ChemPubSoc Europe

DOI: 10.1002/cbic.201300773

Dry Polymerization of 3′,5′-Cyclic GMP to Long Strands of RNA

Matthias Morasch,[a] Christof B. Mast,[a] Johannes K. Langer,[a] Pierre Schilcher,[b] and Dieter Braun*[a]

Recent progress in the synthesis of nucleotides from prebiotically plausible precursors has opened up new ways to explain the origin of genetic matter. Mechanisms for the polymerization of nucleotides without the help of catalysts are, however, rare. Complementary to the experiments done by Costanzo et al., we found that drying 3′,5′-cyclic GMP leads to poly-G RNA strands with lengths of up to 40 nucleotides. We also show that the polymerization to long RNA strands is considerably more efficient under dry conditions than for cGMP polymerization in water. The length depends on the incubation time of dry nucleotides at temperatures of 40–80 °C. No enzymes or other catalysts are needed for successful polymerization.

Introduction

For decades, many facets of the molecular origin of life have been studied. Ribonucleic acid (RNA) has turned out to be a plausible starting point for molecular evolution as it exhibits both catalytic and storage capabilities.[1] In order to support this RNA world hypothesis,[2] many scenarios for the formation of nucleotides as RNA building blocks have been realized with great success.[3] In this context, alternative xeno-nucleobases (XNA) have also been discussed as RNA progenitors,[4] and chemically modified nucleobases have been used to study polymerase fidelity.[5] Also, a six-letter synthetic genetic system was developed and shown to work in terms of amplification, mutation and sequencing.[6] The next step towards an understanding of Darwinian molecular evolution of RNA was to research polymerization of ribonucleotides to long strands and the replication of RNA strands by template-directed copying mechanisms.[7–9] Such replication mechanisms were shown to work within protocells, and thus could have initiated basic competition for overall replication fitness.[10,11]

Here we have addressed the problem of nucleotide polymerization, which has been a subject of study for decades.[2] Although early experiments yielded only short oligomers,[12] later studies demonstrated the synthesis of strands with lengths of up to at least 55-mers, with the help of surface catalysis and monomer renewal.[13] Costanzo et al. reported a mechanism for fast polymerization of the free-acid form of 3′,5′-cyclic GMP to poly-G RNA, with lengths of up to 25 bases, by incubation of 6 mM cGMP in aqueous solution at 80 °C and in the absence of enzymes.[7,8] Although the prebiotic plausibility of cGMP is still under debate (because of its high energy content),[14] the simplicity of cGMP polymerization to long RNA strands without the need of any catalysts is appealing in the context of the origin of life. Despite the simplicity of this mechanism, however, these results have not been reproduced by other groups. The reported fast initial polymerization kinetics upon dissolving cGMP in water were unclear, and this raised speculation that there might have been contamination with polynucleotides formed in the production process.

We show here that the "drying" of cGMP monomers is responsible for this initial polymerization process and leads to the formation of very long poly-G strands. "Dry" conditions are defined as the absence of bulk water (residual hydration shells around the polymers are likely to remain). To avoid contamination with preformed polymers, custom-made cGMP samples that were never dried throughout the production process were used for our experiments. The proposed enzyme-free polymerization reaction yields at least deca- to 40-mers (low-micromolar range), depending on the incubation time in the dry state. It proceeds at temperatures as low as 40 °C and up to at least 80 °C. Within 2 h, long poly-G polymers are produced from pure cGMP without enzymes or catalysts.

Results

Unlike in a previous investigation,[8] we used the intercalating dye SYBR Gold to stain single-stranded RNA in polyacrylamide gels. This allowed us to visualize long strands (down to several nanograms), provided these were at least ten nucleotides in length. For shorter polymers, we also obtained data from ESI-MS-calibrated MALDI-TOF measurements, which are ideal to

[a] M. Morasch, Dr. C. B. Mast, J. K. Langer, Prof. D. Braun
Systems Biophysics, Center for NanoScience, Fakultät für Physik
Ludwig-Maximilians-Universität München
Amalienstrasse 54, 80799 München (Germany)
E-mail: dieter.braun@lmu.de
Homepage: http://www.biosystems.physik.uni-muenchen.de

[b] P. Schilcher
Zentrallabor für Proteinanalytik, Adolf-Butenandt-Institute
Ludwig-Maximilians-Universität München
Schillerstrasse 42, 80336 München (Germany)

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cbic.201300773.

resolve lengths of fewer than ten nucleotides. As both techniques are prone to show false-positive polymerization signals due to stacking of nucleotides (particularly pronounced with purine bases),[15] we also performed digestion by RNAse T1, which is specific to 3′–5′-linked RNA strands.

