

Metal-ion catalyzed polymerization in the eutectic phase in water–ice: A possible approach to template-directed RNA polymerization

Pierre-Alain Monnard^{a,*}, Jack W. Szostak^b

^a *Los Alamos National Laboratory, EES-6, P.O. Box 1663, MS-D462, Los Alamos, NM 87545, USA*

^b *Howard Hughes Medical Institute, Center for Computational and Integrative Biology, and Department of Molecular Biology, Massachusetts General Hospital, CCIB 7215, Simches Research Center, 185 Cambridge Street, Boston, MA, USA*

Received 2 October 2007; received in revised form 17 January 2008; accepted 18 January 2008

Available online 1 February 2008

Abstract

The emergence of an RNA world requires among other processes the non-enzymatic, template-directed replication of genetic polymers such as RNA or related nucleic acids, possibly catalyzed by metal-ions. The absence of uridilate derivative polymerization on adenine containing templates has been the main issue preventing an efficient template-directed RNA polymerization. We report here the investigation of template-directed RNA polymerization in the eutectic phase in water–ice. In particular, it was found that activated uridilate monomers in the presence of metal-ion catalysts could efficiently elongate RNA hairpins whose 5′-overhangs served as the templating sequence. The same applies for every other pyrimidine and purine nucleobase. Moreover, the initial elongation rates were always higher in the presence of a template complementary to the nucleotide than in systems without proper base-pairing opportunities. These results suggest that a template-directed RNA polymerization catalyzed by metal-ions could be carried out under eutectic phase in water–ice conditions.

© 2008 Elsevier Inc. All rights reserved.

Keywords: Template-directed RNA polymerization; Metal-ion catalyzed polymerization; Eutectic phase in water–ice; Micro-environment

1. Introduction

The RNA world hypothesis proposes that RNA likely preceded and promoted the emergence of a DNA–RNA–protein system that defines contemporary organisms. Indeed, RNA alone can function as a genetic information repository, and can also perform the catalytic functions of several catalytic proteins [1]. The emergence of the RNA world must have initially required four non-enzymatic processes likely catalyzed by dissolved metal-ions: (1) monomer synthesis and activation followed by (2) monomer polymerization (also called self-condensation), (3) elongation of existing polymers, and (4) template-directed

polymer replication. The replication process, essential for the amplification of functional RNA sequences, requires the hybridization or base-pairing of complementary ribonucleotide monomers with an existing RNA strand called the template and the subsequent polymerization of those monomers into a continuous fully complementary strand.

In previous experiments that were carried out with metal-ion catalysts and activated monomers in liquid aqueous solutions (above 0 °C), the major obstacle to an efficient template-directed reaction was the poor Uridine-derivative reactivity [2–6]. In fact, an efficient template-directed polymerization from activated monomers was only observed with activated G nucleotides on templates containing a large fraction of C residues (over 60%) [7,8]. In addition, if the template sequence contained more than one other residue (A, U) between stretches of C residues,

* Corresponding author. Tel.: +1 505 664 0718; fax +1 505 667 3285.

E-mail addresses: pmonnard@lanl.gov (P.-A. Monnard), szostak@molbio.mgh.harvard.edu (J.W. Szostak).

replication became extremely slow or non-existent. DNA template sequences (used in most studies of template-directed RNA polymerization) like GT or TG significantly inhibit replication and sequences like AT, TA, AA, GA and AG completely block the formation of complementary strands [3,4,8–10]. The use of RNA templates did not improve the reaction [11]. Furthermore, the template-directed replication of templates rich in G-residues would only occur if the formation of structures in the template was suppressed [12]. Together these restrictions essentially prevent true amplification of RNA sequences in homogeneous aqueous solutions.

It was previously shown that the eutectic phase in water–ice at $-18.4\text{ }^{\circ}\text{C}$ yields a micro-structured environment that supports the polymerization of RNA monomers into chains at least as long as 22-mers by condensation of imidazole-activated monomers (ImpN or nucleoside 5'-monophosphoimidazoles) using dissolved Mg^{2+} and Pb^{2+} -ions as catalysts [13]. However, whereas the self-condensation of activated Uridine monomers almost quantitatively yielded polymers with a length exceeding that of a 20 mer, previous attempts at polyadenylic acid (a homopolymeric template) directed polymerization of polyuridylic acid, the homopolymeric product, in the same micro-structured environment, but using only Mg^{2+} -ion catalysts in a higher starting ionic strength medium, were unsuccessful [14]. In the light of the self-condensation of uridine results, the re-examination of the template-directed polymerization in the eutectic phase seemed essential.

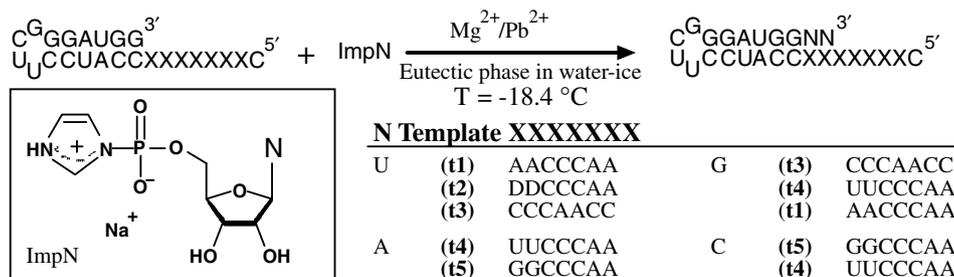
This micro-structured environment was obtained by cooling a solution containing activated monomers and dissolved metal-ion catalysts below its depressed freezing point, but above its eutectic point (the temperature at which the whole solution freezes solid). In this system, two water phases coexist: a solid one (the ice crystals) and a liquid one.

