RNA Chemistry

International Edition: DOI: 10.1002/anie.201506592
German Edition: DOI: 10.1002/ange.201506592

Copying of RNA Sequences without Pre-Activation

*Mario Jauker, Helmut Griesser, and Clemens Richert**

Abstract: *Template-directed incorporation of nucleotides at the terminus of a growing complementary strand is the basis of replication. For RNA, this process can occur in the absence of enzymes, if the ribonucleotides are first converted to an active species with a leaving group. Thus far, the activation required a separate chemical step, complicating prebiotically plausible scenarios. Here we show that a combination of a carbodiimide and an organocatalyst induces near-quantitative incorporation of any of the four ribonucleotides. Upon in situ activation, adenosine monophosphate was found to also form oligomers in aqueous solution. So, both de novo strand formation and sequence-specific copying can occur without an artificial synthetic step.*

Ribonucleic acid (RNA) is found in all cells. It can encode genetic information, but it is also involved in protein synthesis,[1] catalysis,[2] and the regulation of gene expression.[3] Because RNA can fulfil several pivotal roles in biochemistry, it is possible that life started with a so-called "RNA world".[4] It is therefore important to ask how oligoribonucleotides may form in the absence of enzymes, and how the genetic information they contain may be copied into complementary strands without the catalytic action of a polymerase. Current-day metabolism generates nucleoside triphosphates for replication, transcription, and encoded protein synthesis, but nucleoside triphosphates are largely unreactive in the absence of enzymes.[5]

The most common way to induce enzyme-free oligomerization of a ribonucleotide is to activate it in a separate chemical reaction, producing a monomer with an organic leaving group, or an anhydronucleotide.[6] The product is isolated and then used in a subsequent oligomerization step (Figure 1). Following this protocol, strands have been shown to form in the presence of mineral surfaces[7] or when exposed to elevated temperatures and/or organic solvents.[8,9] Heterogeneous media favor the incorporation of all four nucleotides,[10,11] and long polymers were found in eutectic phases.[11] Pre-activated nucleotides were also used to demonstrate that copying of a given template sequence into a complementary strand can occur without enzymes, mostly in the form of

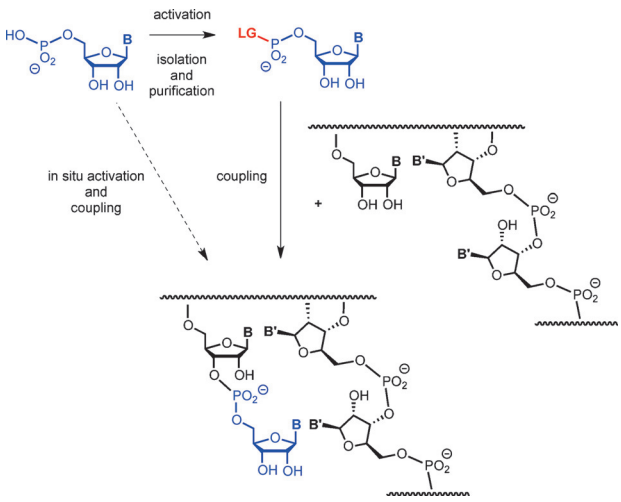

Figure 1. Copying of an RNA sequence via enzyme-free primer extension, with or without pre-activation of the ribonucleotide monomer. LG = Leaving group.

enzyme-free primer extension. Pre-activated nucleotides typically used for copying include imidazolides,[12] methylimidazolides,[13] and oxyazabenzotriazolides.[14] We recently showed that when the latter react with immobilized template–primer duplexes, near-quantitative incorporation of any of the four nucleotides (A/C/G/U) is found.[15]

Discontinuous, two-step syntheses require complicated prebiotic scenarios. Conditions that induce activation and chain extension simultaneously make presumed prebiotic processes more likely. It is therefore important to ask whether such conditions exist and what activating chemistry supports them. Uronium salts are known to activate nucleotides[16,17] for subsequent coupling, but they are usually employed in organic solvents, and it is unclear whether they are prebiotically relevant. A combination of a phosphine and pyridyldisulfide has also been used to activate nucleotides,[18,19] but this approach is not suitable for in situ activation. Simple inorganic activation agents like COS have been shown to induce the formation of aminoacylnucleotides,[20] but not RNA oligomers. Simple reagents are also problematic because the potential for side reactions in complex reaction mixtures involving highly reactive reagents or elevated temperatures is very high. Complex one-pot reactions often lead to intractable mixtures or tar.[21]

