated.

Here, we investigated the ability of recombinant human Polθ (polymerase domain; residues 1792–2590) to transfer to RNA 2′-deoxy and 2′-hydroxy ribonucleotides, and several types of 2′-hydroxy, 2′-deoxy and 2′,3′-dideoxy ribonucleotide analogs containing functional groups *in vitro*. First, we confirmed the ability of Mn^{2+} to stimulate Polθ ssDNA terminal 2′-deoxy nucleotidyl transferase activity, as in previous studies (9). All ssDNA and RNA substrate sequences are listed in Table 1. We note that our use of the term nucleotidyl transferase activity does not exclusively refer to a template-independent mechanism of nucleotide addition as this process may utilize templated or non-templated mechanisms (9,12). As expected, Polθ 3′ terminal ssDNA extension activity with deoxyribonucleotides is slightly enhanced in the presence of Mn^{2+} (Figure 1A, compare lanes 2 and 3). Remarkably, Mn^{2+} showed even greater stimulation of Polθ ssDNA extension activity with ribonucleotides, indicating reduced discrimination against these substrates (Figure 1A, compare lanes 4 and 5). We next examined the ability of Polθ to add nucleotides to RNA with Mg^{2+} versus Mn^{2+}. Here, again Mn^{2+} stimulated Polθ addition of 2′-deoxy and 2′-hydroxy nucleotides to RNA, yet the overall 3′ terminal extension activity of Polθ was less efficient than that observed on ssDNA (compare Figure 1A and 1B). Consistent with this, the rate of Polθ 2′-deoxyribonucleotide addition under steady-state conditions is significantly slower on RNA versus ssDNA substrates composed of identical sequence (Figure 1F). A positive control for Polθ extension activity on the RNA substrate in Figure 1F at higher enzyme concentrations is presented in Supplementary Figure 1A. We examined Polθ RNA extension activity with 2′-hydroxy and 2′-deoxy ribonucleotides on multiple substrates to initially assess whether there are sequence specific effects. For example, because RNA often forms complex structures, it is conceivable that particular RNA sequences (and structures) may impose restrictions on Polθ RNA extension. Three different RNA sequences were examined in Figure 1C on the same day and using the identical aliquot of Polθ. The results show that Polθ exhibits highly efficient 3′ terminal 2′-deoxy and 2′-hydroxy ribonucleotide extension activity on RNA substrates 458R and 454R, whereas these activities are less efficient on RNA 478R (Figure 1C). Remarkably, Polθ extends the 458R and 454R RNA substrates with nearly 100% efficiency in the presence of 2′-deoxy and 2′-hydroxy ribonucleotides. Quantitative addition of 2′-deoxy ribonucleotides is also observed on RNA 478R, despite fewer addition events (i.e. addition of 3–4 nt). In each case, Polθ shows a preference for adding 2′-deoxy ribonucleotides which is consistent with its cellular function as a DNA polymerase (1,3–5,8,9,22). Figure 1E further compares Polθ extension activity in the presence of dNTPs versus ribonucleoside triphosphates (NTPs). Here, a time course reaction was performed with dNTPs and NTPs using the identical ssDNA substrate. As expected from previous studies, Polθ strongly discriminates against NTPs even under our conditions which utilize $MnCl_2$ that increases the promiscuity of Polθ (Figure 1E). Intriguingly, Polθ surprisingly possesses a low discrimination against 2′-hydroxy ribonucleotides in certain RNA sequence contexts. For example, as noted above Polθ exhibits a relatively high efficiency of RNA extension activity with 2′-hydroxyl nucleotides on the substrates used in the center and right panels compared to the substrate in the left panel in Figure 1C. This phenomenon is further discussed below.

Next, we examined whether X-family polymerase TdT exhibits similar RNA extension activity. Using identical RNA substrates and optimal reaction conditions for TdT, we failed to observe any nucleotide additions on either substrate, even after 60 min of incubation (Figure 1D, left). As a positive control, we demonstrate that the same stock of TdT used on the same day is fully active on ssDNA as expected (Figure 1D, right). Taken together, these data demonstrate that Polθ possesses unique properties that allow it to efficiently add nucleotides to the 3′ terminus of RNA.

Because 3′ terminal RNA extension activity is likely to be useful for a variety of biotechnology and synthetic biology applications, we identified optimal conditions for this function of Polθ (Supplementary Figure S2). For example, ≥2 mM Mn^{2+}, buffer pH 8.2 and excess amounts of Polθ over RNA substrate are required for maximal RNA extension activity using 2′-hydroxy ribonucleotides (Supplementary Figure S2). We utilized these optimal conditions (i.e. 2 mM $MnCl_2$, Tris–HCl pH 8.2, 5–10 fold excess of Polθ over RNA) throughout the remainder of the study. We note that under more physiologically relevant conditions where Polθ is added at low amounts (i.e. 5 nM) along with 2 mM $MgCl_2$ instead of $MnCl_2$, the enzyme exhibits very poor RNA extension activity (Supplementary Figure S1B). Thus, we do not anticipate that Polθ performs this function in cells.

We next examined whether the observed 3′ terminal RNA extension activity occurs with high processivity or whether the polymerase acts distributively. To establish initial conditions to test for Polθ processivity, we performed a simple time course of RNA extension activity, first using 200

Table 1. Oligonucleotide sequences

RNA (5′-3′)
454R GUCCAGCCCAAACU
456R GGAGAGAAAAGG
457R ACAGUUUUUUUUGCCCUUUUCUCUCC
458R GUCCAGCCCAAACUCU
465R GCUCACCUGACUUUUUUUUGACCUUUUCUCUCC
467R AGAACCUGUUGAACAAAAGC
468R ACAGUUUUUUUUGCUUUUGUUCAACAGGUUCU
470R CAUCAUGCAGGACAGUCGGAUCGCAGUCAG
477R CAUCAUGCAGGACAGUUUUUUUUUUUUUUUGCUUUUGUUCAACAGGUUCU
478R UGAAGUACUUCGGUACU
493R UUUUUUUCGCGCUGCGACGUCG
494R UUUUCUGCGCGCUGCGACGUCG
495R CCCCCCGGAGAGAAAAGG
496R UUUUCCCCGGAGAGAAAAGG
SM44R GCUCUGAUGCCGCAUAGUUAAGCCAGCCCCGACACCCG
SM103R GUACACGGCCGCAUAAUCAGAGUUUGGGCUGGAC
SM271R GAGACUCGGCGCUGCGACGUCG
SM275R GAGACUGCGCGCUGCGACGUCG
DNA (5′-3′)
RP334cy3 /5Cy3/TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
RP360 CCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
469D CTGTCCTGCATGATG
470D CATCATGCAGGACAGTCGGATCGCAGTCAG