There are several important aspects to note about the water/ice eutectic phase system. First it is possible to create a super-cooled situation in samples with a small volume where ice should exist but does not. For this reason we flash-froze our solutions in order to ensure ice formation, prior to “warming” the system back to the operating temperature where the liquid solution and the solid ice phase

coexist. Secondly, the solid part of the two-phase system contains at most a small fraction of the solutes. As a result, the solutes become highly concentrated in the residual liquid phase that exists in the space between the solid ice regions and possibly within micro-structures. Based on crude theoretical calculations, rough estimates indicate concentration enhancement by some two orders of magnitude for the solute concentrations used in this study. Furthermore, since the monomer concentration is only a relatively small fraction of the overall starting solute concentration, in spite of the large overall concentration enhancement found in the residual brine, the relative fractional changes in the monomer concentration in the starting solution are still expected to be approximately maintained in the residual brine. Finally, the overall monomer concentrations that can be reached also depends on the starting concentration of the dominant solute, as well as the ratio of the starting monomer concentration to that of the dominant solute.

Since the length of the RNA products obtained in the eutectic phase by condensation of monomers should allow the formation of secondary structural motifs, such as hairpins, that could support a template-directed elongation, RNA hairpins were used in this study. These single-molecule primer-template systems (see Scheme 1) allowed the investigation of the effects of template sequences on elongation rates of each nucleobase by itself (non-competitive conditions). The major technical differences compared to previous attempts [14] were (1) the use of a single-molecule primer-template system, the hairpin, versus the use of homopolymeric template, (2) the use of a lower starting ionic strength, and (3) the composition of metal-ion catalyst system. Furthermore, the initial elongation rates rather than the elongation yields were monitored.

The success of template-directed RNA polymerization depends on the actual polymerization of every nucleobase on its cognate residue. The insertion of even one wrong monomer could clearly impact information transfer, thereby possibly compromising the capacity for RNA activity. Thus, the investigation of the elongation rates of each monomer on non-cognate templates is also important as lower rates would mean that the overall error rates in template-directed polymerization would be reduced espe-



Scheme 1. Template-directed polymerization using RNA hairpins. The 5'-overhang sequence of the hairpin serves as the template for the elongation of its 3'-end which functions as the primer. The hairpin motif with a duplex stretch of 6 bp ensures the relative stability of the construct at temperatures as high as $4\text{ }^{\circ}\text{C}$ according to calculations made using the Vienna package (<http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi>).

cially when all the monomers are present simultaneously. That is, a correct copying of the RNA template in the eutectic phase is more likely to be achieved.

In the studies reported here, initial elongation rates in several systems were compared using imidazole-activated monomers (see Scheme 1), Mg^{2+} and Pb^{2+} -ions as catalysts, and a straight RNA primer or an RNA hairpin, the latter acting as a combined primer and template. The systems studied were: (i) hairpin systems where the single activated monomer species used was cognate with the first two overhanging residues of the template (base-pairing can occur); (ii) hairpin systems where this was not the case; and (iii) linear primer systems to study simple primer elongation.

These studies revealed that the elongation of self-templating RNA hairpins in the eutectic phase in water–ice by an activated Uridine derivative (ImpU, Imidazole-activated Uridilate) is possible opposite A- and D-residues in the template where D is a 2,6-diaminopurine nucleobase that forms three hydrogen bonds with uridine. Furthermore, the initial elongation rates for all four nucleobases opposite to their cognate Watson–Crick base-pair residues in the first two positions of the templating hairpin overhang were higher than in systems containing two non-cognate residues. Thus, template-directed elongation of RNA fragments catalyzed by metal-ions seems possible in the eutectic phase in water–ice.

2. Materials and method

Reagent grade materials were used throughout, and solvents were HPLC grade. The reagents were purchased from the following sources: 2-(N-morpholino)ethane sulfonic acid (MES) from Sigma; magnesium nitrate hexahydrate >99.5%, and lead(II) nitrate 99% from Fluka; $[\gamma\text{-}^{32}\text{P}]$ ATP from Amersham Pharmacia; T4 polynucleotide kinase (T4 PNK) with a specific activity of 10 units/ μl from NEN, and RNAsin from Promega.

RNAs were obtained from the Keck Facility (Yale University, CT, USA) and from Dharmacon RNA technologies (Lafayette, CO, USA).

The samples were directly incubated in a Neslab RTE10 (10 l) refrigerated bath set at $-18.4\text{ }^\circ\text{C}$ with fluctuation limits of $\pm 0.3\text{ }^\circ\text{C}$. This apparatus possesses a thermocouple cooling system that ensures a temperature precision of $\pm 0.1\text{ }^\circ\text{C}$ and its actual temperature was regularly checked for accuracy with an external calibrated thermometer.

