



Cite this: *Chem. Commun.*, 2016, 52, 11905

Received 23rd August 2016,  
Accepted 5th September 2016

DOI: 10.1039/c6cc06925h

www.rsc.org/chemcomm

## Effect of terminal 3'-hydroxymethyl modification of an RNA primer on nonenzymatic primer extension†

Ayan Pal,<sup>‡,ab</sup> Rajat S. Das,<sup>‡,cd</sup> Weicheng Zhang,<sup>‡,ae</sup> Megan Lang,<sup>c</sup> Larry W. McLaughlin<sup>c</sup> and Jack W. Szostak<sup>\*,abe</sup>

**The significance of the precise position of the hydroxyl at the 3'-end of an RNA primer for nonenzymatic template-directed primer extension is not well understood. We show that an RNA primer terminating in 3'-hydroxymethyl-2',3'-dideoxy-guanosine has greatly diminished activity, suggesting that the spatial preorganization of the terminal sugar contributes significantly to the efficiency of primer extension.**

Template-directed primer extension with chemically activated ribonucleotide monomers, such as nucleotide 5'-phosphorimidazolides, has been extensively studied as a model for nonenzymatic RNA replication during the origin of life.<sup>1–4</sup> In this model, activated nucleotide monomers first interact with the RNA primer–template complex by binding reversibly to complementary sites on the template. Extension of the primer occurs by nucleophilic attack of either the 2'- or 3'-hydroxyl of the 3'-terminal nucleotide of the primer on the activated 5'-phosphate of an adjacent template-bound monomer. A 2'-5' or 3'-5' phosphodiester bond is formed when the nucleophile displaces the leaving group on the phosphate of the activated monomer. This reaction is accelerated by the catalytic assistance of a divalent cation, such as Mg<sup>2+</sup> or Mn<sup>2+</sup>, although it is not known whether the metal ion interacts with the nucleophilic hydroxyl, the phosphate, or both. Various leaving groups on the phosphate of the incoming monomers, such as imidazole,<sup>5</sup> 2-methylimidazole,<sup>6</sup> and oxazabenzotriazole,<sup>7</sup> have been studied

in the primer extension reaction. Oligomerization on mineral surfaces has also been carried out using 1-methyladenine<sup>8</sup> as the leaving group. Markedly improved polymerization, in terms of both rate and 3'-5' regioselectivity, is achieved with 2-methylimidazole<sup>9</sup> as the leaving group at a pH of 8 to 8.5, although hydroxyazabenzotriazole gives faster rates at pH 8.9–9.5.<sup>10</sup>

Previous studies have demonstrated that primer extension is more efficient on A-form templates such as RNA compared to B-form templates such as DNA.<sup>11,12</sup> Moreover, our laboratory has shown that the sugar pucker of activated ribonucleotides switches from mostly 2'-endo in the free state to mostly 3'-endo upon binding to an RNA template.<sup>13</sup> These observations suggest that an A-form primer–template complexed with activated ribonucleotides monomers in the 3'-endo conformation represents the optimal conformation for in-line attack of the nucleophile at the 3' end of the primer on the 5'-phosphate of the adjacent activated monomer. A better nucleophile, such as the 3'-amine of N3'–P5'-linked phosphoramidate DNA (3'-NP-DNA), is known to improve the kinetics of the extension reaction.<sup>14,15</sup> However, the importance of the structural position and rigidity of the nucleophile at the 3'-position has not been systematically investigated. Based on the fact that both the 2' and 3' hydroxyls are able to attack the phosphate of the incoming nucleotide (to an extent controlled by leaving group and metal ion identity), it might seem that the precise position of the nucleophile is not critical. Furthermore, primers ending in a 2'-deoxynucleotide can still be extended at about 10% of the rate of primer ending in a ribonucleotides,<sup>16</sup> despite the higher pK<sub>a</sub> of the hydroxyl (~14 vs. 12.5).<sup>16,17</sup> The intrinsic sugar preference for the 2'-endo conformation could be countered by the cooperative induction of A-form helical geometry by the remainder of the RNA primer.<sup>18</sup> On the other hand, alritol and hexitol monomers, which have different sugar moieties, are poor substrates for continued primer extension, although it is unclear whether this reflects poor positioning of the attacking nucleophile or the electrophilic phosphate.<sup>19</sup>

To explore the structural requirements for the placement of the nucleophile, we have replaced the terminal ribonucleotide G of a primer by 3'-hydroxymethyl-2',3'-dideoxy-guanosine,

<sup>a</sup> Howard Hughes Medical Institute, Department of Molecular Biology and Center for Computational and Integrative Biology, Massachusetts General Hospital, 185 Cambridge Street, Boston, Massachusetts 02114, USA.