nM of the enzyme (Figure 1G, lanes 1–6). Unexpectedly, we observed a fast initial extension phase followed by a significantly slower extension phase. For instance, Polθ initially extends the RNA by ∼4 nt in <1 min, but then further extends the substrate by several nucleotides at a significantly reduced rate. This biphasic rate of RNA extension activity is further discussed below. Despite the two apparent different phases of RNA extension activity on this substrate, we next repeated the time course assay but added a 150-fold excess of unlabeled RNA trap 30 s after the reaction was initiated on the initial radio-labeled substrate. In the event that Polθ acts distributively, we would expect the enzyme to dissociate from the initial template and then preferentially act on the excess unlabeled RNA trap which would prevent further nucleotidyl addition activity on the initial radio-labeled substrate. We observed that the second slow phase of RNA extension was suppressed, demonstrating that at least this phase of terminal extension is non-processive (Figure 1G, lanes 7–12). We note that the initial RNA extension phase is likely to be rapid and occur within just a few seconds. Thus, the addition of the cold RNA trap after 30 s would not influence this initial fast rate of RNA extension. To properly asses the processivity of the initial fast phase, we established steady-state conditions in which a significantly lower concentration of Polθ (2 nM) was added which allows for only a small fraction of the 50-fold excess of the initial RNA substrates to be extended over a longer period of time (i.e. 16 min) (Figure 1H, lanes 1–6). As expected, only the initial extension phase is observed under these enzyme limiting conditions. Next, the identical time course reaction was repeated; however, a 100-fold excess of cold RNA trap was added 30 s after the reaction was initiated. Here, the results show that the initial RNA extension reaction is suppressed in the presence of the excess cold RNA trap, which shows that this reaction phase is also non-processive (Figure 1H, lanes 7–12). The overexposed gel in the right panel shows that the limited activity observed in the presence of the cold RNA trap is not increased over the 16 min time course. Hence, these data taken together demonstrate that Polθ acts distributively during both phases of RNA extension.

We proceeded to test whether Polθ consistently exhibits 3′ terminal 2′-deoxy and 2′-hydroxy ribonucleotide addition activity on several different RNA substrates of various lengths and sequences (Figure 2 and Table 1). The data demonstrate that Polθ possesses robust 2′-deoxyribonucleotide extension activity on all substrates. For example, Polθ adds >40 deoxyribonucleotides to 454R (Figure 2A and lane 3). Similar extension efficiencies are observed on the other substrates with the exception of 477R which gives rise to fewer addition events (Figure 2A and lane 18). Nevertheless, in all cases Polθ quantitatively extends all the RNA substrates in the presence of 2′-deoxyribonucleotides. We note that the RNA substrates contain either pyrimidines or purines at their 3′ termini in order to prevent sequence bias at the 3′ terminus (Figure 2 and Table 1). The lengths of the RNA used are also wide ranging (14–49 nt) (Figure 2 and Table 1). In the case of ribonucleotides, only the 454R substrate results in efficient extension activity (Figure 2A and lane 2). For example, here Polθ quantitatively adds ∼4–5 ribonucleotides, whereas little or no activity is observed on the other substrates in the presence of ribonucleotides. These data suggest that the sequence context of RNA can influence Polθ ribonucleotide extension activity. Consistent with this idea, we observed that Polθ exhibits efficient ribonucleotide extension activity on RNA substrates consisting of similar sequence (see substrates 454R and 458R in Figure 1C). We speculate that 454R and 458R 3′ terminal sequences may influence the conformation of the active site of Polθ in such a way that results in lowering the enzyme's discrimination against NTPs. For example, the 5′-CCCAAACU-3′ motif is present in both substrates at or near the 3′ terminus and therefore may have this specific effect. Consistent with this idea, a recent study demonstrates that mutation of particular residues within the active site of Polθ, for example the