2.1. Sample preparation

2.1.1. Preparation of the monomers

The sodium salts of nucleoside 5'-phosphoimidazolides, ImpN (N = Adenosine, Cytidine, Guanosine, Uridine; see structure in Scheme 1), were prepared as previously described [16] with a $98 \pm 1\%$ purity. Samples were always prepared from freshly made stock solutions of the substrates to preserve limit the extent of possible spontaneous

deactivation prior to use. The concentrations of the stock solutions were determined by UV spectroscopy.

2.1.2. Preparation of the RNAs

First the RNAs were deprotected according to the manufacturer's recommended protocols. RNAs were then purified by HPLC chromatography on an anion-exchange column, DNAPac PA100, from Dionex (Sunnyvale, CA, USA) (sodium perchlorate gradient: 0–250 mM in 40 min; column at $55\text{ }^\circ\text{C}$). The collected RNA solutions containing the longer RNAs (i.e., RNA hairpins) were then run through NAP-25™ columns (GE Healthcare Life Science, PA, USA). The collected solutions containing the shorter RNA (LR8, 5'-GCUUCGGA-3') were desalted on a reverse-phase column (C-18) using a 50 mM TEAA solution with 25% v/v acetonitrile. In both cases the eluates were finally lyophilized overnight.

RNAs were labeled at their 5' end using T4 polynucleotide kinase and $[\gamma\text{-}^{32}\text{P}]$ ATP. The usual phosphorylation solution contained 3 nmol RNA, 5.0 μl $[\gamma\text{-}^{32}\text{P}]$ ATP ($\approx 50\text{ }\mu\text{Ci}$), 120 units T4 PNK, and 480 units of RNAsin (final volume = 300 μl). After a 1–1.5 h incubation period, the enzyme was removed by phenol/chloroform extraction and the RNAs were precipitated using NaOAc/EtOH at $-20\text{ }^\circ\text{C}$. The RNAs were then suspended in RNase-free water. Both the specific activity and the RNA concentration were measured by scintillation counting and UV spectroscopy, respectively.

2.1.3. Composition of the samples

The solution used in a run was prepared from stock solutions of 10–50 mM ImpN except for the ImpG (1–12.5 mM), 30 mM $Pb(NO_3)_2$, 60 mM $Mg(NO_3)_2$, and 50 mM MES, pH 6.5. RNAs and monomers were first mixed together, then heated at $50\text{ }^\circ\text{C}$ for 2 min, cooled on ice, and finally added to the buffer which already contained the metal-ions. Millipore water was used to achieve the final desired concentrations. A typical solution contained a monomer at the specified concentration, 1 μM RNA to be elongated (the primer or the hairpin), 5.2 mM $Mg(NO_3)_2$, 1.2 mM $Pb(NO_3)_2$, and 5 mM MES buffer, pH 6.5. The pH of all reaction mixtures was in the range of 6.6–6.9 at room temperature and remained stable during incubation.

2.1.4. Elongation rate determination

Analysis of the samples was done by gel electrophoresis (20% polyacrylamide gel – length 40 cm, 40–56 W, 2–3 h), exposed on image plates (molecular dynamics) and analyzed using ImageQuant Software (molecular dynamics).

The experimental results are classified in terms of runs, where a single run refers to the preparation and use of a common solution consisting of a specific RNA and a specific monomer at a specific concentration. After preparation, the solution was subdivided into separate containers, each containing 20–30 μl of solution that yielded at least 40,000 counts per minute (cpm) from the

radioactively labeled RNA hairpin or linear oligomer. Each container represents a single sample. A single run consists of multiple samples. All the samples representing a single run were prepared in a single batch and then simultaneously flash-frozen in acetone/dry ice and then directly incubated at $-18.4\text{ }^{\circ}\text{C}$ in the refrigerated bath for various periods of time. The flash-freezing process ensured the ice nucleation (less than 10 s) in all samples, while the final temperature of $-18.4\text{ }^{\circ}\text{C}$ ensured partial “re-melting” of the samples resulting in a two-phase system. Furthermore, when samples slowly cooled from $0\text{ }^{\circ}\text{C}$ (within 5 min) were compared to the ones prepared using the flash-frozen approach, the elongation rates were within the standard variation.

Samples were then withdrawn at specific incubation times, thawed, then immediately quenched with EDTA (final concentration = 9.6 mM) and kept at $-30\text{ }^{\circ}\text{C}$ until the last samples were collected. All the samples from a given run were subsequently analyzed in parallel. Each sample from a run thus represented the reaction extent at the specific quenching time. For some runs only a single sample was withdrawn for each specified time, while for other runs several (typically 3) samples were withdrawn simultaneously to help determine the repeatability and error associated with the overall process.

The runs were conducted in series, with a given series referring to an entire set of runs that were conducted in parallel for a specific monomer and specific template, but variable starting monomer concentrations. At least three series of runs were performed for each combination of monomer and template type. Each series was analyzed independently to yield a rate. The 3 or more individual elongation rates for a specific template–monomer combination were then averaged to yield the values reported here. As some of run series had multiple samples for each time point, the rate value from such a series was given a higher weight in determining the average rate.