E-mail: szostak@molbio.mgh.harvard.edu

<sup>b</sup> Department of Genetics, Harvard Medical School, 77 Avenue Louis Pasteur, Boston, Massachusetts 02115, USA

<sup>c</sup> Boston College, Department of Chemistry, Merkert Chemistry Center, 2609 Beacon Street, Chestnut Hill, Massachusetts 02467, USA

<sup>d</sup> Alnylam Pharmaceuticals, 300 Third Street, Cambridge, Massachusetts 02142, USA

<sup>e</sup> Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford St., Cambridge, Massachusetts 02138, USA

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c6cc06925h

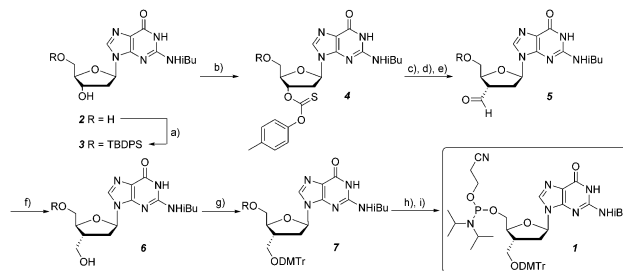
‡ These authors contributed equally to this work.



represented here as dG\*. Different arguments suggest that this modified 3'-terminus could either enhance or interfere with primer extension. The hydroxymethyl group of dG\*, a primary alcohol, should be a better nucleophile than the 3'-hydroxyl of a 2'-deoxynucleotide, because of its lower  $pK_a$  and decreased steric hindrance (although hydrogen bonding of the hydroxyl to a phosphate oxygen could increase its  $pK_a$  as well as causing steric problems). On the other hand, the displacement of the hydroxyl of dG\* away from the sugar, and its greater conformational flexibility, could decrease the rate of primer extension if precise spatial positioning is a critical factor in the reaction.

In order to investigate the effects of structural perturbation at the 3'-position on the primer extension reaction, we first had to prepare an RNA primer terminating in dG\*. Here we describe the synthesis of 5'-( $\beta$ -cyanoethyl) phosphoramidite of 3'-hydroxymethyl-2',3'-dideoxy-guanosine and its use in the reverse solid-phase preparation of a primer terminating with dG\* (Fig. 1). Several previous reports have described the synthesis of nucleotide analogs with an extended sugar phosphate backbone, typically beginning with a modified abasic sugar or carbocycle<sup>20–25</sup> followed by glycosylation to incorporate the purine or pyrimidine nucleobase. This approach leads to the formation of unwanted regioisomers that are difficult to separate.<sup>26</sup> In addition, maintenance of the correct stereochemistry at the 3'-carbon has been difficult.<sup>27,28</sup> In our study, the 3'-hydroxymethyl-dG phosphoramidite was synthesized, purified, and characterized by a modification of the procedure originally reported by Mesmaeker *et al.*<sup>29</sup> that avoids these problems by 3'-homologation of 2'-deoxyguanosine (Scheme 1).

To ensure the stereoselective introduction of the 3'- $\alpha$ -hydroxymethyl functionality, the  $\beta$ -face of the sugar ring was blocked by both the bulky *tert*-butyl diphenylsilyl group and the guanine nucleobase, so that the attack of an incoming nucleophile would occur predominantly from the  $\alpha$ -face. The 3'-hydroxyl was converted into its corresponding *o*-(*p*-tolyl)thiocarbonate ester **4**. The ester **4** was reacted with  $\beta$ -tributylstannylstyrene in the presence of AIBN to form a new C–C bond *via*  $\beta$ -elimination of the tributylstannyl group. Since the  $\beta$ -face of the sugar ring is