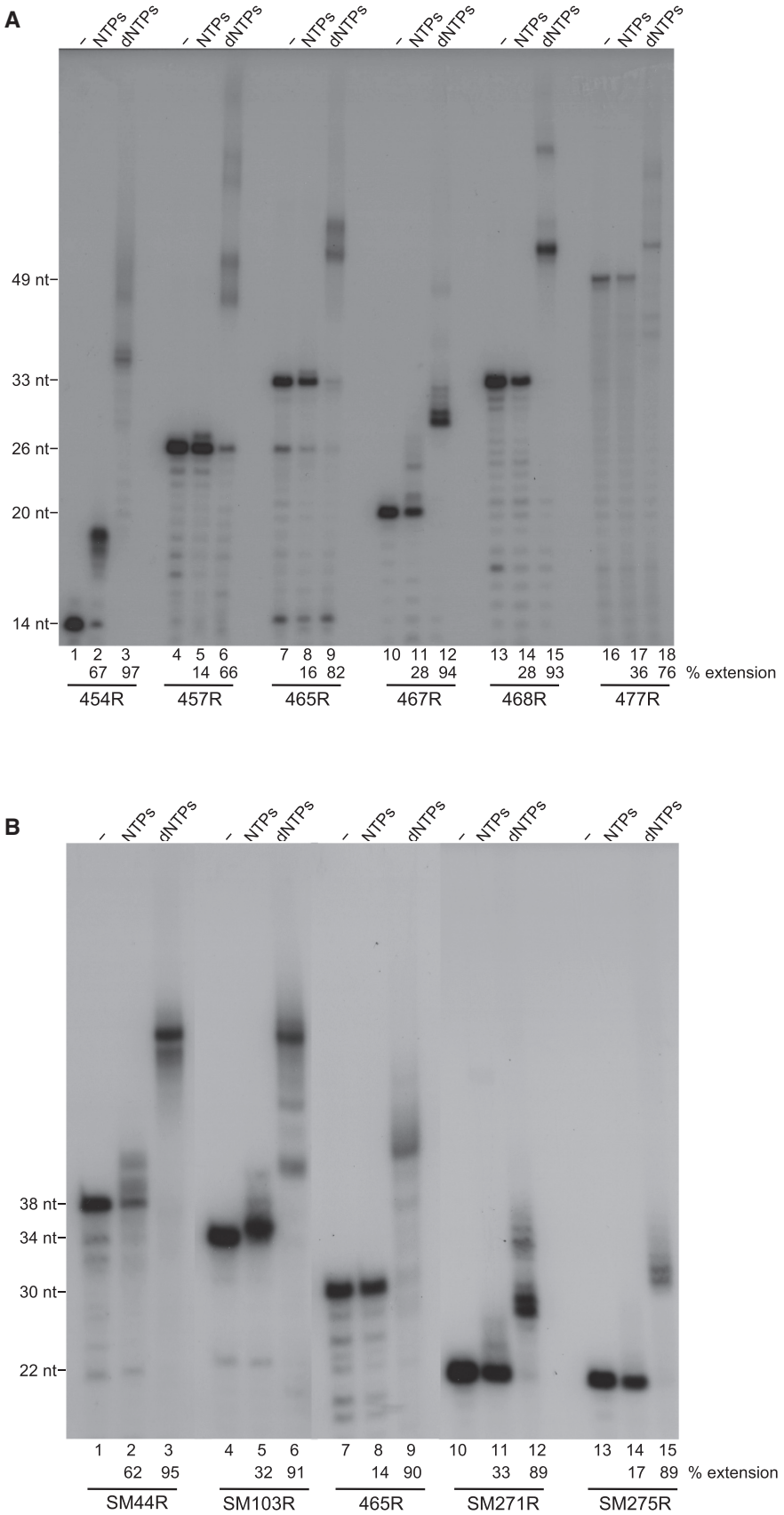

Figure 2. Polθ exhibits proficient RNA terminal 2′-deoxy nucleotidyl extension activity. (**A** and **B**) Denaturing gels showing Polθ RNA terminal nucleotidyl extension activity in the presence of Mn^{2+} on the indicated substrates with dNTPs or NTPs. Percentage extension is indicated.

steric gate, significantly enhances its ability to incorporate ribonucleotides, similar to other A-family Pols (23). Despite the relatively efficient extension of RNA substrates 454R and 458R in the presence of NTPs, overall our data in Figures 1 and 2 demonstrate that Polθ exhibits a strong preference for adding deoxyribonucleotides to RNA.

In Figure 3, we evaluated how RNA sequence influences individual 2′-deoxy and 2′-hydroxy nucleobase preferences during Polθ RNA 3′ terminal extension activity. Polθ 2′-deoxyribonucleotide addition was first analyzed on substrates 454R and 458R which are identical in sequence except for the 3′ terminal two nucleotides (see Table 1). The pattern of extension activity for individual 2′-deoxyribonucleotides was nearly identical (compare Figure 3A and B). For example, Polθ clearly shows a preference for adding deoxyguanosine monophosphate (dGMP), and a higher preference for extending with deoxycytidine monophosphate (dCMP) compared to deoxyadenosine monophosphate (dAMP) and deoxythymidine monophosphate (dTMP) (dG>>dC>dC>dA>dT)(Figure 3A and B). The preference for adding 15–20 consecutive dGMPs as observed in Figure 3A and B may be beneficial for applications involving enzymatic synthesis of DNA. We note that in our previous studies on Polθ DNA extension activity under similar conditions with $MnCl_2$, that the polymerase exhibited the following three different mechanisms: (i) a template-independent mechanism; (ii) a snap-back replication mode in which the polymerase utilizes the upstream portion of the substrate as a template *in cis* and; (iii) a template-dependent *in trans* mode in which the enzyme occasionally utilizes a separate substrate as a template to extend the 3′ terminus (9). We present the snap-back replication model for Polθ RNA 3′ terminal extension activity in Figure 3I. This model requires the 3′ terminal bases to transiently base pair with complementary bases near the 5′ region of the RNA substrate as shown previously for Polθ DNA 3′ terminal extension activity. Once Polθ promotes microhomology-mediated annealing of the 3′ and 5′ terminal bases, it is capable of extending the 3′ terminus by using the 5′ portion of the substrate as a template *in cis*. Consistent with the snap-back replication mode observed in our previous studies with DNA, we identified a potential base-pairing opportunity between the 3′ termini of substrates 454R and 458R (5′-CU-3′) and an internal site toward the 5′ end of these RNAs (5′-AG-3′). Figure 3J models this potential loop within the 454R substrate. This structure which is facilitated by microhmology annealing would allow Polθ to use the consecutive cytosines (CC) just upstream of the minimally base paired 3′ terminus as template bases to transfer dGMPs which show preferential activity (Figure 3J and 3A). MMEJ of the 3′ termini of DNA is a known function of Polθ, and this activity may facilitate the RNA extension activity observed herein (24). Importantly, supplementary data rule out a model in which Polθ utilizes a template *in trans* to extend the RNA 3′ terminus (Supplementary Figure S1C). Our previous research indicated that Polθ could perform multiple repetitive annealing and extension steps during snap-back replication, enabling efficient template-dependent DNA extension activity that depends on transient microhmology-mediated annealing and extension steps. Thus, we propose the same mechanism for Polθ RNA 3′ terminal extension activity in Figure 3I and J.

To further assess the potential mechanism of snap-back replication in facilitating Polθ RNA extension, we evaluated additional RNA substrates in Figure 3C and D. Although the RNA substrate in Figure 3C has the same two 3′ terminal bases as those in panels A and B, the preference for dNMP addition is different; dAMP and dTMP are preferentially added compared to dCMP and dGMP (i.e. A>T>>C>G) (compare Figure 3C with Figure 3A and B). A fourth RNA sequence was investigated which resulted in relatively inefficient extension overall and showed a preference for dTMP (i.e. T>C>G>A) (Figure 3D). Interestingly, this substrate was unable to support dAMP addition altogether; this is further discussed below (Figure 3D, lane 2).