Our experiments showed that highly effective polymerization takes place in the dry state. To avoid contamination by preformed RNA polymers, we started with 0.72 mM aqueous cGMP (custom-made by BioLog, Bremen, Germany; produced from its sodium salt, and HPLC purified by the manufacturer). The sodium salt version is the optimum starting material as it does not polymerize under aqueous or dry conditions.[8] Although standard techniques for cGMP production use a drying step after the conversion, our custom-made cGMP was specifically manufactured without sample drying after the formation of the polymer-free acid form of cGMP. Samples were kept frozen until use. We concentrated the sample in water to 12 mM and could not find any RNA strands (Figure 1 A, lane 1). To prove that drying is the determining step for initial polymerization, we incubated dried samples at 40–80 °C. Lanes 2–5 in Figure 1 A show stained cGMP in a polyacrylamide gel after incubation in a vacuum centrifuge at 50 °C for 20 min to 15 h in the dry state. Drying and incubation of all samples within this temperature range resulted in similar polymerization, usually after 4–5 h in the dry state. We could not detect RNA strands longer than ten bases in samples that contained small amounts of residual bulk water. This suggests that cGMP polymerization under (near) dry conditions is significantly more efficient than polymerization in water. To quantify the results we analyzed polymerization kinetics in the dry state (Figure 1 B). Although very weak signals for deca- to 13-mers were found shortly after drying (20 min), more and longer strands formed after further incubation of the dry sample. After 15 h at 50 °C, deca- to 17-mers formed in considerable amounts (a few nanograms per microlitre for 12 mM cGMP). Figure 1 C shows the increased concentration for the deca-, 14-, and 17-mers in detail, estimated from separate calibration measurements.

The ribonucleic character of these polymers was shown by incubation of the samples with RNase T1, an endonuclease that cleaves at the 3′-ends of G residues (Figure 2 A). Our results confirm a previously RNase sensitivity report for incubated samples that were dried for 10 h and resuspended to 12 mM.[8] Different amounts of RNase T1 were added, and the samples were incubated at 22 °C for 20 h; the polymers were fully digested at high RNase concentrations. To compare the RNase activity with that for pure poly-G RNA, predried 2 mM cGMP and 10 μM commercial poly-G 16-mer samples were incubated for 5 h at 37 °C in Tris·HCl (pH 7.4) with 2 mM EDTA. The total mass of the commercial 16-mer was estimated to be

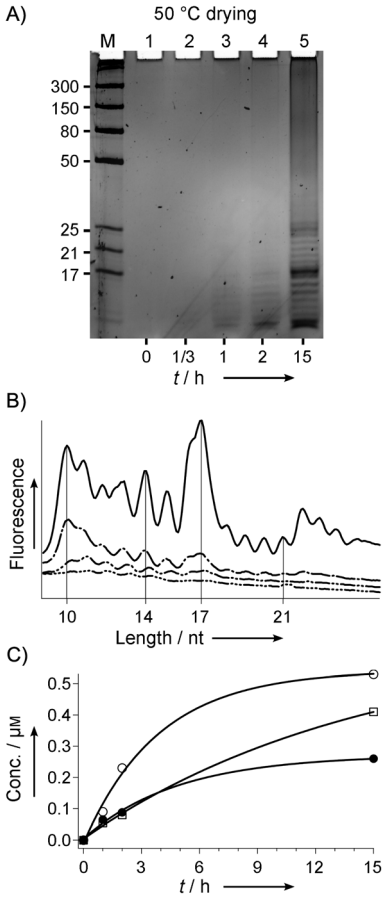

Figure 1. Polymerization of cGMP in the dry state. Samples were dried for different durations in a vacuum centrifuge at 50 °C and subsequently resuspended to 12 mM. A) Gel electrophoresis. No polymers were found before drying (lane 1); some deca- to 13-mers were visible shortly afterwards (lane 2: 20 min, lane 3: 60 min). For longer incubation times in the dry state (lane 4: 120 min, lane 5: 15 h) more and longer polymers formed. ssRNA markers in the leftmost lane. B) Fluorescence analysis of the same samples (–····–: 20 min, –···–: 1 h, –·–: 2 h, ——: 15 h). C) Concentration profile for three chosen lengths (○: decamer, ●: 14-mer, □: 17-mer) from the above samples.