2.1.5. Data analysis

For each time value of a run, the elongation of the RNA hairpin template was monitored by quantifying the fraction of extended versus non-extended templates. Initial elongation rates were determined by fitting a plot of the fraction of elongated templates during the first 10–20 h of a particular run series using the linear regression fitting function with the force through zero option of DeltaGraph (SPSS Inc. and RedRock Software, USA). The force through zero option was used as the analysis of samples immediately thawed and quenched after formation of the eutectic phase showed little or no detectable elongation.

The initial elongation rates were calculated separately for each independent run series as mentioned above. Only regression results with a confidence (R^2) of more than 0.8 were considered for the calculation of the average rates reported below.

To rule out elongations resulting from the formation of 5'-pyrophosphate-linkages between monomers and the 5'-

phosphate of the RNAs, reactions were also conducted using non-labeled templates under the same conditions used for the labeled templates. The resulting RNA was then labeled and the initial elongated fraction was compared to that obtained at the same monomer concentration with pre-labeled templates. This post-incubation labeling ensured that no molecules with a 5'-pyrophosphate cap [5'-NppN(pN)] would be detected. In all cases, these rates were within the standard deviations reported in Table 1. This confirmed the already observed low pyrophosphate formation under eutectic conditions for monomer condensation [13].

2.2. Example of initial elongation rate determination

The radioactive material migrating above the band representing the non-elongated hairpin in lane 4 and 17 (see Fig. 1) was subtracted from the elongated material in the

Table 1
Initial elongation rates

ImpN	Template	[ImpN] (mM)	Initial elongation rate $\times 10^{-2}$ (h^{-1})	Ratio of rates ^a	To template	
U	t1	1.0	1.14 ± 0.15	1.11	t3	
		1.3	1.17 ± 0.21	N.D. ^b		
		1.5	1.69 ± 0.16	1.59		
		1.74	1.73 ± 0.15	N.D.		
		2.0	1.46 ± 0.14	1.30		
	t2	1.0	1.66 ± 0.2	1.61 & 1.46	t3 & t1	
		1.3	1.93 ± 0.1	N.D. & 1.64	t3 & t1	
		1.74	2.10 ± 0.1	N.D. & 1.21	t3 & t1	
		2.0	1.85 ± 0.02	1.65 & 1.27	t3 & t1	
	t3	1.0	1.03 ± 0.26			
		1.5	1.06 ± 0.17			
		2.0	1.12 ± 0.23			
	G	t3	0.1	1.02	1.82 & 1.10	t4 & t1
			0.25	1.13 ± 0.1	N.D. & 1.27	t4 & t1
			0.5	1.44 ± 0.02	2.36 & 2.67	t4 & t1
1.0			1.37 ± 0.15	2.28 & 2.11	t4 & t1	
t4		0.1	0.93			
		0.5	0.54			
t1		1.0	0.65			
		0.1	0.56			
		0.25	0.89 ± 0.06			
		0.5	0.61 ± 0.12			
A	t4	1.0	0.60 ± 0.20			
		0.5	1.31 ± 0.09	1.72	t5	
		0.75	1.84 ± 0.13	1.67	t5	
	t5	1.0	1.52 ± 0.14	1.38	t5	
		0.5	0.76 ± 0.15			
		0.75	1.10 ± 0.14			
C	t5	1.0	1.10 ± 0.15			
		1.0	1.87 ± 0.20	1.53	t4	
		1.25	1.96 ± 0.19	1.41	t4	
	t4	1.5	0.96 ± 0.16	0.86	t4	
		1.0	1.22 ± 0.10			
		1.25	1.39 ± 0.15			
		1.5	1.11 ± 0.12			

^a The ratio is calculated as follows: initial elongation rate of column 4 divided by that of the template(s) listed in last column.

^b No data available.

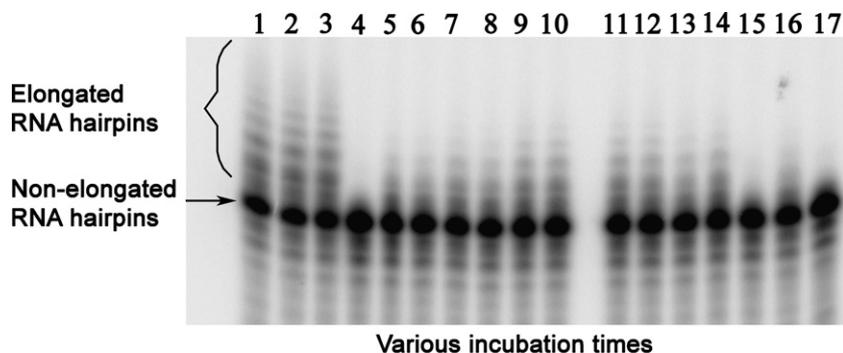


Fig. 1. Electrophoresis gel of 1.3 mM ImpU elongation on t2 (DD-template). This gel is part of run series with three samples per time point: Lanes 1–10 represent the first sample series, and lanes 11–17 are from the second sample series. Lanes 1–3 are three samples taken at 47-h of incubation, while lanes 4–10 are, respectively, for incubation times of 0, 3, 4, 6, 7, 9, and 11 h. Lanes 17–11 are the second run series time points for the same 0–11 h incubation points, but are in reverse order.

subsequent bands in the determination of the elongated hairpin fraction. The hairpins elongated beyond the second cognate residues were also included in the yield determination, as these products were first correctly elongated by two monomers.

