Adenosine residues in the template do not block spontaneous replication steps of RNA†

Stephanie R. Vogel and Clemens Richert*

Received (in Cambridge, UK) 22nd February 2007, Accepted 20th March 2007
First published as an Advance Article on the web 12th April 2007
DOI: 10.1039/b702768k

Sub-freezing temperatures, azabenzotriazolide activation, multiple monomer addition, and helper displacement help to overcome what seemed like an intrinsic block of adenine-templated RNA replication steps in the absence of enzymes.

The extension of a primer strand by individual nucleotides, directed by the sequence of a template, is the fundamental reaction of replication. In today's cells, this reaction requires catalysis by polymerases. At some point in prebiotic evolution, replication must have begun. It is likely that RNA was the first encoding system, as it also forms catalysts (ribozymes) for a primitive metabolism, and the "RNA world" hypothesis dominates the current debate on prebiotic evolution.[1] Whether RNA can undergo spontaneous replication is unclear, though. Elegant work demonstrated that RNA-based replication steps can indeed occur spontaneously with chemically activated nucleotides in the absence of enzymes.[2–6] Further, spontaneous oligomerization of ribonucleotides, catalyzed by minerals, has been shown to generate RNA strands long enough to serve as templates and primers.[7]

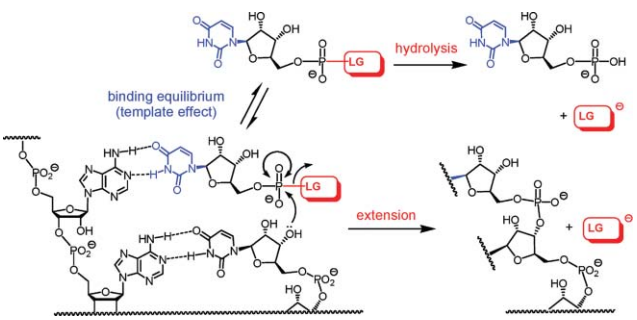
Severe difficulties remain, though, that complicate enzyme-free replication, and the chance of demonstrating replication in an all-RNA system has been called "remote".[8] A number of sequences do not template well. It has been reported that the template must contain at least 60% cytidylic acid residues, and that no part may contain more than one adenosine or thymidine nucleotide, which must be separated by several cytidines, for continuous copying steps to occur.[3,9] Semi-conservative replication requires that one nucleobase constitutes no more than 50% of a template, however. Sequences like TT,‡ GT and TG appear to be partial barriers and sequences like AT, TA, AA, GA and AG have been described as total barriers to spontaneous formation of complementary strands.[2,10,11] Also, stretches of guanosines may only be copied if the formation of self-structures of the template can be suppressed.[10a,12]

Given the difficulties with RNA itself, most recent studies have focused on structural analogs. Diaminopurine, an adenine analog forming three hydrogen bonds,[13] has been shown to promote copying more successfully than adenine,[14,15] particularly when combined with 5-propynyluracil as templating base,[6] but full replication has remained elusive.[16] Distantly related analogs of RNA have also been tested as templates for enzyme-free replication steps, including peptide nucleic acids, hexitol nucleic acids, and altritol nucleic acids.[17–20] Ligation experiments with diaminopurine-containing oligomers of RNA or TNA (an RNA analog with a shortened backbone)[21] are promising,[22] but ligation-based replication will require a large set of building blocks for diverse sequences.

It is interesting to ask what the physical basis of the "block" is that weakly base-pairing residues appear to induce in copying reactions. Since there is little reason to believe that nucleobases like adenine actively inhibit the esterification reaction underlying primer extension (Scheme 1), reduced templating efficiency is a likely cause. We have recently shown that spontaneous replication steps involving adenine or thymine as templating base can occur with near-quantitative yield in DNA-based systems containing a highly reactive 3′-amino primer and a "helper oligonucleotide" that binds downstream of the templating base.[23] The strength of base pairing should be roughly the same for DNA- and RNA-based duplexes, and so, there should be no intrinsic block to RNA replication steps templated by adenine. We were able to show that an acceleration in an RNA-based system can be achieved with azaoxybenzotriazolides of ribonucleotides as monomers, but the yields for adenine-templated reactions remained below 40%.[24] Here we show that such replication steps can be high yielding and that the dinucleotide AA is not a total block to replication.