Analysis of Polθ RNA extension activity on the same substrates in the presence of individual NTPs instead of dNTPs is presented in Figure 3 panels E–H. Consistent with the snap-back replication model (see Figure 3I and J), a strong preference for guanosine monophosphate (GMP) addition was observed for RNA substrates 454R and 458R (Figure 3E and F). Adenosine monophosphate (AMP) and uridine monophosphate (UMP) were added more efficiently than GMP and cytidine monophosphate (CMP) on RNA substrate 478R which is consistent with more effective addition of dAMP and dTMP on the same substrate shown in panel C (compare Figure 3G and C). SM275R showed relatively low extension activity overall, but demonstrated preferential addition of CMP and UMP which is consistent with the results in panel D in the presence of dCTP and deoxythymidine triphosphate (dTTP) (Figure 3H) (C>U>G>A).

To further test the snap-back replication model, we evaluated Polθ 3′ terminal extension on modified versions of SM275R that contain different sequences exclusively within their 5′ terminal region (Figure 4A–C). Figure 4A represents Polθ extension activity on the SM275 substrate initially shown in Figure 3D as a reference. As stated above, the snap-back replication model predicts that efficient RNA extension activity is dependent on the ability of the 3′ terminus to base pair with a complementary region within the 5′ portion of the same RNA substrate, and that the 5′ portion of the RNA substrate immediately downstream from the minimally base-paired 3′ terminus can act as a template *in cis* to encode for the incoming complementary nucleotide (Figure 3I). To further test this model, the 5′ portion of SM275 was substituted with 7 or 4 uridines in Figure 4B and C, respectively. Consistent with the model, we observed highly efficient addition of dAMP on both modified substrates as a result of the 5′ terminal uridine sequence tract just upstream from the 5′-CG-3′ motif which presents a base-pairing opportunity for the 3′ terminus (Figure 4D). In sharp contrast, no dAMP addition was observed on the original SM275 substrate, due to a lack of 5′ terminal uridines (Figure 4A). We presume that multiple transient annealing and extension cycles and/or slippage events between the extended 3′ terminus, consisting of adenosines, and the complementary 5′ uridine tract enable the long extension products observed in the presence of dATP (Figure 4B and C lane 2). dGMP addition was also enhanced on the modified sequence contexts versus the original SM275 substrate, likely due to misincor-

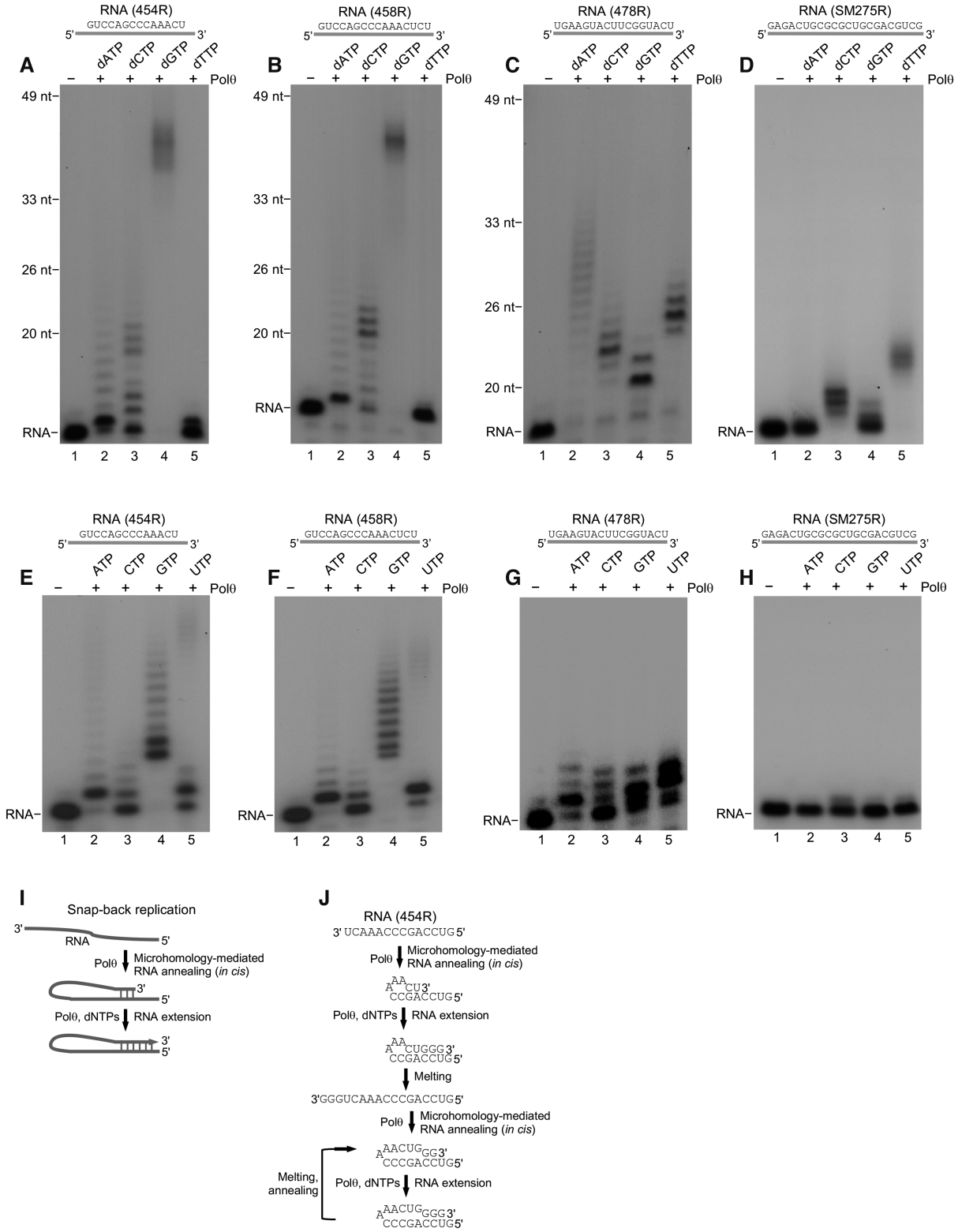

Figure 3. Sequence effects on Polθ RNA extension activity (**A–H**) Denaturing gels showing Polθ RNA extension activity in the presence of Mn^{2+} on the indicated substrates and with the indicated 2′-deoxy (A-D) or 2′-hydroxy (E-H) ribonucleotides. (**I**) Snap-back replication model of Polθ RNA extension activity. (**J**) Model of Polθ RNA extension activity via snap-back replication on substrate 454R.