The linear regression shown in Fig. 2 was calculated using seven average time points and forcing the regression line through zero. The initial elongation rate was $1.79 \times 10^{-2} \text{ h}^{-1}$ with a confidence (R^2) of 0.921. This result was then combined with two other rates (representing 5–9 samples per incubation time) to obtain the average elongation rate of $1.93 \pm 0.1 \times 10^{-2} \text{ h}^{-1}$ reported later in this paper.

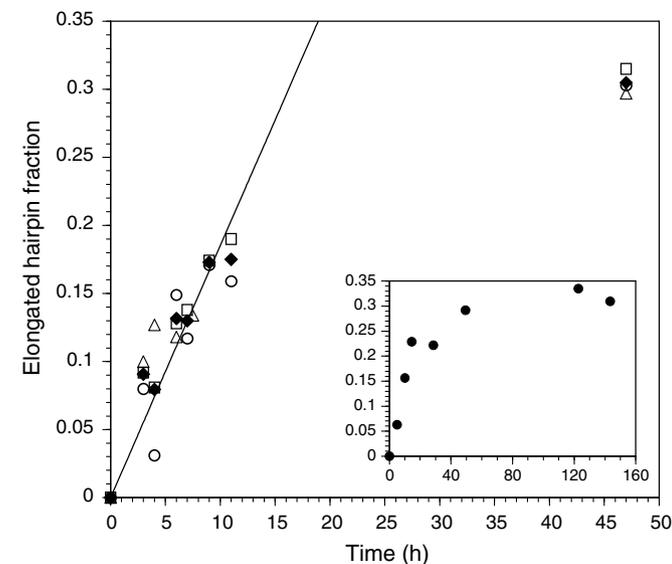


Fig. 2. Numerical analysis: fraction of elongated hairpin template and determination of the initial elongation rate for 1.3 mM ImpU elongation on t2 (DD-template). This is the same run series as that for the gel shown in Fig. 1. Open symbols: individual elongation yields; solid diamond: average yields for each time point. (Inset) Representative kinetic of the elongation of 1.0 mM ImpU on t2 (solid circles). This example shows the general elongation pattern over longer periods of time. In particular, it highlights the contrast between the early linear elongation phase used to determine the initial rates and the much slower later rates, as well as intermediate rates in the transition region.

3. Results

In this report, the possibility of a metal-ion catalyzed, template-directed RNA polymerization in the eutectic phase in water–ice at $-18.4 \text{ }^\circ\text{C}$ was investigated. This temperature is slightly below the freezing point of samples ($-17.8 \pm 0.3 \text{ }^\circ\text{C}$) and permits us to maintain the eutectic phase in water–ice samples for long time periods as established previously [13,15]. This temperature is also above the eutectic point as polymerization was observed at temperatures as low as $-20 \text{ }^\circ\text{C}$ but not at $-21 \text{ }^\circ\text{C}$. In particular, the elongation rates of various RNA hairpins with a 5'-overhang serving as the template were studied using imidazole-activated ribonucleoside monophosphates (ImpNs) as activated monomers. This was done for systems where the two first residues on the template are either the cognate or a non-cognate nucleobase for the activated monomers. The elongation rates for ImpG and ImpU on a linear RNA oligomer (non-template-directed elongation) were also determined to establish how the absence of a template affects elongation.

The initial elongation rates in all systems were determined using samples that were incubated for up to two days at $-18.4 \text{ }^\circ\text{C}$. This time limit was imposed to prevent possible interference resulting from the hydrolysis of either activated monomers or RNAs that can occur over longer periods of incubation. Indeed, hydrolysis of activated monomers remained low during first 48 h, representing less than 8.5% and 5% of the initial monomer concentration (data not shown) in the presence of and the absence of 5.2 mM magnesium nitrate. (Magnesium is known to promote the hydrolysis, but not the polymerization [15]). The decomposition of the hairpins and primers was also negligible during such a short incubation whereas it became apparent, but nonetheless remained small (less than 5% of the total RNA were affected) on the gels after 72–120 h incubation. In addition, as seen in Fig. 2 for the template-directed hairpin case, the elongation results showed that these reactions proceeded rapidly during the initial 10–20 h of incubation, but then tended to slow down at later times.

3.1. Non-template-directed primer elongation

A linear primer (LR8) 5'-GCUUCGGA-3' was used to establish the elongation rates in the absence of a template. Because U and G nucleotides have the most distinct behavior with respect to polymerization, we limited our elongation studies to imidazole-activated U and G monomers in the non-templated case.

The initial elongation rates for the reactions without templates increased with increasing monomer concentration until a monomer concentration value of 1.5–2.0 mM was reached and thereafter declined (see Fig. 3).