One class of activation reagents that is of prebiotic relevance is carbodiimides. Carbodiimide is a tautomer of cyanamide, a compound formed under presumed prebiotic conditions.[22–24] Ligations between strands terminating in amino groups and phosphates have been induced by carbodiimides, including replication reactions.[25–27] It is known that pre-activation can be induced with a conventional condensa-

[*] Dipl.-Chem. M. Jauker, Dipl.-Chem. H. Griesser, Prof. C. Richert
Institut für Organische Chemie, Universität Stuttgart
70569 Stuttgart (Germany)
E-mail: lehrstuhl-2@oc.uni-stuttgart.de

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201506592.

tion agent, such as *N*-ethyl-*N*′-(3-dimethylaminopropyl)carbodiimide (EDC) at pH 5.5,[28] but no genetic copying occurs under these conditions. Untemplated oligomerization up to tetramers was recently reported in homogeneous solution at pH 6.5, accompanied by massive side reactions,[29] but not genetic copying. In the 1960s, template-directed oligomerizations, not genetic copying, had been studied using in situ activation, without an organocatalyst, but the yields were low and the oligomers obtained were too short for duplex formation.[30,31] These results are understandable because free ribonucleotides were shown to act as inhibitors of enzyme-free primer extension.[15,32]

Here we show that a combination of a carbodiimide and an *N*-alkyl heterocycle as catalyst induces efficient copying reactions on RNA templates using unactivated, free ribonucleotides. While additives like free imidazole give fairly unreactive imidazolides, alkylated imidazole as the organocatalyst can give a highly reactive imidazolium species. Both primer extension on preformed RNA strands and the untemplated de novo oligomerization of ribonucleotides to RNA strands were observed. This shows for the first time that genetic copying can occur without the synthetic step of preactivation of the ribonucleotide. Both strand formation and copying of genetic information occur without the intervention of a synthetic chemist. Further, it was found that the conditions described here unleash other intrinsic reactivities of nucleotides and amino acids, leading to the spontaneous formation of peptidyl RNAs and cofactors, as reported in Ref. [33].

We started by studying primer extension with guanosine 5′-monophosphate (GMP) as the monomer and cytosine as the templating base in the RNA template (Figure 2a). This is the most favorable case among the four copying reactions, as GMP pairs more strongly than the other three ribonucleotides (AMP, CMP, and UMP).[32] Still, copying is demanding, as it requires both intrinsic reactivity and delicate molecular recognition in order to be sequence selective.[34] Initially the challenge of achieving activation and coupling in one solution seemed all but insurmountable. The unactivated ribonucleotide was known to act as a competitive inhibitor to primer extension, blocking the extension site.[15] Since NMR monitoring indicated that the vast majority of ribonucleotide molecules in solution remain unactivated upon treatment with EDC, reactions were expected to be inefficient. Further, the temperature optimum is different for the two steps, with coupling best performed at low temperatures,[14] but activation at room temperature or above. Also, while activation is favored by organic solvents and acidic pH,[28] coupling requires aqueous buffer to achieve the template effect and neutral or basic pH to ensure nucleophilic reactivity at the primer terminus.

We monitored primer extension with MALDI-TOF mass spectrometry under conditions that allow for quantitative detection.[35] Initially, no extension of primer **2** by GMP was detectable, even in high-salt buffer containing 0.8 M EDC, and our experimental work aimed at high-yielding reactions was unsuccessful, in agreement with the low yields reported in the early literature.[30,31] Only when an organocatalyst was used did detectable copying set in.