In the current study, we focused on a template containing three consecutive adenosines. Stretches of adenosine are known to be particularly problematic for chemical self-replicating systems.[25] Since hydrolysis of activated nucleotides competes with primer extension, we hoped to increase relative rates of the desired (primer extension) *vs.* undesired reaction (hydrolysis) with azaoxybenzotriazolides as leaving groups for the activated nucleotides[24] (Fig. 1).

First, we focused on enhancing the yield of a single extension step templated by adenine (Scheme 2). Primer **1**, which is complementary to the 3′-terminal portion of template **2**, was reacted with activated mononucleotides in the presence of downstream-binding "micro helper" **3**. Primer extension was monitored by MALDI-TOF mass spectrometry (see ESI† for details). Activated ribothymidine or uridine monophosphate (**4**/**5**) were employed as



Scheme 1 Primer extension by a uridylic acid residue or hydrolysis.

Institute for Organic Chemistry, University of Karlsruhe (TH), 76131 Karlsruhe, Germany. E-mail: cr@rrg.uka.de; Fax: +49-721 608 4825; Tel: +49-721 608 2860

† Electronic supplementary information (ESI) available: Protocols and kinetic data. See DOI: 10.1039/b702768k

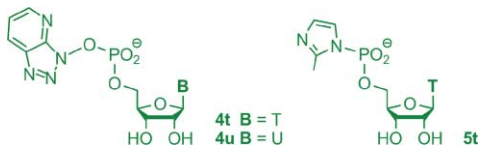
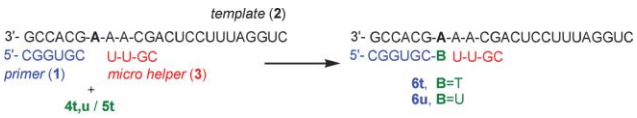


Fig. 1 Structures of activated monomers.



Scheme 2 Primer extension templated by an adenosine residue.

monomers. The former gives products easier to distinguish mass spectrometrically in studies with monomer mixtures,[24] and engages in slightly stronger stacking interactions. A half-life time of **4t** in aqueous solution has been reported.[24]

When methylimidazolide **5t** was reacted with **1/2** at 50 mM monomer concentration and 0 °C (conditions suggested by the literature),[26] <20% primer extension to **6t** was detected after 12.5 d (Table 1). Further, no multiple extension of primer **1** was observed. Conversion increased to 26% in the presence of **3**, but the $t_{1/2}$ remained >1000 h. More favorable results were observed with azaoxybenzotriazolides **4t** or **4u**. In our earlier work, at room temperature, **4t** gave <40% primer conversion for adenine as templating base, even when fresh aliquots of monomer were added after the initial reaction phase.[24] At 4 °C, 5 mM **4u** was sufficient to induce up to 45% extension of primer **1**. When the monomer concentration was raised to 20 mM and ribothymidine monomer **4t** was employed, up to 78% conversion of **1** to **6t** was achieved in the presence of **3**.

Cooling the reaction mixture to −20 °C, combined with a lower monomer concentration for the first reaction phase was tested next. Ligation,[27] oligomerization,[28] and template-directed polymerization[29] have been performed below 0 °C in the past, though not at quite so low a temperature. At −20 °C, the solidified mixture containing **3** gave between 27% and 92% primer extension at monomer concentrations between 1 mM and 5 mM (Table 1). Also, at 5 mM monomer, rate constants were surprisingly high at −20 °C, suggesting that decreases in reactivity can be compensated

Table 1 Effect of leaving group, concentration of monomer, and temperature on extension of **1** templated by **2** in the presence of **3**[a]

Oligos	Monom.	c_{monom}/ mM	T/°C	k_1[b]/10^{-2} h^{-1}	$t_{1/2}$/h	Conversion[c] (%)
1/2/–	**5t**[d]	50	0	<0.1	>2300	17[e]
1/2/3	**5t**[d]	50	0	<0.1	>1300	26[e]
1/2/–	**4t**	20	4	5.2 ± 0.2	13.4	47
1/2/3	**4t**	20	4	9.2 ± 0.4	7.6	78
1/2/–	**4u**	5	4	5.7 ± 0.3	12.2	19
1/2/3	**4u**	5	4	5.8 ± 0.2	11.9	45
1/2/3	**4t**	1	−20	2.2 ± 0.1	31.9	27
1/2/3	**4t**	2	−20	2.5 ± 0.1	27.5	61
1/2/3	**4t**	5	−20	4.8 ± 0.1	14.5	92
1/2/3	**4u**	5	−20	8.5 ± 0.3	8.2	94