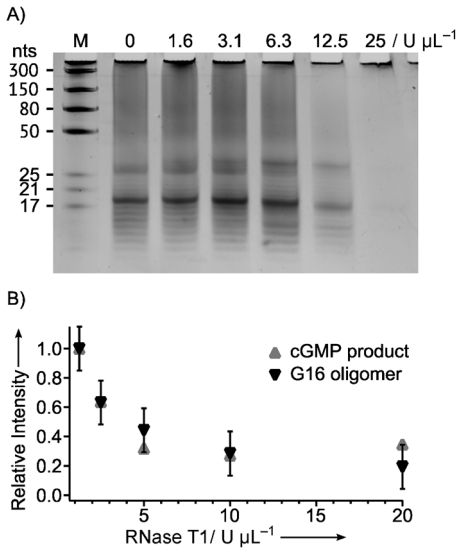

Figure 2. RNase T1 treatment of the created polymers. A) Complete digestion of polymers created by cGMP drying. RNase (0–25 U μL^{-1}) was incubated (20 h, 22 °C) with a 12 mM cGMP sample that was dried for 10 h. At high RNase concentrations, the sample was completely digested. B) Different amounts of RNase T1 were added to samples of 2 mM cGMP (Sigma–Aldrich) and 10 μM commercial 16-mer poly-G for 5 h at 37 °C in 100 mM Tris·HCl (pH 7.4) containing 2 mM EDTA. At high RNase concentration, commercial 16-mer ("G16") and polymers obtained by drying cGMP were similarly digested. (See Figure S1 for the corresponding polyacrylamide gel image.)

similar to that of polymers obtained by drying the cGMP sample. Figure 2B shows that the samples behaved similarly when exposed to the same amounts of RNase (corresponding gel image in Figure S1 in the Supporting Information). Similar tests have been performed in the past; these also proved the sensitivity of the polymerization product to exonuclease (phosphodiesterase I).[8]

To verify the polyacrylamide gel results, a MALDI-TOF analysis of dried and nondried samples was performed (Figure 3). This showed formation of polymers identical to those previously found from a mostly pure sample by drying.[8] The polymers are identified as *n* times the mass of a cGMP molecule plus one H and one OH molecule. ESI-TOF analysis was used to confirm and calibrate the MALDI results (Figure S2). In addition to finding the predicted masses of poly-cGMP plus water, we also observed what we assumed to be stacking of the polymers.[8] It should be noted that the initial sample without dry incubation did not show these mass peaks. We detected mostly monomers and (to a lesser extent) dimers but nothing higher (Figure S3), although the concentration of monomers capable of stacking was the same. This suggests that the poly-cGMP peaks might not be stacked cGMP monomers but more strongly bound molecules. Overall, we conclude that the observed polymers are the same as those found by Costanzo et al.