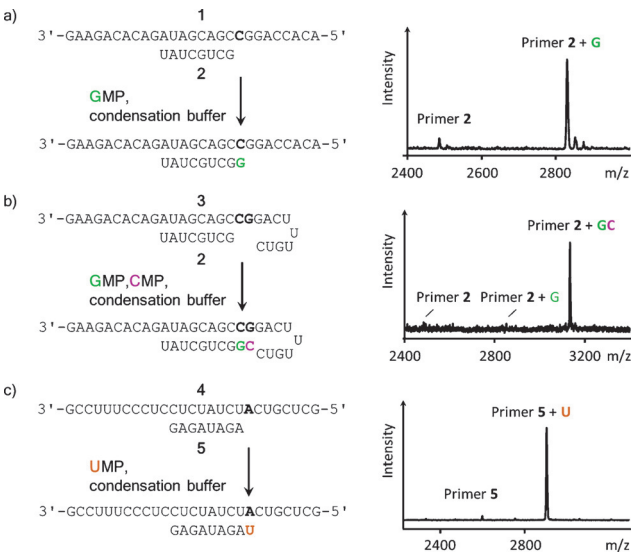

Figure 2. High-yielding incorporation of three of the four ribonucleotides opposite their complementary bases in aqueous condensation buffer with in situ activation. a) Extension of primer by GMP after 2 d at 20 °C; b) consecutive extension of primer by GMP and CMP after 6 d at 20 °C; c) extension of primer by UMP after 5 d at 0 °C. In each case, the reaction scheme is shown next to a MALDI mass spectrum; conditions: 0.8 M EDC, 0.1–0.15 M ethylimidazole, 0.4 M NaCl, 0.2–0.5 M HEPES buffer, and/or 0.08–0.16 M $MgCl_2$. See the Supporting Information for details.

Adding imidazole or 2-methylimidazole led to 2% conversion after 24 h. Presumably, imidazolides of the nucleotide, known monomers for copying,[29,36] were formed in the solution, but steady-state levels of these species were not high enough for efficient extension. With 1-methyladenine as catalyst,[37] 12% primer extension was observed after 1 d at 20 °C. When 1-ethylimidazole was used, the yield of extended primer was 32% after 24 h (Table 1). A subsequent screen of pH values gave an optimum of 7.5.

Next we varied the salt conditions. In the absence of a divalent metal ion, no primer extension was observed. In the presence of 80 mM Mn^{2+}, 10% extended primer was detected after 24 h at pH 8. With 80 mM Ca^{2+}, 25% conversion was found after 24 h. The fastest reaction was observed with Mg^{2+}, with 70% extension after 1 d and 90% conversion after 2 d (mass spectrum in Figure 2a). With the same buffer and a mixture of GMP and CMP, two consecutive high-yielding copying steps occurred on a hairpin template that prevents uncontrolled further extension (Figure 2b). When uridine 5′-monophosphate (UMP), the most weakly pairing nucleotide,[32] was offered to a primer on a template with adenine as the first templating base, 75% incorporation occurred after 2 d. Lowering the temperature to 0 °C, conditions close to those of eutectic ice phases, led to near-quantitative incorporation of UMP after 5 d in 0.5 M HEPES buffer (Figure 2c, last entry in Table 1). High concentrations are characteristic of eutectic ice phases, and the beneficial effect of low temperatures for copying was known.[17] The buffer thus optimized was dubbed "general condensation buffer" (0.8 M EDC, 0.15 M 1-ethylimidazole, 0.08 M $MgCl_2$, 0.5 M HEPES, pH 7.5) and was used in all subsequent assays.

Table 1: Results of primer extension assays in solution.[a]

Template, primer[b]	NMP[c]	Cat.[d]	Buffer[e]	pH	*T* [C°]	*t* [h]	Extension [%]
1, **2**	G	–	A	6.7	20	24	<1
1, **2**	G	Im	A	6.7	20	24	2
1, **2**	G	2-MeIm	A	6.7	20	24	8
1, **2**	G	MeAde	A	7.7	20	24	12
1, **2**	G	1-EtIm	A	6.7	20	24	32
1, **2**	G	1-EtIm	A	5.5	20	24	44
1, **2**	G	1-EtIm	A	6.5	20	24	62
1, **2**	G	1-EtIm	A	7.0	20	24	63
1, **2**	G	1-EtIm	A	7.5	20	24	74
1, **2**	G	1-EtIm	A	7.9	20	24	73
1, **2**	G	1-EtIm	A	7.9	20	48	90
3, **2**	G+C	1-EtIm	A	7.5	20	140	95
4, **5**	U	1-EtIm	A	7.5	20	50	75
4, **5**	U	1-EtIm	B	7.5	0	120	92