[a] Conditions: 268 μM oligonucleotides, 200 mM HEPBS, 400 mM NaCl, 80 mM $MgCl_2$, pH 8.9. [b] Divide by c_{monom} for second-order rate constants. Since hydrolysis is an unavoidable side reaction, the rate constants are not concentration independent and may only be compared quantitatively at the same monomer concentration and temperature. [c] Determined by fitting kinetic data, unless otherwise stated. [d] HEPES instead of HEPBS, pH 7.7. [e] After 300 h.

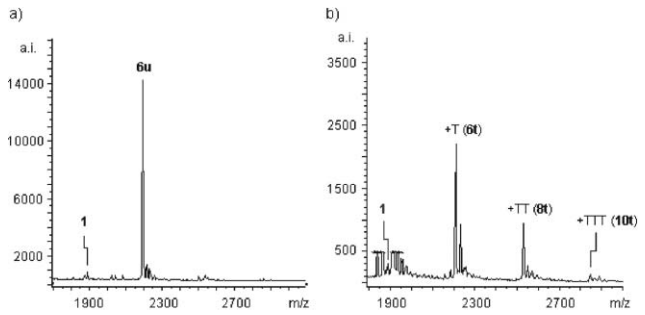
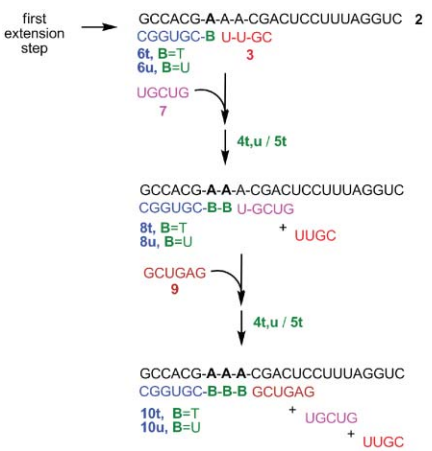


Fig. 2 Typical MALDI-TOF mass spectra showing a) extension of **1** with **4u** at −20 °C, templated by **2** and aided by **3** after 98 h; b) multiple extension of primer **1** with **4t** templated by **2** and aided by helpers **3**, **7** and **9**, total reaction time 580 h at −20 °C; 5 mM **4t** at t_0, additions of **4t** after 74 h and 146 h (11 mM each), 218 h (9 mM), and 341 h (51 mM). Addition of **7** after 74 h, and addition of **9** after 341 h.

by a strengthening of the template effect. Even for uridine **4u**, near-quantitative extension was measured (Table 1, Fig. 2a).

We then performed an exploratory study on multiple extensions. For this, fresh aliquots of activated monomer were added whenever the rate slowed down significantly. Helper displacement was induced by adding a second, longer helper (oligonucleotide **7**, Scheme 3) and a fresh aliquot of **4t** to the solution after the first extension step had progressed to at least 50%. Pentamer **7** has a higher affinity for template **2** than tetramer **3** and can be expected to displace 96% of the latter at 0 °C, according to the predicted[30] hybridization equilibrium (see ESI†). Different monomer concentrations were tested, both during the initial phase and later phases of the reaction, together with different temperatures and time points of helper addition.

The highest yielding conversion of **1** to **10t** was again observed at −20 °C. Fig. 2b shows a mass spectrum of a sample drawn after 24 d at −20 °C. Over time, four aliquots of a stock solution of **4t** had been added. Further, helpers **7** and **9** had been injected after 3 d and 14 d, respectively. The third helper (RNA strand **9**) can be expected to displace 98% of **3** and **7** from template **2** at 0 °C (Fig. S2, ESI†). Significant extension of **1** with up to three ribothymidylic acid residues was observed, demonstrating that even the sequence AAA is not a total block to spontaneous replication. The product distribution after 24 d was 6% of remaining **1**, 64% singly extended **6t**, 26% doubly extended **8t**, and



Scheme 3 Multiple primer extension with helper displacement.