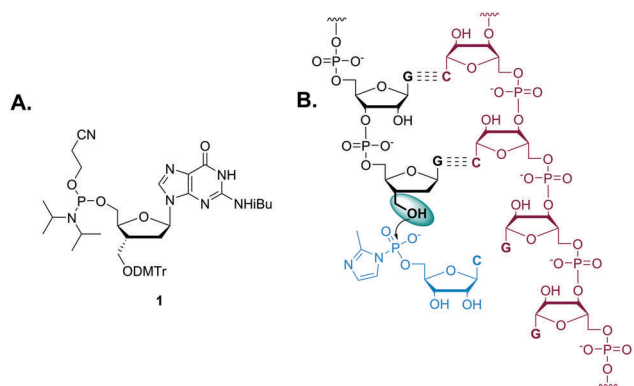


**Scheme 1** Synthesis of phosphoramidite. Reaction conditions: (a) TBDPSCl, DMAP, pyr, 20 °C, 36 h, 92%; (b) *o*-(*p*-tolyl)-chlorothionoformate, DMAP, ACN, 20 °C, 1 h, 94%; (c)  $\beta$ -tributylstannylstyrene, AIBN, toluene, 20 to 80 °C, 48 h; (d) OsO<sub>4</sub>, NMO, dioxane, 20 °C, 2 h; (e) NaIO<sub>4</sub>, 20 °C, 3 h, 28% over 3 steps; (f) NaBH<sub>4</sub>, MeOH, 0 to 20 °C, 30 min, 94%; (g) DMTrCl, DMAP, pyr, 20 °C, 12 h, 93%; (h) TBAF, THF, 20 °C, 8 h; (i) 2-cyanoethyl *N,N,N,N*-tetraisopropylphosphoramidite, tetrazole, DIEA, DCM, 20 °C, 2 h, 89% over 2 steps.

sterically blocked, styryl addition occurs preferentially through the  $\alpha$ -face. The crude reaction mixture was filtered through a short silica pad to remove excess reagents and dried. The resulting crude product was used directly in the next step. The styryl group was cleaved by dihydroxylation using osmium tetroxide, followed by oxidation of the diol with sodium periodate to generate a 3'- $\alpha$ -C-formyl group in **5**. The aldehyde of **5** was efficiently reduced to yield the hydroxymethyl group of **6** by sodium borohydride. This 3'- $\alpha$ -C-hydroxymethyl group was then protected with DMTr, and the 5'-alcohol of **7** retrieved by desilylation, followed by conversion to phosphoramidite **1** in excellent yield. This reverse amidite **1** was incorporated as the terminal residue of an RNA oligomer by reverse solid-phase synthesis. The oligomer was purified by HPLC and characterized by LC-MS. The dG\*-terminated oligomer obtained was at greater than 97% purity with a trace of N-1 oligomer.

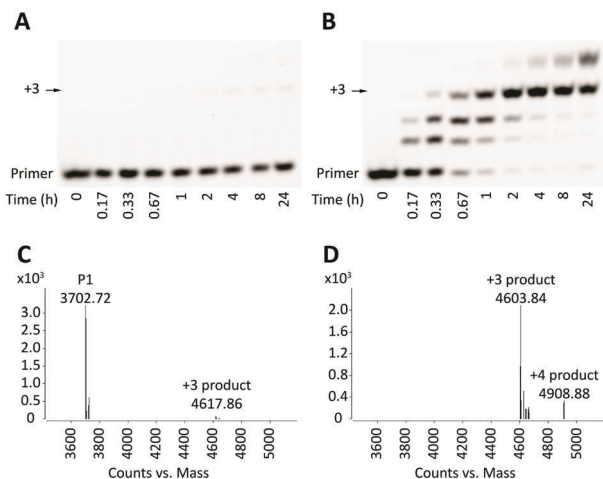
To examine the effect of dG\* on template-directed non-enzymatic primer extension, we used a 5'-fluorophore tagged dG\*-terminated RNA primer (P1, 5'-Cy3-GACUGACUGdG\*-3') complexed with a complementary RNA template strand (T, 3'-CUGACUGACCGGGAA-5') containing a homopolymeric G4 site to be copied by the activated monomer, cytidine-5'-phosphor-(2-methyl)imidazole (2-MeImpC). As a control, a 5'-fluorophore tagged dG-terminated RNA primer (P2, 5'-Cy3-GACUGACUGdG-3') was also prepared and time courses of primer extension with the two primers were compared.