[a] For conditions, see General Protocol 1 in the Supporting Information. [b] 60 μM Template, 50 μM primer. [c] 20 mM GMP; or 150 mM UMP; or 50 mM GMP and 50 mM CMP. [d] Catalysts: Im, imidazole; 2-MeIm, 2-methylimidazole; MeAde, 1-methyladenine; 1-EtIm, 1-ethylimidazole (0.1 M of either for Buffer A; 0.15 M 1-ethylimidazole for Buffer B). [e] Buffer A = 0.2 M HEPES, 0.4 M NaCl, 0.16 M $MgCl_2$, 0.8 M EDC; Buffer B = 0.5 M HEPES, 0.08 M $MgCl_2$, 0.8 M EDC.

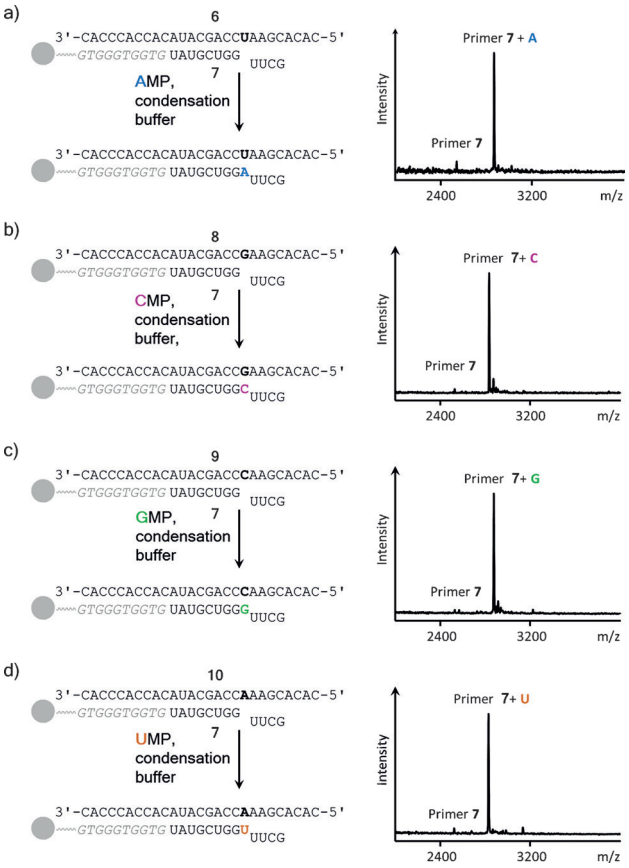

Figure 3. High-yielding incorporation of any of the four ribonucleotides opposite their complementary bases in condensation buffer. Reaction schemes are shown next to MALDI mass spectra. The oligodeoxynucleotide that immobilizes the primer–template duplex on beads is shown in italics; the tetramer downstream of the extension site limits the chain growth to a single step, facilitating analysis.[15] a) Extension by AMP, spectrum after 21 d at 0°C; b) extension by CMP, spectrum after 8 d at 0°C; c) extension by GMP, spectrum after 8 d at 0°C; d) extension by UMP after 21 d at 0°C. Conditions: 0.8 M EDC, 0.15 M 1-ethylimidazole, 0.5 M HEPES buffer, and 0.08 M $MgCl_2$. See the Supporting Information for details.

In mass spectra of copying assays with AMP as the monomer, strong peaks for oligoadenylates were observed. This was unexpected because similar assays with pre-activated monomers had not shown oligomers.[14,15,17] We immobilized template–primer duplexes on magnetic beads via hybridization (Figure 3).[15,38] Now, magnetic separation, followed by washing the beads and denaturation allowed an unobstructed view on copying reactions, even for A (Figure 3a). Kinetics showed that incorporation of AMP opposite U occurred with a half-life time of approximately 3 d. This is less than one order of magnitude slower than similar reactions involving pre-activated forms of AMP.[15,39]

To confirm that the copying conditions found are general, primer **7** was hybridized to three other templates, each offering a different template base (C, G, or U). The resulting primer–template duplexes were allowed to react with the corresponding ribonucleotide (AMP, CMP, GMP or UMP). In each case near-quantitative primer extension was observed (Figure 3b–d). Despite the high concentrations, which are reminiscent of what is found in eutectic ice phases,[11] the reactions were surprisingly well behaved, with little to no side products detectable by MALDI MS. This confirmed that enzyme-free copying of RNA does not require pre-activation.