Table 2 Kinetics for the second primer extension step (**6u** to **8u/t**) with **4t**,[a] with a second addition of **4t** (80 mM) after 7 d[b]

	$c(\mathbf{4t})_{initial}$/mM	k_1[c]/10^{-2} h^{-1}	$Conversion_1$[d] (%)	$c(\mathbf{4t})_{final}$/mM	k_2[e]/10^{-2} h^{-1}	$Conversion_2$[f] (%)
Oligos	first reaction phase (before addition of **4t**)			second reaction phase (after addition of **4t**)		
6u/2/–	5	1.7 ± 0.1	22	80	1.4 ± 0.1	36
6u/2/7	5	2.5 ± 0.1	49	80	1.3 ± 0.2	72

[a] Conditions: 268 μM oligonucleotides, 200 mM HEPBS, 400 mM NaCl, 80 mM $MgCl_2$, pH 8.9, −20 °C. [b] See ESI for further details. [c] Pseudo first order rate constant for the first phase (<7 d) determined graphically as initial rate. [d] Conversion of **6u** after 7 d (first reaction phase) with **4t** (5 mM). [e] Pseudo first order rate constant for the initial rate of the second phase of the reaction (>7 d). [f] Total conversion of **6u** (t >400 h) after single addition of **4t** (200 nmol).

4% triply extended **10t**. With uridine **4u** as monomer, 3% of **1** remained after 10 d, with 84% singly, 12% doubly and 1% triply extended primer.

We then asked why the second primer extension step is less efficient than the first. For this, we performed assays with synthetic primer **6u**. The reactions were run in the presence and in the absence of helper **7** at −20 °C with a 5 mM starting concentration of **4t** (Table 2). Up to 49% extension of **6u** to **8u/t** was observed in the presence of **7** without additional **4t**, and up to 72% conversion after adding one more aliquot of **4t** after 7 d. The yield-enhancing effect of **7** was similar to that seen for **3** with **1/2**. Further, when the kinetics were monitored, the rate of the initial extension reaction involving **6u/2/7** was found to be similar to that for **1/2/3**, suggesting that there is no intrinsic problem with AA templates. The second phase of the reaction, induced by adding fresh **4t**, was a little slower than the first, though, suggesting that the ribothymidine monophosphate formed through hydrolysis affects the reaction.

When the mixture of **4t**, **2** and **7** was allowed to sit for 3 d at −20 °C (5 mM **4t**) or 7 d at +4 °C (20 mM **4t**) in the absence of primer, so that the hydrolysis of **4t** could progress, and primer **6u** was added later, the extension reaction was as slow and as incomplete as initially observed for the assay shown in Scheme 3 (see ESI†). This confirms that it is not the sequence motif in itself that blocks the second extension step, but that accumulation of hydrolyzed monomers interferes with extension. It is reasonable to assume that spent and active monomer bind to the reaction site with similar affinities. This scenario leads us to conclude that replication of sequences containing stretches of weakly base-pairing nucleotides may require re-activation of the monomer during the reaction time and/or removal of spent monomer to achieve high yields.

No multiple extension of **1** beyond **6t** was observed with methylimidazolide **5t** as monomer under any conditions tested, including those suggested by the literature.[26] This confirms that the azabenzotriazolide of the nucleotide affects the ratio between the rates for primer extension *vs*. hydrolysis favorably. So, given a proper combination of monomer activation, temperature and helpers, adenosine residues promote replication steps quite effectively. Helper strands of the length employed are readily formed on minerals.[7] Since the principles established here may also apply to replication steps involving other weakly templating sequences, it is not unreasonable to speculate about the likelihood of potentially prebiotic replication in a permanently frozen, perhaps even Arctic environment. Higher yields for multiple extensions may have resulted from more ideal activation chemistry and conditions favoring removal of "spent" monomers and/or utilized "micro helpers". Experiments to identify such conditions are currently under way in our laboratories.

This work was supported by DFG (grant No. RI 1063/1-4). The authors thank C. Deck for helpful comments.

Notes and references

‡ These studies employed ribo-terminated DNA templates, explaining why thymine as nucleobase was involved, and not uracil.