For both monomers, it was determined that the maximum initial elongation rates were obtained in the same monomer concentration range. The elongation yields after 24 h were as high as 25% and 12% of initial primer concen-

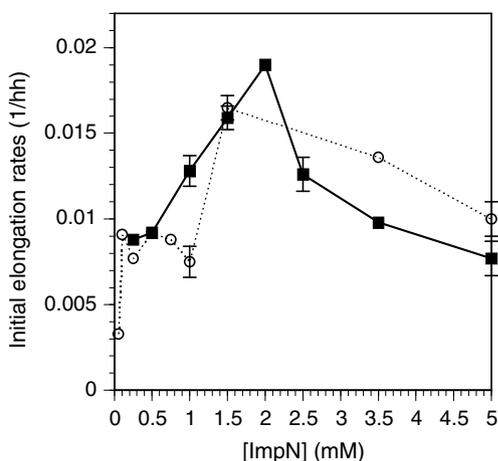


Fig. 3. Dependence of the elongation rates of a linear primer LR8. Elongation with ImpU (filled squares and the uninterrupted line) and ImpG (open circles and the dotted line). Some data points are an average of two independent single sample runs and are an exception to multiple repetitions (5–9 samples) typically used to determine the elongation rates.

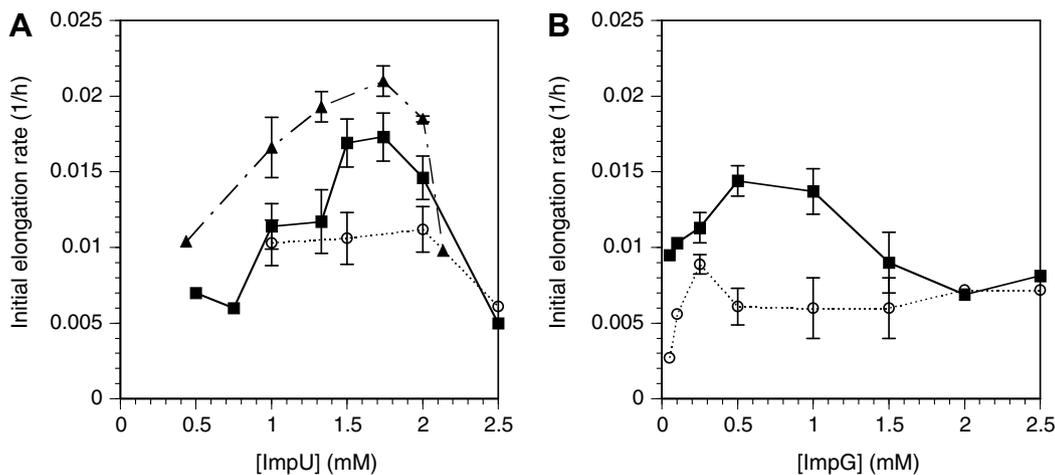


Fig. 4. Dependence of the initial elongation rates of an RNA Hairpin on the starting concentration of activated monomer. (A) ImpU: the filled squares and the solid line represent the rates obtained with t1 (AA template); the filled triangles and the (---) line those on t2 (DD-template) and the open circles and the dotted line those with t3 (CC template). (B) ImpG: The filled squares and the solid lines represent the rates obtained with t3 (CC template) and the open circles and the dotted line those with t1 (AA template).

tration (1 μ M) respectively for ImpU and ImpG. Up to six U and four G nucleotides were added. Furthermore, in the absence of metal-ions, little or no elongation was detected. The different elongation behavior for the two monomers corresponds well to the observations made during the earlier monomer self-condensation experiments where ImpU always tended to produce longer products with higher yields than ImpG [15].

3.2. Template-directed elongation

Next, the template-directed elongation rates were determined using RNA hairpins with a 5'-overhang templating sequence (see Scheme 1). The same general elongation rate dependence on the monomer concentration was established (see Fig. 4A and B) as seen in the non-template elongation study. In the system ImpU-t1 (containing the cognate nucleobase, see Fig. 4A), the highest measured rates were observed in the same range of monomer concentration (1.5–2.0 mM) as in the non-template-directed elongation experiments. Comparable elongation rate values were also found. In contrast, primer elongation on template t3 (templating sequence containing CC residues) was slower, showing that the elongation was inhibited by the presence of non-cognate residues at the beginning of the templating sequence. The non-cognate template residues may prevent the monomer from assuming the correct positioning at the 3'-end of the hairpin required for the elongation to occur.

As expected, the additional H-bond from the D-residue in the ImpU-t2 system significantly increased the initial elongation rate at low monomer concentrations relative to that for the ImpU-t1 system. The fractional enhancement however moderated as the optimum concentration ranges for t1 and t2 were reached. This suggests that, at first, the additional H-bond improved the formation of the ImpU-template complex and thus the polymerization. Then, when the optimum monomer concentration range

was reached, the polymerization rates depended mostly on the reactivity of the imidazole activating group.

In the case of ImpG on t3, the hairpin containing CC in the templating region (see Fig. 4B), the maximum elongation rate was reached at a monomer concentration approximately three times lower than observed for the maximum elongation rate of the primer LR8 (without template). Nonetheless, the measured maximum elongation rate was comparable to that of LR8 by ImpG. Again, when elongation of a hairpin containing a non-cognate residue (t1) with ImpG was attempted, lower elongation rates were recorded.

Elongation rates were also determined for imidazole-activated A and C monomers using a base-pairing and a non base-pairing template sequence (see Table 1). To test the effect of a G–U wobble base-pair, a measurement of t4 with ImpG was also carried out.

In all the cases studied, the initial elongation rates peaked at some intermediate concentration values before declining at higher values. This is in contrast to kinetics studies in pure liquid phase aqueous media where the rates simply plateau at the higher concentration values [17].