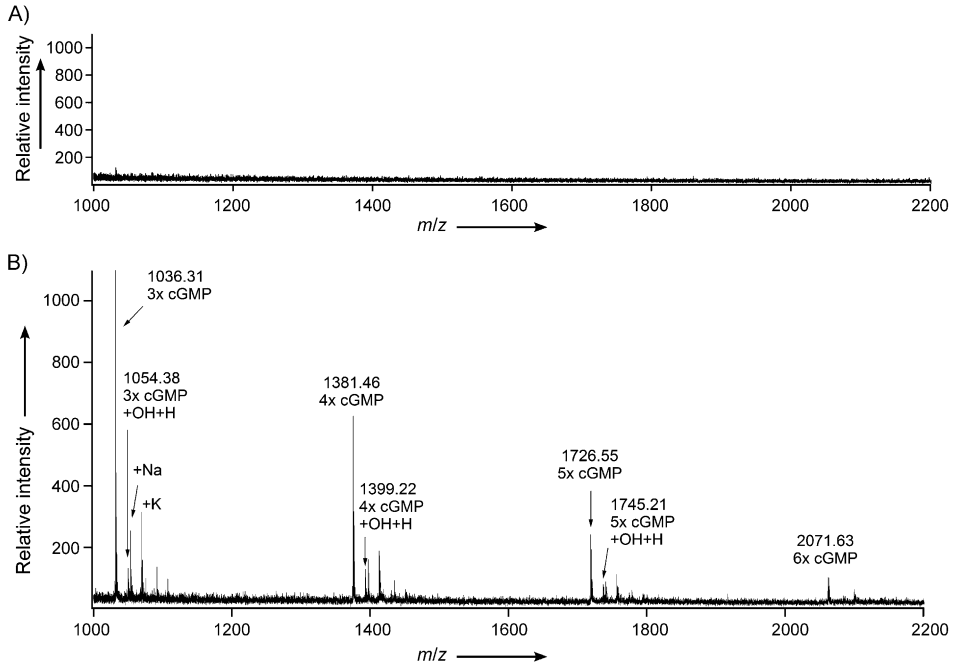

Figure 3. MALDI-TOF MS analysis of A) pure and B) polymerized samples. From the trimer onwards, the pure sample does not show any signs of polymerization. The dried sample exhibits the expected spectrum of *n* times the mass of cGMP, together with a water molecule or sodium/potassium ions. The peaks for cGMP alone were assumed to be stacked monomers, with OH and H added during the proposed polymerization reaction.[8]

Dry samples from other manufacturers have shown the same polymer length distributions in polyacrylamide gel analysis. A sample that was predried by the manufacturer was used in experiments with cGMP in water.[7,8] In our sample we found the same polymerization and intensity profile as for the custom-made and subsequently dried version. Figure 4 shows the result for 6 mM cGMP (Sigma–Aldrich #G7504; dried by the manufacturer), which was also used by Constanzo et al. Lane 1 shows the signal from the sample directly after suspension in water; lane 2 is the same but with a longer electrophoresis run to allow a better resolution for longer bands. Both show long strands (up to 40 nucleotides) with intensity peaks at around 17 and 24 bases. Note that the poly-G strands migrated more slowly than strands with diverse sequences (used as markers). Counting from the lowest to highest visible bands, however, revealed lengths around 40 bases. Based on the fluorescence signals, longer strands seem very likely but could not be resolved or detected in the much less sensitive MALDI-TOF analysis. After incubation at 80 °C for 1 h (lane 3; lane 4 for better resolution at longer bands), a fluorescence increase for longer strands is visible. However, it is unclear whether this elongation effect is due to polymerization of cGMP during incubation or, for example, a consequence of better solubility of clustered poly-G constructs. As the majority of the found RNA strands were already present before incubation, a valid blank measurement was not possible for this sample.

We also tested the polymerization of our polymer-free cGMP samples in aqueous solution according to the protocols of Costanzo et al.[7] Figure 5 shows

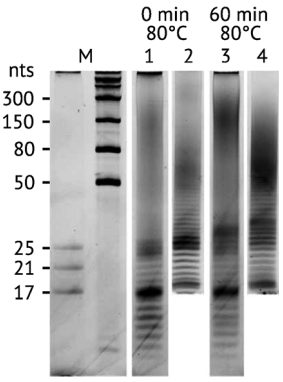

Figure 4. Dried cGMP shows 3′–5′-linked oligomers without further treatment. Gel electrophoresis of dried cGMP before (lanes 1, 2) and after (lanes 3, 4) 1 h incubation at 80 °C. Size markers ("M", 17–300 bases) at the left show sizes for all four lanes. For each pair, the right-hand lane is a longer run of the same sample giving a better resolution of longer strands.

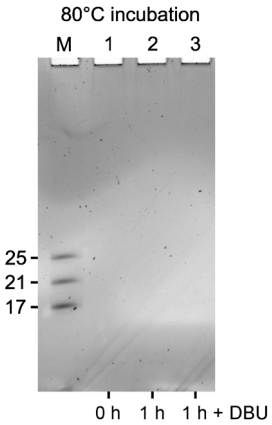

Figure 5. No polymerization (within the achievable sensitivity) is observed in aqueous solution. Incubation of cGMP in water (lane 2), and after addition of 10 mM DBU (lane 3) at 80 °C for 1 h. The control (lane 1) shows no pre-existing RNA strands; no formation of decamers (or higher) was observed.

a fresh 0.72 mM cGMP sample that was dried in a rotary evaporator, resuspended in RNase free water to 12 mM, and subsequently incubated at 80 °C for 1 h. The final drying step in the rotary evaporator took only 5 min, and was therefore short enough to avoid a measurable dry polymerization. In all our experiments we did not observe polymers within the sensitivity limit after incubation for 1 h at 80 °C, even with prior addition of 10 mM 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU; Sigma–Aldrich #139009), which has been reported to enhance cGMP polymerization.[8]

Discussion

Based on the gel data we estimate that at least 2% of the cGMP monomers polymerized to strands longer than ten nucleotides, after only 15 h in the dry state at 50 °C. The formation of 50- to 80-mers (in small amounts) is also suggested by the gel images. Formation of these long poly-G strands is triggered by drying the free acid form of 3′,5′-cyclic GMP. The determining factor is the time in the dry state, at moderately elevated temperatures.