1 (*a*) R. F. Gesteland, T. R. Cech and J. F. Atkins, *The RNA World*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 3rd edn, 2006; (*b*) W. Gilbert, *Nature*, 1986, **319**, 618.
2 (*a*) J. Sulston, R. Lohrmann, L. E. Orgel and H. T. Miles, *Proc. Natl. Acad. Sci. U. S. A.*, 1968, **59**, 726.
3 T. Inoue and L. E. Orgel, *Science*, 1983, **219**, 859.
4 M. Kurz, K. Göbel, C. Hartel and M. W. Göbel, *Helv. Chim. Acta*, 1998, **81**, 1156.
5 A. Kanavarioti, C. F. Bernasconi, J. A. Diann and E. E. Baird, *J. Am. Chem. Soc.*, 1993, **115**, 8537.
6 J. C. Chaput, S. Sinha and C. Switzer, *Chem. Commun.*, 2002, 1568.
7 (*a*) J. P. Ferris and G. Ertem, *Science*, 1992, **257**, 1387; (*b*) W. Huang and J. P. Ferris, *J. Am. Chem. Soc.*, 2006, **128**, 8914.
8 M. Zielinski, I. A. Kozlov and L. E. Orgel, *Helv. Chim. Acta*, 2000, **83**, 1678.
9 G. F. Joyce, *Cold Spring Harbor Symp. Quant. Biol.*, 1987, **52**, 41.
10 R. Lohrmann and L. E. Orgel, *J. Mol. Biol.*, 1980, **142**, 555.
11 A. R. Hill, L. E. Orgel and T. Wu, *Origins Life Evol. Biosphere*, 1993, **23**, 285.
12 L. E. Orgel, *Acc. Chem. Res.*, 1995, **28**, 109.
13 W. L. Jorgensen and J. Pranata, *J. Am. Chem. Soc.*, 1990, **112**, 2008.
14 I. A. Kozlov and L. E. Orgel, *Helv. Chim. Acta*, 1999, **82**, 1799.
15 C. Hartel and M. W. Göbel, *Helv. Chim. Acta*, 2000, **83**, 2541.
16 M. Hey, C. Hartel and M. W. Göbel, *Helv. Chim. Acta*, 2003, **86**, 844.
17 M. Koppitz, P. E. Nielsen and L. E. Orgel, *J. Am. Chem. Soc.*, 1998, **120**, 4563.
18 I. A. Kozlov, P. K. Politis, A. van Aerschot, R. Busson, P. Herdewijn and L. E. Orgel, *J. Am. Chem. Soc.*, 1999, **121**, 2653.
19 I. A. Kozlov, B. Bouvere, A. van Aerschot, P. Herdewijn and L. E. Orgel, *J. Am. Chem. Soc.*, 1999, **121**, 5856.
20 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.
21 L. Zhang, A. Peritz and E. Meggers, *J. Am. Chem. Soc.*, 2005, **127**, 4174.
22 X. Wu, G. Delgado, R. Krishnamurthy and A. Eschenmoser, *Org. Lett.*, 2002, **4**, 1283.
23 P. Hagenbuch, E. Kervio, A. Hochgesand, U. Plutowski and C. Richert, *Angew. Chem., Int. Ed.*, 2005, **44**, 6588.
24 S. R. Vogel, C. Deck and C. Richert, *Chem. Commun.*, 2005, 4922.
25 T. Wu and L. E. Orgel, *J. Am. Chem. Soc.*, 1992, **114**, 7963.
26 (*a*) T. Wu and L. E. Orgel, *J. Am. Chem. Soc.*, 1992, **114**, 317; (*b*) T. Wu and L. E. Orgel, *J. Am. Chem. Soc.*, 1992, **114**, 5496; (*c*) T. Wu and L. E. Orgel, *J. Am. Chem. Soc.*, 1992, **114**, 7963.
27 A. V. Vlassov, B. H. Johnston, L. F. Landweber and S. A. Kazakov, *Nucleic Acids Res.*, 2004, **32**, 2966.
28 P.-A. Monnard, A. Kanavarioti and D. W. Deamer, *J. Am. Chem. Soc.*, 2003, **125**, 13734.
29 M. Renz, R. Lohrmann and L. E. Orgel, *Biochim. Biophys. Acta*, 1971, **240**, 463.
30 K. Siegmund, U. E. Steiner and C. Richert, *J. Chem. Inf. Comput. Sci.*, 2003, **43**, 2153. See also: http://chipcheck.chemie.uni-karlsruhe.de/chipcheck/.