To determine the optimal conditions for the primer extension reaction, a series of experiments were performed with 1  $\mu$ M primer, 5  $\mu$ M template, and 100 mM of different divalent cations including Mg<sup>2+</sup>, Mn<sup>2+</sup> and Ca<sup>2+</sup>, in the presence or absence of monovalent cation (Na<sup>+</sup> or K<sup>+</sup>) and in the presence of 50 mM of activated monomers and 200 mM buffer (pH from 5.5 to 8.5). Surprisingly, we failed to observe any significant primer extension from the dG\* primer P1 even after 1 day under most conditions. The best condition, with 100 mM Mg<sup>2+</sup> at pH 7.5, yielded less than 5% extended products after 24 hours, with +3 extended product being the predominant contributor (Fig. 2A). The  $k_{obs}$  was approximately 0.012 h<sup>-1</sup>. We were unable to detect any +1 or +2 extended products, most likely because after the



**Fig. 1** (A) Structure of 5'-( $\beta$ -cyanoethyl) phosphoramidite of 3'-hydroxymethyl-2',3'-dideoxy-guanosine. (B) Schematic representation of template-directed extension of 3'-hydroxymethyl-2',3'-dideoxy-guanosine-terminated primer in the presence of 2-MeImpC.





**Fig. 2** Gel electrophoresis and LC-MS analysis of primer extension reactions. Reaction conditions: 1  $\mu\text{M}$  primer, 5  $\mu\text{M}$  template T, 200 mM HEPES pH 7.5, 100 mM  $\text{MgCl}_2$ , 50 mM 2-MelmpC. (A) and (B) Gel analysis of primer extension of primers P1 and P2 respectively; (C) and (D) LC-MS analysis of primer extension of primers P1 and P2 after 24 hours respectively.

very slow addition of the first monomer, addition of the following two residues is fast as the terminal residue of the +1 or +2 extended products is then a standard ribonucleotide. On the other hand, P2 extended to varying extents depending on pH, metal ions and time (Fig. 2B). At pH 7.5 with  $\text{Mg}^{2+}$ , full-length +3 products appeared after 2 hours with  $k_{\text{obs}}$  of  $\sim 2.8 \text{ h}^{-1}$ . After 24 hours, more than 95% of the P2 was extended to +3 or +4, in striking contrast to the minimal extent of primer extension of P1.

To confirm that the small amount of P1 +3 product observed by PAGE analysis did indeed result from extension from the  $\text{dG}^*$  residue (and not from trace levels of N-1 oligomer), we used LC-MS to analyse the 24 hour time point of the primer extension reaction. While unreacted P1 was the predominant species (Fig. 2C), approximately 3% of primer extended to +3 product (Fig. 2C; see also ESI,† Fig. S1), consistent with the results of the gel electrophoresis analysis. The mass of the P1 +3 product was 14 daltons greater than that of the P2 +3 product, as expected from replacement of the primer 3'-hydroxyl with a 3'-hydroxymethyl group.

In conclusion, template-directed nonenzymatic extension of a primer with a 3'-terminal  $\text{dG}^*$  proceeds very slowly, at a rate less than 1% that of a primer with a terminal dG residue. We suggest that two major factors may cause the slow primer extension from  $\text{dG}^*$ . First, the location of the hydroxyl in  $\text{dG}^*$  is  $\sim 1.5 \text{ \AA}$  from the position of the normal 3'-hydroxyl. This altered location may make nucleophilic attack of the hydroxyl on the phosphate of the incoming activated nucleotide difficult either due to the altered position *per se* or due to possible hydrogen bonding of the hydroxyl to a phosphate oxygen, or

may hinder coordination with the catalytic  $\text{Mg}^{2+}$  ion. Second, the conformational flexibility of the primary hydroxyl group may decrease the probability of the required spatial positioning of the nucleophile for an in-line attack on the 5'-phosphate of the activated monomer. Our results suggest that the normal location of the nucleophile, on the 3'-carbon of the ribose or deoxyribose sugar, is required for efficient nonenzymatic template directed primer extension.

We thank Drs Victor Lelyveld, Wen Zhang, Noam Prywes, Neha Kamat and Mr Travis Walton for helpful discussions. J. W. S. is an Investigator of the Howard Hughes Medical Institute. This work was supported in part by a grant (290363) from the Simons Foundation to J. W. S.