The elongation rates in Table 1 were in general lower than those observed in reactions conducted in aqueous solutions at 25 °C ($\Delta T \approx 43$ °C), except for the elongation by ImpU on both templates t1 and t2 [3,4]. This is an important result because the uridilate monomer template-directed polymerization was never efficient in previous studies at 25 °C, while here up to 25% elongation of (t1) by at least one nucleotide and up to 13.0% by at least two nucleotides was observed after 24 h in the eutectic phase. For ImpG on (t3), yields of up to 30% were recorded for an elongation by at least one and up to 8.5% by at least two nucleotides. Interestingly, the system ImpG:t4 (UU) in which wobble base-pairs can be formed did not yield increased rates compared to the non base-pairing system (t1). The likely cause of this was that wobble base pair did not permit a correct positioning of the monomer along the 3'-end of the hairpin either. Note also that the difference in the elongation yields between ImpG and ImpU on the LR8 RNA completely disappeared in base-pairing systems.

All four activated nucleotides in base-pairing systems elongated the hairpin with a templating 5'-overhang containing their cognate residue. In these eutectic phase experiments, the similar rates we observed at each specific optimum monomer concentration range indicated no preferential elongation with purine monomers relative to pyrimidine monomers. This is in contrast to previous studies for homogeneous aqueous systems [3–5]. The sequence of optimum starting monomer concentration ranges $G \leq A < C < U$ observed for the elongation on a hairpin template with the cognate residues was consistent with the known nucleotide association tendencies in aqueous solutions [18,19], but these concentration ranges remained within a 1 mM interval.

When the elongation rates in base-pairing systems are compared to those of non base-pairing or wobble base-

pairing systems, the ratios indicated an improved elongation in the former systems (see Table 1). Under these non-competitive conditions (only one nucleobase species present at a time), the rates reported for the non base-pairing systems might be slightly higher than under competitive conditions (monomer mixtures). These conditions also explain why some elongation beyond the cognate residues was observed to have occurred.

4. Discussion

The studies presented here determined the elongation rates achieved by activated ribonucleotide monophosphates (ImpNs) when using an RNA hairpin with a 5'-overhang serving as a template in the eutectic phase in water–ice at -18.4 °C. The hope was that elongation would not only be achieved, but would also depend on the base pairing between the monomers and the templating sequence within the system. This process, called template-directed polymerization, would have been essential in the amplification/replication and evolution of RNAs with catalytic or informational functions ultimately required for the emergence of an RNA world.

The studies reported here established that a template-directed elongation of an RNA primer can occur both for pyrimidines and purines in clear contrast to the results obtained in homogeneous aqueous solutions or in previous attempts in the eutectic phase in water–ice [14]. Under the chosen non-competitive conditions used here, comparison of the initial elongation rates for the various RNA oligomers established that each of the four monomers elongated an RNA hairpin with a 5'-overhang template sequence containing its cognate nucleobase as the first two residues more efficiently than an RNA hairpin containing non-cognate residues. The initial elongation rates were also comparable for the four base-pairing systems. In addition, the elongation rates of the RNA hairpins containing non-cognate residues with ImpG and ImpU were lower than those of a linear primer LR8 (a system devoid of templating sequence). That is, the wrong residues in sequences of the template partially inhibited the elongation. Furthermore, the starting ImpG concentration for an optimal hairpin elongation (t3) was lower by a factor of three than that observed for the extension of a linear primer (LR8). For ImpU, the template effect on the optimal starting concentration for elongation was less pronounced. Indeed, the experiments conducted with the DD-template (t2) that can form an additional hydrogen bond with uridine resulted in an enhancement of the elongation rates at slightly lower starting concentrations (1.0 and 1.5 mM ImpU) compared to those needed for the LR8 system. These results indicate that base-pairing between monomers and the hairpin templates increased the probability of bond formation, thereby increasing the elongation yields. However, the elongation rates in RNA hairpin system containing non-cognate residues remain important and could lead

to high error rates if observed in elongation reactions in the presence of four activated monomers.

The difference in polymerization rates reported for ImpU here and previously [14] might be explained by the distinct reaction compositions and conditions. The previous attempt at the template-directed polymerization of activated U derivatives were carried out on poly(A) homopolymeric templates without primers for up to 12-week incubation and at high ionic strength. Recent reports for the polymerization of 2-MeImpA on poly(U) templates under similar conditions [20] have shown that such a reaction only proceeded after a lag phase of three months. The authors therein postulated that the formation of short poly(A) fragments which could act as primers was the rate limiting step in their polymerization, a observation that might apply to the Stribling and Miller experiments as well.

The ability to still achieve elongation in spite of the very low starting monomer concentrations (compared to those in studies at 25 °C) can be directly linked to the established concentrating power of the eutectic phase [13,15]. The loss of liquid water to the ice phase likely increases the effective concentrations by several (most likely by two as theoretical calculations suggest) orders of magnitude as highlighted in epifluorescence micrographs of eutectic phase in water–ice [15].