No polymerization of cGMP within the detectable limit (~1 ng) could be detected in water; this difference from previous results could be explained by the different labeling method used in this work:[7,8] RNA strands shorter than ten bases cannot be excluded with SYBRGold labeling, so polymerization reactions in the short regime of mostly octamers (as found by Costanzo et al.) could not be detected. Also, the mechanism that causes changes in the fluorescent profile after incubation of the predried version is unclear (Figure 4). It might arise from polymerization of monomers, but is more likely a consequence of better solubility of the samples in water. Mass spectrometry, however, showed clear emergence of the previously observed polymers after drying a pure sample. We therefore suggest that for all polymerization experiments with cGMP, samples that are produced by an all-liquid manufacturing process should be used, to avoid false-positive results for RNA polymers.

Experimental Section

Monomer samples: Guanosine 3′,5′-cyclic monophosphate (free acid) was from Sigma–Aldrich (powder, #G7504), and custom-made cGMP (free acid) was from BioLog (Bremen, Germany; 0.72 mM in aqueous solution). The latter had been manufactured from cGMP sodium salt, which has been reported not to polymerize.[8] For higher concentrations, samples were dried, or concentrated in a Laborota 4000 rotary evaporator (Heidolph Instruments, Schwabach, Germany) at 37 °C and resuspended in RNase-free water (Sigma–Aldrich, #W4502) to 11–13 mM.

Polymerization procedure: Highly concentrated samples (~12 mM, 50–200 μL) were placed in an RVC 2-25 vacuum centrifuge (Martin Christ Gefriertrocknungsanlagen, Osterode, Germany) equipped with a CT 02-50 SR condensation trap (Christ) and an MZ 2C vacuum pump (Vacuubrand, Wertheim, Germany), or allowed to air dry at various temperatures. Resuspended samples were briefly heated to 60 °C to promote dissolving of the powder.

Gel electrophoresis: Polyacrylamide gels were mixed from a 40% acrylamide/bis-acrylamide (29:1) stock solution (Carl Roth, #A515.1). The gels (~25% acrylamide) contained urea (~7 M; Carl Roth, #X999.2) and TBE (1×; Carl Roth, #3050.1). Gels were labeled with SYBR Gold (1×, from 10000× stock; Invitrogen, #S-11494) for 10–15 min and analyzed in a self-modified Darkroom Hood (Herolab, Wiesloch, Germany, #RH-5).

Gel staining efficiency with SYBR Gold was checked by using a decamer and a poly-G RNA ladder (kindly provided by E. Di Mauro, Istituto di Biologia e Patologia Molecolari, Rome). We were able to detect deca- to 24-mers at a few nanograms; a pentamer could not be detected. Concentration calibration was performed with a 40-mer of known concentration. We assumed length-dependent but sequence-independent staining with SYBR Gold. The fluorescence of this 40-mer was compared with that for the 25 bp ladder and the poly-G bands, thus allowing calculations of poly-G concentration to obtain a rough estimate of polymerization efficiency (systematic errors ~50%).

RNase analysis: RNase T1 (Invitrogen, #AM2280) treatment of the samples was performed at room temperature for 20 h, or at 37 °C for 5 h, in Tris·HCl (100 mM, pH 7.4). T1 cleaves at the 3′-ends of G residues; our polymers are sensitive to this. Units are defined as given by the manufacturer. A commercially synthesized poly-G 16-mer (Biomers, Ulm, Germany) was used to compare the RNase T1 sensitivity of the cGMP products to commercial poly-G RNA.