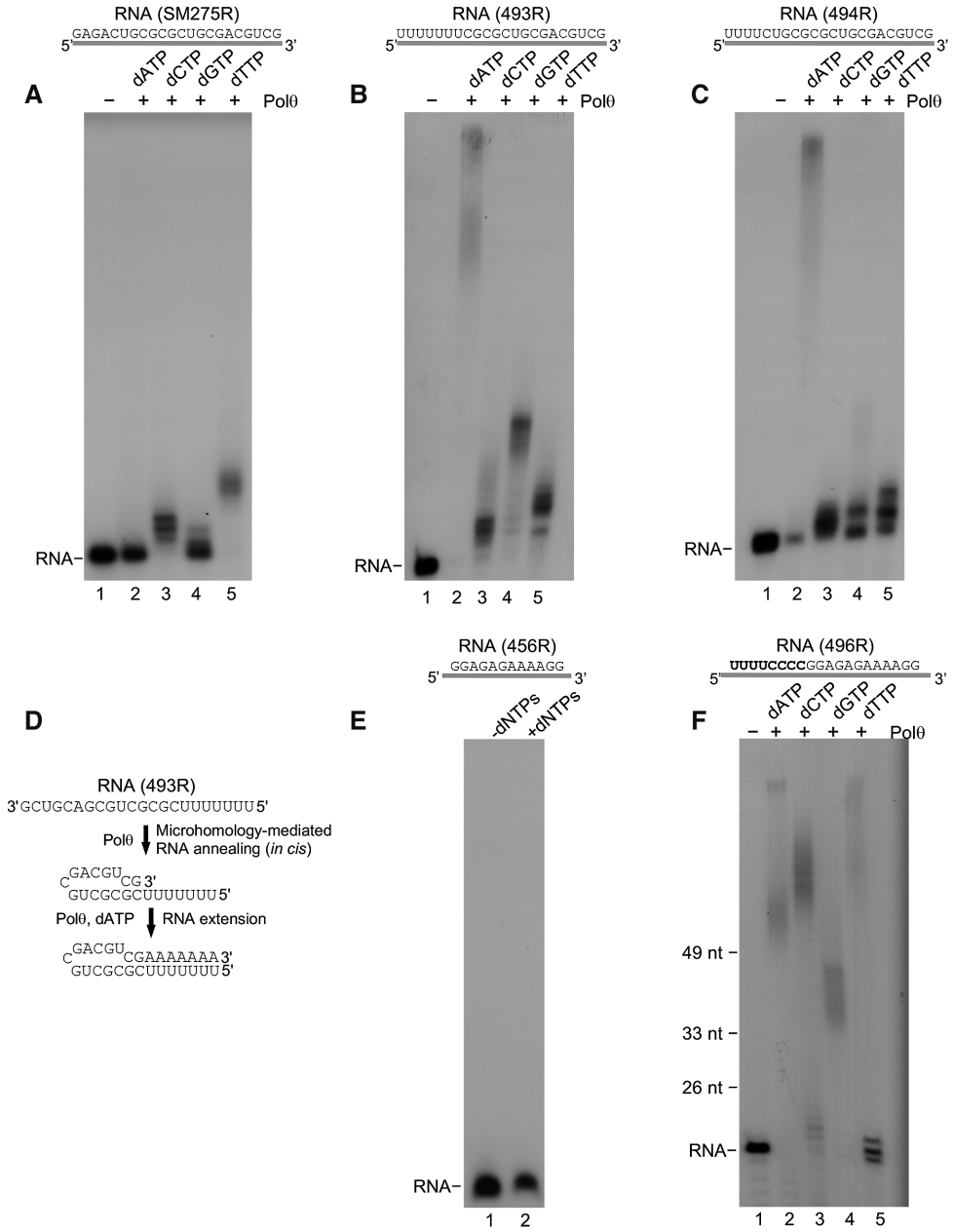

Figure 4. Polθ RNA 3′ terminal extension activity is directly influenced by the 5′ region of the RNA substrate. (**A–F**) Denaturing gels showing Polθ RNA 3′ terminal extension activity in the presence of Mn^{2+} on the indicated substrates and in the presence of the indicated dNTPs. D Snap-back replication model of Polθ RNA extension activity on substrate 493R.

poration of dGMP opposite the 5′ terminal uridine tracts as a result of wobble base-pairing (compare lanes 4 in Figure 4B and C to Figure 4A).

As further proof for the snap-replication model, we identified a substrate in which no Polθ RNA extension activity was observed (Figure 4E). We note that this RNA substrate is short (12 nt) and exclusively composed of guanine and adenine (homo-purine sequence) and therefore lacks any base-pairing opportunities between the 3′ terminus and 5′ terminal region. The addition of short cytidine and uridine tracts at the 5′ end of this substrate enabled robust Polθ 3′ terminal extension activity (Figure 4F). Here, dGMP and dAMP are efficiently added as a result of their hydrogen bonding potential opposite cytosine and uracil, respectively, at the 5′ terminus (Figure 4F, lanes 2 and 4). Efficient dCMP addition is also observed, likely due to misincorporation opposite cytosine or uracil (Figure 4F, lane 3).

Taken together, the data presented in Figures 3 and 4 strongly support that Polθ RNA extension activity is facilitated by a snap-back replication mechanism modeled in Figure 3I. We note that this mechanism requires Polθ to utilize ribonucleotides as template bases which is typically performed by RNA-dependent DNA polymerases, such as those encoded by retroviruses. However, because related A-

family polymerases lacking 3′-5′ exonuclease activity can exhibit robust RNA-dependent DNA polymerase activity, the ability of Polθ, which is devoid of exonuclease activity, to similarly utilize ribonucleotides as template bases is not inconceivable (25). Because Polθ exhibits relatively poor extension activity on double-strand DNA (9), we presume the enzyme is also inefficient on double-strand RNA. Based on this rationale, we speculate that Polθ RNA and DNA 3′ terminal extension activity becomes highly inefficient or even terminates upon fully extending the looped 3′ terminus to the 5′ end of the substrate (see Figure 3I). This mechanism provides an explanation for the biphasic extension activity observed in Figure 1G. For example, a rapid rate of RNA extension, likely due to snap-back replication, is followed by a slow and inefficient extension mechanism which is presumably due to the inability of Polθ to effectively perform extension of double-stranded substrates. Further studies are required to fully elucidate the mechanistic differences between the rapid and slow 3′ terminal extension processes identified herein.