The apparent inhibition of the elongation at higher starting monomer concentrations in the eutectic phase is intriguing and reminiscent of substrate inhibition observed in enzymatic catalysis. This is also likely due to the intrinsic properties of eutectic systems and was observed for all monomers on every RNA template. The understanding of the eutectic phase supported RNA polymerization is at this time rather limited [13,15,20–22]. The freezing of the reaction mixtures clearly allows the formation of monomer-hairpin-metal catalyst complexes conducive to the template elongation that either do not exist or are much less stable at 25 °C, as is highlighted by the low temperature RNA elongation by ImpU, which is almost absent in room temperature studies even when high concentrations are used (0.1–1 M). This reaction promotion within the eutectic phase in water–ice occurs once the eutectic systems are formed and not during the freezing process as samples directly thawed and quenched after the formation of the eutectic phase showed no significant elongation. Furthermore, it is also known that the volume of liquid phase in the eutectic system depends on the total solute concentration present in water–ice at a specific temperature, which in this study is dominated by the ions from the buffer and the metal salts. The monomer salts contributed at most 14.6% of the total ion solutes (at 2.5 mM), a relatively small fraction of the final concentration. Nonetheless, a slight increase of the liquid phase is expected for the higher starting monomer concentrations, which could reduce the concentration enhancement achieved in the system and hence the extent of complex formation between monomers and hairpins.

The increased elongation rate in base-pairing systems reported here for the four nucleobases suggests that the template-directed RNA polymerization mediated by metal

catalysts may be possible for both pyrimidines and purines in the eutectic phase in water–ice and in general that heterogeneous media might have been essential for the polymerization of an early genetic material.

5. Abbreviations

ImpN	nucleoside (N) 5'-phosphoimidazolides
2-MeImpN	nucleoside (N) 5'-phospho-2-methyl-imidazolides
A; C; D; G; U; T and N	adenine; cytosine; 2,6-diaminopurine; guanine, uracil; thymine and N stands for any nucleobase.
EDTA	ethylenediaminetetraacetic acid
MES	2-(N-morpholino) ethane sulfonic
T4 PNK	T4 polynucleotide kinase
TEAA	triethylammonium acetate

Acknowledgement

This work was supported by the Grant (20050064DR) from the Laboratory-Directed Research and Development (LDRD) program at Los Alamos National Laboratory. The initial phases of this work were carried out at MGH, and were supported in part by NSF grant CHE 0434507 to JWS. JWS is an Investigator of the Howard Hughes Medical Institute. We thank Hans Ziock for his comments.

References

- [1] G.F. Joyce, *Nature* 418 (2002) 214–221.
- [2] G.F. Joyce, L.E. Orgel, *J. Mol. Biol.* 202 (1988) 677–681.
- [3] T. Wu, L.E. Orgel, *J. Am. Chem. Soc.* 114 (1992) 7963–7969.
- [4] A.R. Hill Jr., L.E. Orgel, *Orig. Life Evol. Biosph.* 23 (1993) 285–290.
- [5] H. Sawai, S. Totuka, H. Yamamoto, *Orig. Life Evol. Biosph.* 27 (1997) 525–533.
- [6] H. Sawai, M. Wada, *Orig. Life Evol. Biosph.* 30 (2000) 503–511.
- [7] T. Inoue, L.E. Orgel, *Science* 219 (1983) 859–862.
- [8] G.F. Joyce, Cold Spring Harbor, 1987, pp. 41–51.
- [9] T. Wu, L.E. Orgel, *J. Am. Chem. Soc.* 114 (1992) 5496–5501.
- [10] T. Wu, L.E. Orgel, *J. Am. Chem. Soc.* 114 (1992) 317–322.
- [11] A. Van Aerschot, R. Busson, P. Herdewijn, L.E. Orgel, *J. Am. Chem. Soc.* 121 (1999) 2653–2656.
- [12] L.E. Orgel, *Crit. Rev. Biochem. Mol. Biol.* 39 (2004) 99–123.
- [13] P.-A. Monnard, A. Kanavarioti, D.W. Deamer, *J. Am. Chem. Soc.* 125 (2003) 13734–13740.
- [14] R. Stribling, A.L. Miller, *J. Mol. Evol.* 32 (1991) 282–288.
- [15] A. Kanavarioti, P.-A. Monnard, D.W. Deamer, *Astrobiology* 1 (2001) 271–281.
- [16] A. Kanavarioti, L.F. Lee, S. Gangopadhyay, *Orig. Life Evol. Biosph.* 29 (1999) 473–487.
- [17] A. Kanavarioti, C.F. Bernasconi, D.J. Alberas, E.E. Baird, *J. Am. Chem. Soc.* 115 (1993) 8537–8546.
- [18] T.N. Solie, J.A. Schellman, *J. Mol. Biol.* 33 (1968) 61–77.
- [19] P.O.P. Ts'o, I.S. Melvin, A.C. Olson, *J. Am. Chem. Soc.* 85 (1968) 1289–1296.
- [20] H. Trinks, W. Schroder, C.K. Biebricher, *Orig. Life Evol. Biosph.* 35 (2005) 429–445.
- [21] A. Vlassov, B.H. Johnston, L.F. Landweber, S.A. Kazakov, *Nucl. Acids Res.* 32 (2004) 2966–2974.
- [22] S.R. Vogel, C. Richert, *Chem. Commun. (Camb)* (2007) 1896–1898.