Finally, we performed an exploratory study on strand formation occurring under our in situ activation conditions (vide supra). Using a combination of ion-exchange HPLC and MALDI MS, we tested for the formation of oligomers in solutions of AMP, UMP, or CMP (150 mM each) after one week at 0°C. For GMP, a 20 mM solution was used to avoid precipitation and aggregation. Only for AMP >90% of the monomer was converted after 7 d, whereas for CMP ≥25% of the monomer remained, and ≥40% of UMP was unreacted after 7 d. In the case of GMP, peaks were too broad for unambiguous assignment, but it appeared that significant concentrations of both monomers and short oligomers were present. Figure 4 shows results from an assay with AMP run for 30 d at 0°C. Peaks for chains of up to at least nine AMP residues are discernable. The formation of mixed sequences that can encode genetic information from any of the four ribonucleotides (A/C/G/U) is presented in Ref. [33]. Taken together, our results confirm that "general condensation conditions" induce both high-yielding copying reactions and the untemplated de novo formation of RNA strands.

In conclusion, we report that the combination of a carbodiimide and an N-alkylated heterocycle as the catalyst induces high-yielding genetic copying. The efficiency of the reaction strongly depends on the choice of the covalent catalyst. A positively charged imidazolium species, expected to form with 1-ethylimidazole, is apparently much more reactive than the

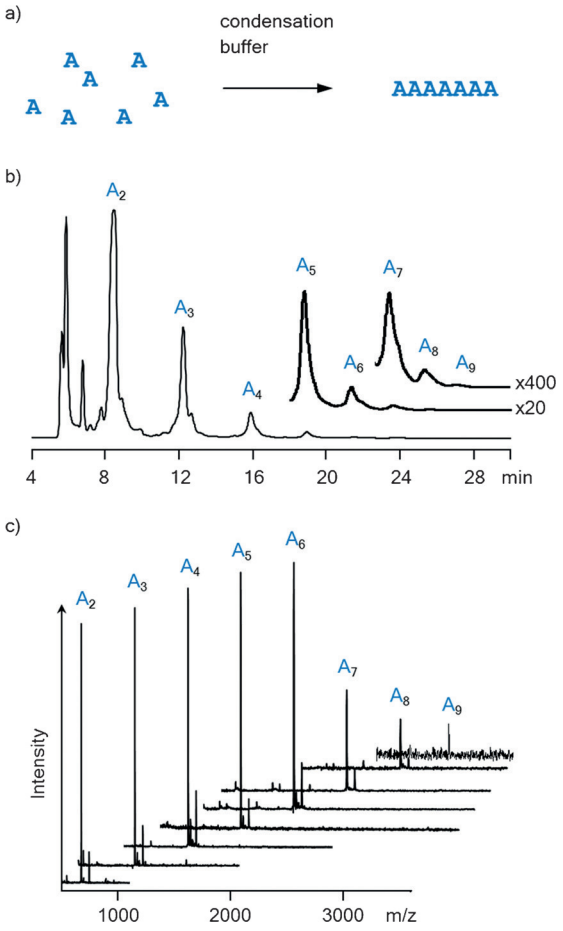

Figure 4. Oligomerization of AMP in condensation buffer at 0 °C after 30 d. a) Reaction scheme. b) Ion-exchange HPLC chromatogram at $\lambda_{det} = 260$ nm. c) Overlay of MALDI-TOF mass spectra of HPLC fractions, showing peaks of oligoribonucleotides with 2–9 residues.

well-established imidazolides. Chain growth is not limited to template-directed reactions. Oligomerization of ribonucleotides occurs in the absence of a mineral surface. It takes place at a fast enough rate to provide strands that may act as templates or primers, but slowly enough not to dominate the reaction landscape, so that is still allows for sequence-selective copying steps. The fast oligomerization of AMP is an interesting contrast to the slow primer extension on A-rich templates,[40,41] the second step of replication scenarios with strand formation and subsequent copying, suggesting a kinetic compensation. Based on our results, a more conclusive picture of the emergence of RNAs can be formulated, and much simpler experimental setups can be used to study their formation. A number of other processes leading to pivotal biomolecules occur under the same general condensation conditions, as reported in Ref. [33].