Because the ability to modify the 3′ terminus of RNA is desirable for a variety of biotechnology and synthetic biology applications, we next examined the ability of Polθ to add to RNA different 2′-deoxy, 2′,3′-dideoxy and 2′-hydroxy ribonucleotide analogs with functional modifications such as fluorophores and attachment moieties (e.g. biotin, digoxigenin). The results show that Polθ efficiently adds to multiple RNA molecules a variety of 2′-deoxy and 2′,3′-dideoxy ribonucleotide analogs (Figure 5). For example, 2′-deoxy uridine nucleobases conjugated with biotin or digoxigenin are added efficiently to most of the RNA substrates (see lanes 4 and 5 in Figure 5B–F). The substrate in Figure 5G which contains a poly-uridine tract at its 5′ terminus, however, shows low efficiency of adding the 2′-deoxy uridine analog conjugated with digoxigenin (lane 4). Remarkably, several consecutive biotin and digoxigenin 2′-deoxy uridine nucleobases are added to RNA substrates 454R, 478R and 467R (Figure 5B, C and E). In contrast, only a few biotin and digoxigenin modified 2′-deoxyuridine nucleobases are added to RNA 457R SM275R (Figure 5D and F), and a single biotin conjugated dUMP is added to the 493R substrate (Figure 5G). Hence, the sequence context of the RNA also influences the ability of Polθ to add particular nucleotide analogs. Variable addition efficiency of a Cy3 conjugated 2′-deoxyuridine nucleotide analog is also observed (compare lane 3 in Figure 5B–G). We note that Polθ typically adds only one or two Cy3 conjugated dUMPs, likely due to the large fluorophore modification impeding further extension (lane 3 in Figure 5B–G). This is useful for biotechnology applications requiring Cy3 labeled RNA to avoid potential quenching as a result of multiple fluorophores in close proximity.

Because the controlled attachment of a single nucleobase with a functional chemical motif to RNA may be desirable for particular biotechnology and synthetic biology applications, we examined Polθ addition of 2′,3′-dideoxy uridine nucleotide analogs conjugated with different Atto fluorophores which exhibit high fluorescence quantum yield and high photostability. The results show that Polθ efficiently adds all three 2′,3′-dideoxyuridine nucleotide analogs, which demonstrates the ability to enzymatically modify RNA 3′ termini with a single chemical motif using a simple one step reaction (lanes 8–10, Figure 5B–G). We note the appearance of some lower molecular weight nucleotide transfer events in panels B and C (lanes 3–5 and 8–10), which is likely due to nucleotide impurities lacking the large functional motifs.

Next, we examined the ability of Polθ to add 2′-hydroxy ribonucleotide analogs to RNA, despite its preference for using 2′-deoxy ribonucleotides as substrates. The 2′-hydroxy 4-thio-uridine nucleotide analog was added with slightly lower efficiency compared to canonical ribonucleotides on all substrates (compare lanes 2 and 7, Figure 5B–G). In the case of N6-(6-Azido)hexyl-ATP which allows for ‘click’ chemistry—a widely used site-specific attachment moiety—the Polθ RNA extension reaction resulted in a faster migrating band on all RNA substrates (lane 6, Figure 5B–G). This is likely due to a net positive charge resulting from addition of this particular nucleotide analog. Hence, although Polθ shows a preference for adding 2′-deoxy ribonucleotides, certain 2′-hydroxy ribonucleotide analogs (e.g.N6-(6-Azido)hexyl-ATP) can be effectively used as substrates for Polθ RNA extension activity. Furthermore, particular mutations within the active site of Polθ can be introduced to dramatically lower the enzyme’s discrimination against ribonucleotides (23). Thus, we envisage that mutational engineering of Polθ will enable this enzyme to promote a wider variety of RNA based applications such as 3′ RNA terminal chemical modification demonstrated herein and potentially sequence-specific synthesis of synthetic RNA using reversible nucleotide chain terminator technology (26,27).

DISCUSSION

Because TdT is unable to act on RNA, Polθ exhibits a unique functionality as a RNA 3′ terminal extension enzyme that is capable of efficiently adding canonical nucleotides and a variety of nucleotide analogs to RNA. Importantly, Polθ is closely related to bacterial A-family Pol I enzymes (e.g. Klenow fragment), which have been extensively characterized by random and structure based mutagenesis (28). Based on these prior studies, we anticipate that Polθ can be further developed via site-directed mutagenesis to add an even wider assortment of ribonucleotide analogs to RNA 3′ terminal ends. Consistent with this idea, a recent study on Polθ demonstrated that mutation of conserved residues known to suppress incorporation of ribonucleoitdes in A-family Pols (i.e. steric gate) significantly increase Polθ terminal 2′-hydroxy nucleotidyl transferase activity on ssDNA (23). Further studies are required to determine whether similar mutations significantly increase Polθ 3′ terminal extension activity on RNA substrates.

Further development of Polθ as a RNA 3′ terminal polymerase may allow for large-scale and low cost enzymatic production of sequence-specific synthetic RNA using reversible ribonucleotide terminator technology (26,27). This potential function for Polθ would most likely require RNA 3′ terminal 2′-hydroxy nucleotidyl extension activity that is not strongly influenced by the sequence context of the RNA substrate. A recent study indicates that steric gate mutants