Mass spectrometry: MALDI-TOF MS analysis was performed on a Voyager STR system (AB SCIEX, Framingham, MA). Pure (12 mM) and polymerized cGMP samples were mixed with a 3-hydroxypicolinic acid/diammonium hydrogencitrate (9:1) matrix (Fluka/Sigma–Aldrich, #56197 and #09831, respectively) and analyzed. Before the analysis, a standard cleanup of each sample was performed with ZipTip (Millipore, #ZTC18M960); this removed most of the monomers. ESI-TOF MS analysis was performed on a QSTAR XL quadrupole-TOF hybrid mass spectrometer (AB SCIEX). The samples were again purified with ZipTips, and analyzed in a methanol (50%, LC-MS grade; Roth, #AE71.1), elution in trifluoroacetic acid (0.1%; Roth, #P088.1).

Acknowledgements

We thank Ernesto Di Mauro, Giovanna Costanzo, and Samanta Pino for discussions and valuable comments on the manuscript

at various stages, as well as Axel Imhof and Ignasi Forné for support with the mass spectrometry analysis. We thank John Sutherland for helpful comments on the polymerization process. Financial support from the NanoSystems Initiative Munich, the Ludwig-Maximilians-Universität Munich Initiative Functional Nanosystems, and the European Research Council Starting Grant is acknowledged.

Keywords: cGMP · molecular evolution · nucleotides · polymerization · RNA

[1] W. Gilbert, *Nature* **1986**, *319*, 618.
[2] L. E. Orgel, *Crit. Rev. Biochem. Mol. Biol.* **2004**, *39*, 99 – 123.
[3] a) M. W. Powner, B. Gerland, J. D. Sutherland, *Nature* **2009**, *459*, 239 – 242; b) M. W. Powner, J. D. Sutherland, J. W. Szostak, *J. Am. Chem. Soc.* **2010**, *132*, 16677 – 16688; c) H. D. Bean, F. A. L. Anet, I. R. Gould, N. V. Hud, *Origins Life Evol. Biospheres* **2006**, *36*, 39 – 63.
[4] H. Yu, S. Zhang, J. C. Chaput, *Nat. Chem.* **2012**, *4*, 183 – 187.
[5] P. Strazewski, C. Tamm, *Angew. Chem. Int. Ed. Engl.* **1990**, *29*, 36 – 57; *Angew. Chem.* **1990**, *102*, 37 – 59.
[6] Z. Yang, F. Chen, J. B. Alvarado, S. A. Benner, *J. Am. Chem. Soc.* **2011**, *133*, 15105 – 15112.
[7] G. Costanzo, S. Pino, F. Ciciriello, E. Di Mauro, *J. Biol. Chem.* **2009**, *284*, 33206 – 33216.
[8] G. Costanzo, R. Saladino, G. Botta, A. Giorgi, A. Scipioni, S. Pino, E. Di Mauro, *ChemBioChem* **2012**, *13*, 999 – 1008.
[9] a) N. V. Hud, S. S. Jain, X. Li, D. G. Lynn, *Chem. Biodiversity* **2007**, *4*, 768 – 783; b) L. E. Orgel, *Nature* **1992**, *358*, 203 – 209; c) M. Hey, C. Hartel, M. W. Göbel, *Helv. Chim. Acta* **2003**, *86*, 844 – 854; d) K. Adamala, J. W. Szostak, *Science* **2013**, *342*, 1098 – 1100; e) C. Deck, M. Jauker, C. Richert, *Nat. Chem.* **2011**, *3*, 603 – 608.
[10] a) S. S. Mansy, J. P. Schrum, M. Krishnamurthy, S. Tobé, D. A. Treco, J. W. Szostak, *Nature* **2008**, *454*, 122 – 125; b) J. Attwater, A. Wochner, V. B. Pinheiro, A. Coulson, P. Holliger, *Nat. Commun.* **2010**, *1*, 1 – 8.
[11] I. A. Chen, R. W. Roberts, J. W. Szostak, *Science* **2004**, *305*, 1474 – 1476.
[12] J. Morávek, *Tetrahedron Lett.* **1967**, *8*, 1707 – 1710.
[13] a) J. P. Ferris, A. R. Hill Jr., R. Liu, L. E. Orgel, *Nature* **1996**, *381*, 59 – 61; b) W. Huang, J. P. Ferris, *Chem. Commun.* **2003**, 1458 – 1459.
[14] R. N. Goldberg, Y. B. Tewari, *J. Chem. Thermodyn.* **2003**, *35*, 1809 – 1830.
[15] B. T. Burcar, L. M. Cassidy, E. M. Moriarty, P. C. Joshi, K. M. Coari, L. B. McGown, *Origins Life Evol. Biospheres* **2013**, *43*, 247 – 261.

Received: December 11, 2013
Published online on February 27, 2014