Experimental Section

Condensation buffer: The optimized reaction medium, referred to as "general condensation buffer" was an aqueous solution of HEPES (0.5 M), $MgCl_2$ (0.08 M), 1-ethylimidazole (0.15 M), and the appropriate concentration of reactants, adjusted to pH 7.5. A fresh aliquot (65 μL) of this solution was added to EDC hydrochloride (10 mg, 52 μmol) to give an initial EDC concentration of approximately 0.8 M.

Primer extension: A suspension of beads with capture oligonucleotide (5 μL, 5 mg mL^{-1}) in HEPES buffer (0.5 M with 0.08 M $MgCl_2$, pH 7.5) was treated by addition of solutions of template (0.6 μL, 100 μM, 60 pmol) and primer (0.5 μL, 100 μM, 50 pmol) and left at 0 °C for 15 min. The supernatant was aspirated, and 5 μL of a freshly prepared condensation buffer containing either AMP (0.15 M), GMP (0.02 M), CMP (0.15 M), or UMP (0.15 M) at pH 7.5 was added. The mixture was transferred to a vessel containing 5 nmol of the downstream-binding tetramer, vortexed for 5 s, and incubated at 0 °C. At 5 d intervals the supernatant was drawn and replaced with a fresh solution. The reaction was monitored with MALDI-MS.[15]

Acknowledgements

We thank Dr. C. Deck and Dr. E. Kervio for discussions. A grant from DFG (No. RI 1063/8-2 to C.R.) and the EU COST action CM1304 supported this work.

Keywords: genetic copying · nucleotides · oligomerization · replication · RNA

How to cite: *Angew. Chem. Int. Ed.* **2015**, *54*, 14559–14563
Angew. Chem. **2015**, *127*, 14767–14771

[1] M. M. Yusupov, G. Z. Yusupova, A. Baucom, K. Lieberman, T. N. Earnest, J. H. Cate, H. F. Noller, *Science* **2001**, *292*, 883–896.
[2] T. R. Cech, *Angew. Chem. Int. Ed. Engl.* **1990**, *29*, 759–768; *Angew. Chem.* **1990**, *102*, 745–755.
[3] C. C. Mello, *Angew. Chem. Int. Ed.* **2007**, *46*, 6985–6994; *Angew. Chem.* **2007**, *119*, 7114–7124.
[4] W. Gilbert, *Nature* **1986**, *319*, 618.
[5] F. Westheimer, *Science* **1987**, *235*, 1173–1178.
[6] M. S. Verlander, R. Lohrmann, L. E. Orgel, *J. Mol. Evol.* **1973**, *2*, 303–316.
[7] J. P. Ferris, A. R. Hill, R. H. Liu, L. E. Orgel, *Nature* **1996**, *381*, 59–61.
[8] S. Pino, F. Ciciriello, G. Costanzo, E. Di Mauro, *J. Biol. Chem.* **2008**, *283*, 36494–36503.
[9] G. Costanzo, R. Saladino, G. Botta, A. Giorgi, A. Scipioni, S. Pino, E. Di Mauro, *ChemBioChem* **2012**, *13*, 999–1008.
[10] P.-A. Monnard, J. W. Szostak, *J. Inorg. Biochem.* **2008**, *102*, 1104–1111.
[11] P.-A. Monnard, A. Kanavarioti, D. W. Deamer, *J. Am. Chem. Soc.* **2003**, *125*, 13734–13740.
[12] B. J. Weimann, R. Lohrmann, L. E. Orgel, H. Schneider-Bernloehr, J. E. Sulston, *Science* **1968**, *161*, 387.
[13] G. F. Joyce, T. Inoue, L. E. Orgel, *J. Mol. Biol.* **1984**, *176*, 279–306.
[14] S. R. Vogel, C. Deck, C. Richert, *Chem. Commun.* **2005**, 4922.
[15] C. Deck, M. Jauker, C. Richert, *Nat. Chem.* **2011**, *3*, 603–608.
[16] P. Hagenbuch, E. Kervio, A. Hochgesand, U. Plutowski, C. Richert, *Angew. Chem. Int. Ed.* **2005**, *44*, 6588–6592; *Angew. Chem.* **2005**, *117*, 6746–6750.
[17] S. R. Vogel, C. Richert, *Chem. Commun.* **2007**, 1896.
[18] T. Mukaiyama, M. Hashimoto, *J. Am. Chem. Soc.* **1972**, *94*, 8528–8532.
[19] B. C. Chu, L. E. Orgel, *Nucleic Acids Res.* **1988**, *16*, 3671–3691.
[20] L. J. Leman, L. E. Orgel, M. R. Ghadiri, *J. Am. Chem. Soc.* **2006**, *128*, 20–21.
[21] S. A. Benner, H.-J. Kim, M. A. Carrigan, *Acc. Chem. Res.* **2012**, *45*, 2025–2034.