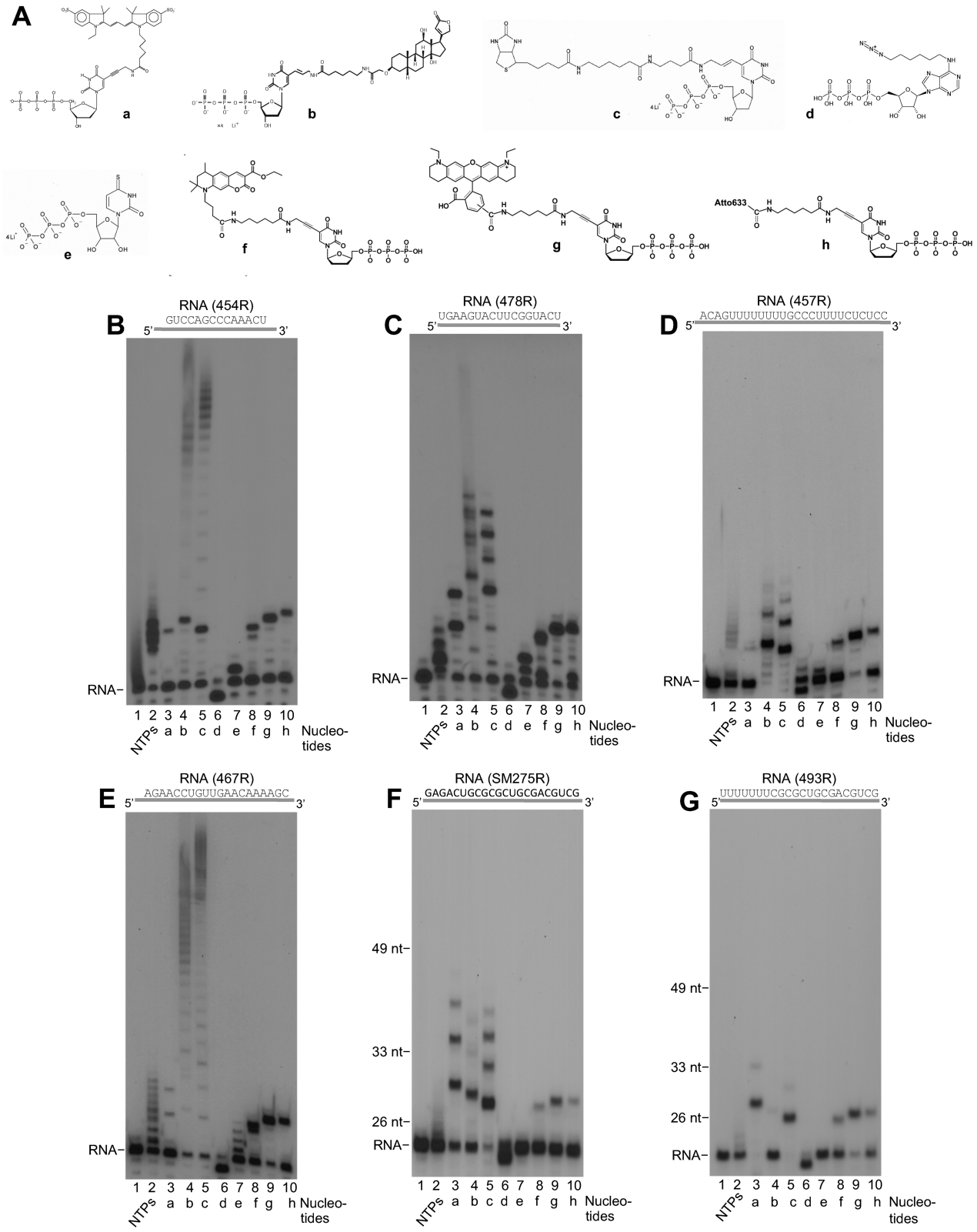

Figure 5. Polθ adds various types of nucleotide analogs to RNA 3′ termini. (**A**) Structures of nucleotide analogs. **a,** cy3-deoxyuridine triphosphate (dUTP); **b,** Digoxigenin-11-dUTP; **c,** Biotin-16AA-dUTP; **d,** N6-(6-Azido)hexyl-ATP; **e,** 4-Thio-UTP; **f,** Atto 425-11-ddUTP; **g,** Atto 565-11-ddUTP; **h,** Atto 633-11-ddUTP. (**B-G**) Denaturing gels showing Polθ RNA 3′ terminal extension activity on the indicated substrates and in the presence of Mn^{2+} and the indicated nucleotide analogs.

of Polθ are able to add ribonucleotides to DNA in a random manner (23). Yet, the results of our study indicate that RNA extension activity by wild-type Polθ is directly influenced by the RNA substrate sequence. Thus, it will be interesting to determine whether similar steric gate Polθ mutants or other variants are able to perform random ribonucleotide additions to RNA. Finally, we anticipate that rapid chemical modification of RNA via Polθ 3′ terminal addition of nucleotide analogs will enable new methods for isolating and/or quantitating various forms of cellular RNA.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

Author contributions: C.T., T.R. and T.A. performed experiments and interpreted data. T.K. and T.H. purified proteins. I.G. interpreted data and provided editorial input. R.T.P. designed experiments, interpreted data and wrote the manuscript.

FUNDING

National Institutes of Health [1R01GM115472-01 to R.T.P.]; Modification of 3′ Terminal Ends of Nucleic Acids by DNA Polymerase Theta (Patent) [WO2017075421A1]. Funding for open access charge: National Institutes of Health [1R01GM115472-01].

Conflict of interest statement. A patent entitled, Modification of 3′ Terminal Ends of Nucleic Acids by DNA Polymerase Theta (WO2017075421A1) has been filed.