## Notes and references

- 1 L. E. Orgel, *Crit. Rev. Biochem. Mol. Biol.*, 2004, **39**, 99.
- 2 T. Wu and L. E. Orgel, *J. Am. Chem. Soc.*, 1992, **114**, 5496.
- 3 T. Wu and L. E. Orgel, *J. Am. Chem. Soc.*, 1992, **114**, 7963.
- 4 S. S. Mansy, J. P. Schrum, M. Krishnamurthy, S. Tobé, D. A. Treco and J. W. Szostak, *Nature*, 2008, **454**, 122.
- 5 R. Lohrmann and L. E. Orgel, *Tetrahedron*, 1978, **34**, 853.
- 6 G. F. Joyce, T. Inoue and L. E. Orgel, *J. Mol. Biol.*, 1984, **176**, 279.
- 7 P. Hagenbuch, E. Kervio, A. Hochgesand, U. Plutowski and C. Richert, *Angew. Chem., Int. Ed. Engl.*, 2005, **44**, 6588.
- 8 K. J. Prabahar and J. P. Ferris, *J. Am. Chem. Soc.*, 1997, **119**, 4330.
- 9 T. Inoue and L. E. Orgel, *J. Mol. Biol.*, 1982, **162**, 201.
- 10 S. R. Vogel, C. Deck and C. Richert, *Chem. Commun.*, 2005, 4922.
- 11 M. Kurz, K. Göbel, C. Hartel and M. W. Göbel, *Angew. Chem., Int. Ed. Engl.*, 1997, **36**, 842.
- 12 M. Kurz, K. Göbel, C. Hartel and M. W. Göbel, *Helv. Chim. Acta*, 1998, **81**, 1156.
- 13 N. Zhang, S. Zhang and J. W. Szostak, *J. Am. Chem. Soc.*, 2012, **134**, 3691.
- 14 S. Zhang, N. Zhang, J. C. Blain and J. W. Szostak, *J. Am. Chem. Soc.*, 2013, **135**, 924.
- 15 S. Zhang, J. C. Blain, D. Zielinska, S. M. Gryaznov and J. W. Szostak, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 17732.
- 16 T. Wu and L. E. Orgel, *J. Am. Chem. Soc.*, 1992, **114**, 317.
- 17 G. I. Birnbaum, J. Giziewicz, C. P. Huber and D. Shugar, *J. Am. Chem. Soc.*, 1976, **98**, 4640.
- 18 M. C. Wahl and M. Sundaralingam, *Nucleic Acids Res.*, 2000, **28**, 4356.
- 19 I. A. Kozlov, M. Zielinski, B. Allart, L. Kerremans, A. Van Aerschot, R. Busson, P. Herdewijn and L. E. Orgel, *Chem. – Eur. J.*, 2000, **6**, 151.
- 20 H. Boumchita, M. Legraverend and E. Bisagni, *Heterocycles*, 1991, **32**, 1785.
- 21 J. Wachtmeister, B. Classon, I. Kvarnström and B. Samuelsson, *Nucleosides Nucleotides*, 1997, **16**, 809.
- 22 J. Wachtmeister, A. Mühlman, B. Classon and B. Samuelsson, *Tetrahedron*, 1999, **55**, 10761.
- 23 H. R. Moon, H. O. Kim, M. W. Chun, L. S. Jeong and V. E. Marquez, *J. Org. Chem.*, 1999, **64**, 4733.
- 24 J. Wachtmeister, C. Björn, S. Bertil and I. Kvarnström, *Tetrahedron*, 1997, **53**, 1861.
- 25 M. Janson, L. Svansson, S. C. T. Svensson, I. Kvarnström, B. Classon and B. Samuelsson, *Nucleosides Nucleotides*, 1992, **11**, 1739.
- 26 G. S. Buenger and V. E. Marquez, *Tetrahedron Lett.*, 1992, **33**, 3707.
- 27 T. Kofoed, P. B. Rasmussen, P. Valentin-Hansen and E. B. Pedersen, *Acta Chem. Scand.*, 1997, **51**, 318.
- 28 J. J. Vasseur, F. Debart, Y. S. Sanghvi and P. D. Cook, *J. Am. Chem. Soc.*, 1992, **114**, 4006.
- 29 Y. S. Sanghvi, R. Bharadwaj, F. Debart and A. De Mesmaeker, *Synthesis*, 1994, 1163.