[22] A. Schimpl, R. M. Lemmon, M. Calvin, *Science* **1965**, *147*, 149 – 150.
[23] F. Duvernay, T. Chiavassa, F. Borget, J.-P. Aycard, *J. Am. Chem. Soc.* **2004**, *126*, 7772 – 7773.
[24] M. L. Kilpatrick, *J. Am. Chem. Soc.* **1947**, *69*, 40 – 46.
[25] G. von Kiedrowski, *Angew. Chem. Int. Ed. Engl.* **1986**, *25*, 932 – 935; *Angew. Chem.* **1986**, *98*, 932 – 934.
[26] W. S. Zielinski, L. E. Orgel, *Nucleic Acids Res.* **1987**, *15*, 1699 – 1715.
[27] W. S. Zielinski, L. E. Orgel, *Nature* **1987**, *327*, 346 – 347.
[28] J. A. Rojas Stütz, E. Kervio, C. Deck, C. Richert, *Chem. Biodiversity* **2007**, *4*, 784 – 802.
[29] B. T. Burcar, M. Jawed, H. Shah, L. B. McGown, *Origins Life Evol. Biospheres* **2015**, *45*, 31 – 40.
[30] J. E. Sulston, R. Lohrmann, L. E. Orgel, H. T. Miles, *Proc. Natl. Acad. Sci. USA* **1968**, *59*, 726 – 733.
[31] J. Sulston, R. Lohrmann, L. E. Orgel, H. T. Miles, *Proc. Natl. Acad. Sci. USA* **1968**, *60*, 409 – 415.
[32] E. Kervio, B. Claasen, U. E. Steiner, C. Richert, *Nucleic Acids Res.* **2014**, *42*, 7409 – 7420.
[33] M. Jauker, H. Griesser, C. Richert, *Angew. Chem. Int. Ed.* **2015**, DOI: 10.1002/anie.201506593; *Angew. Chem.* **2015**, DOI: 10.1002/ange.201506593.
[34] A. Kaiser, C. Richert, *J. Org. Chem.* **2013**, *78*, 793 – 799.
[35] D. Sarracino, C. Richert, *Bioorg. Med. Chem. Lett.* **1996**, *6*, 2543 – 2548.
[36] R. Lohrmann, L. E. Orgel, *J. Mol. Biol.* **1980**, *142*, 555 – 567.
[37] K. J. Prabahar, J. P. Ferris, *J. Am. Chem. Soc.* **1997**, *119*, 4330 – 4337.
[38] A. Kaiser, S. Spies, T. Lommel, C. Richert, *Angew. Chem. Int. Ed.* **2012**, *51*, 8299 – 8303; *Angew. Chem.* **2012**, *124*, 8424 – 8428.
[39] K. Adamala, J. W. Szostak, *Science* **2013**, *342*, 1098 – 1100.
[40] A. R. Hill, L. E. Orgel, T. Wu, *Origins Life Evol. Biospheres* **1993**, *23*, 285 – 290.
[41] E. Kervio, A. Hochgesand, U. E. Steiner, C. Richert, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 12074 – 12079.

Received: July 16, 2015
Revised: September 3, 2015
Published online: October 5, 2015