REFERENCES

1. Black,S.J., Kashkina,E., Kent,T. and Pomerantz,R.T. (2016) DNA polymerase theta: a unique multifunctional End-Joining machine. *Genes (Basel)*, **7**, 67.
2. Kent,T.C., Chandramouly,G., McDevitt,S.M., Ozdemir,A.Y. and Pomerantz,R.T. (2015) Mechanism of microhomology-mediated end-joining promoted by human DNA polymerase θ. *Nat. Struct. Mol. Biol.*, **22**, 230–237.
3. Sfeir,A. and Symington,L.S. (2015) Microhomology-mediated end joining: a back-up survival mechanism or dedicated pathway? *Trends Biochem. Sci.*, **40**, 701–714.
4. Mateos-Gomez,P.A., Gong,F., Nair,N., Miller,K.M., Lazzerini-Denchi,E. and Sfeir,A. (2015) Mammalian polymerase theta promotes alternative NHEJ and suppresses recombination. *Nature*, **518**, 254–257.
5. Yousefzadeh,M.J., Wyatt,D.W., Takata,K., Mu,Y., Hensley,S.C., Tomida,J., Bylund,G.O., Doublie,S., Johansson,E., Ramsden,D.A. *et al.* (2014) Mechanism of suppression of chromosomal instability by DNA polymerase POLQ. *PLoS Genet.*, **10**, e1004654.
6. Chan,S.H., Yu,A.M. and McVey,M. (2010) Dual roles for DNA polymerase theta in alternative end-joining repair of double-strand breaks in Drosophila. *PLoS Genet.*, **6**, e1001005.
7. Koole,W., van Schendel,R., Karambelas,A.E., van Heteren,J.T., Okihara,K.L. and Tijsterman,M. (2014) A Polymerase Theta-dependent repair pathway suppresses extensive genomic instability at endogenous G4 DNA sites. *Nat. Commun.*, **5**, 3216.
8. Wood,R.D. and Doublie,S. (2016) DNA polymerase theta (POLQ), double-strand break repair, and cancer. *DNA Repair (Amst)*, **44**, 22–32.
9. Kent,T., Mateos-Gomez,P.A., Sfeir,A. and Pomerantz,R.T. (2016) Polymerase theta is a robust terminal transferase that oscillates between three different mechanisms during end-joining. *Elife*, **5**, e13740.
10. Hogg,M., Sauer-Eriksson,A.E. and Johansson,E. (2012) Promiscuous DNA synthesis by human DNA polymerase theta. *Nucleic Acids Res.*, **40**, 2611–2622.
11. Motea,E.A. and Berdis,A.J. (2010) Terminal deoxynucleotidyl transferase: the story of a misguided DNA polymerase. *Biochim. Biophys. Acta*, **1804**, 1151–1166.
12. Loc'h,J., Rosario,S. and Delarue,M. (2016) Structural basis for a new templated activity by Terminal deoxynucleotidyl Transferase: implications for V(D)J recombination. *Structure*, **24**, 1452–1463.
13. Jarchow-Choy,S.K., Krueger,A.T., Liu,H., Gao,J. and Kool,E.T. (2011) Fluorescent xDNA nucleotides as efficient substrates for a template-independent polymerase. *Nucleic Acids Res.*, **39**, 1586–1594.
14. Moon,A.F., Garcia-Diaz,M., Batra,V.K., Beard,W.A., Bebenek,K., Kunkel,T.A., Wilson,S.H. and Pedersen,L.C. (2007) The X family portrait: structural insights into biological functions of X family polymerases. *DNA Repair (Amst.)*, **6**, 1709–1725.
15. Dominguez,O., Ruiz,J.F., Lain de Lera,T., Garcia-Diaz,M., Gonzalez,M.A., Kirchhoff,T., Martinez,A.C., Bernad,A. and Blanco,L. (2000) DNA polymerase mu (Pol mu), homologous to TdT, could act as a DNA mutator in eukaryotic cells. *EMBO J.*, **19**, 1731–1742.
16. Ramadan,K., Maga,G., Shevelev,I.V., Villani,G., Blanco,L. and Hubscher,U. (2003) Human DNA polymerase lambda possesses terminal deoxyribonucleotidyl transferase activity and can elongate RNA primers: implications for novel functions. *J. Mol. Biol.*, **328**, 63–72.
17. Bertocci,B., De Smet,A., Berek,C., Weill,J.C. and Reynaud,C.A. (2003) Immunoglobulin kappa light chain gene rearrangement is impaired in mice deficient for DNA polymerase mu. *Immunity*, **19**, 203–211.
18. Bertocci,B., De Smet,A., Weill,J.C. and Reynaud,C.A. (2006) Nonoverlapping functions of DNA polymerases mu, lambda, and terminal deoxynucleotidyltransferase during immunoglobulin V(D)J recombination in vivo. *Immunity*, **25**, 31–41.
19. Laishram,R.S. (2014) Poly(A) polymerase (PAP) diversity in gene expression–star-PAP vs canonical PAP. *FEBS Lett.*, **588**, 2185–2197.
20. Hall,T.M. (2002) Poly(A) tail synthesis and regulation: recent structural insights. *Curr. Opin. Struct. Biol.*, **12**, 82–88.
21. Martin,G. and Keller,W. (2007) RNA-specific ribonucleotidyl transferases. *RNA*, **13**, 1834–1849.
22. Seki,M., Marini,F. and Wood,R.D. (2003) POLQ (Pol theta), a DNA polymerase and DNA-dependent ATPase in human cells. *Nucleic Acids Res.*, **31**, 6117–6126.
23. Randrianjatovo-Gbalou,I., Rosario,S., Sismeiro,O., Varet,H., Legendre,R., Coppee,J.Y., Huteau,V., Pochet,S. and Delarue,M. (2018) Enzymatic synthesis of random sequences of RNA and RNA analogues by DNA polymerase theta mutants for the generation of aptamer libraries. *Nucleic Acids Res.*, **46**, 6271–6284.
24. Kent,T., Chandramouly,G., McDevitt,S.M., Ozdemir,A.Y. and Pomerantz,R.T. (2015) Mechanism of microhomology-mediated end-joining promoted by human DNA polymerase theta. *Nat. Struct. Mol. Biol.*, **22**, 230–237.
25. Shi,C., Shen,X., Niu,S. and Ma,C. (2015) Innate reverse transcriptase activity of DNA polymerase for isothermal RNA direct detection. *J. Am. Chem. Soc.*, **137**, 13804–13806.
26. Gardner,A.F., Wang,J., Wu,W., Karouby,J., Li,H., Stupi,B.P., Jack,W.E., Hersh,M.N. and Metzker,M.L. (2012) Rapid incorporation kinetics and improved fidelity of a novel class of 3′-OH unblocked reversible terminators. *Nucleic Acids Res.*, **40**, 7404–7415.
27. Guo,J., Xu,N., Li,Z., Zhang,S., Wu,J., Kim,D.H., Sano Marma,M., Meng,Q., Cao,H., Li,X. *et al.* (2008) Four-color DNA sequencing with 3′-O-modified nucleotide reversible terminators and chemically cleavable fluorescent dideoxynucleotides. *Proc. Natl. Acad. Sci. U.S.A.*, **105**, 9145–9150.
28. Loh,E. and Loeb,L.A. (2005) Mutability of DNA polymerase I: implications for the creation of mutant DNA polymerases. *DNA Repair (Amst.)*, **4**, 1390–1398.