Prebiotic Chemistry

International Edition: DOI: 10.1002/anie.201610650
German Edition: DOI: 10.1002/ange.201610650

Ribonucleotides and RNA Promote Peptide Chain Growth

*Helmut Griesser, Peter Tremmel, Eric Kervio, Camilla Pfeffer, Ulrich E. Steiner, and Clemens Richert**

Abstract: *All known forms of life use RNA-mediated polypeptide synthesis to produce the proteins encoded in their genes. Because the principal parts of the translational machinery consist of RNA, it is likely that peptide synthesis was achieved early in the prebiotic evolution of an RNA-dominated molecular world. How RNA attracted amino acids and then induced peptide formation in the absence of enzymes has been unclear. Herein, we show that covalent capture of an amino acid as a phosphoramidate favors peptide formation. Peptide coupling is a robust process that occurs with different condensation agents. Kinetics show that covalent capture can accelerate chain growth over oligomerization of the free amino acid by at least one order of magnitude, so that there is no need for enzymatic catalysis for peptide synthesis to begin. Peptide chain growth was also observed on phosphate-terminated RNA strands. Peptide coupling promoted by ribonucleotides or ribonucleotide residues may have been an important transitional form of peptide synthesis that set in when amino acids were first captured by RNA.*

All known forms of life produce proteins through RNA-directed formation of polypeptides, catalyzed by the ribosome, the catalytic core of which consists of RNA.[1] It is clear that a molecular machinery of the complexity of the ribosomal apparatus did not arise in one step, starting from random sequences.[2] There must have been a simpler form of RNA-induced peptide synthesis that preceded translation in prebiotic evolution. This simpler system cannot have depended on protein enzymes encoded in genes, such as the known aminoacyl tRNA synthetases,[3] because genes could not yet be translated. Something as simple as intrinsic reactivity and/or molecular recognition may have driven the formation the first peptides produced in RNA-promoted reactions.

Numerous hypotheses on the origin of the translational system have been formulated,[4,5] and there are famous proposals on the even more difficult problem of the origin of the genetic code,[6–8] but an experimental system that starts from free amino acids and ribonucleotides and that shows peptide growth in the absence of enzymes has been lacking. As a consequence, the questions of how amino acids were first recruited by the "molecular agents" of the "RNA world"[9] and what drove peptide formation have remained open, although oligomerization pathways producing depsipeptides through wet–cool/dry–hot cycles have been found.[10] Interest in this question is increasing, as amino acids are being found in ever more remote places, such as the coma of comets,[11] and new pathways for the enzyme-free formation of ribonucleotides are being described.[12,13] So it is important to know what reactions may have favored a "peptide–RNA world".[14]

Covalent attachment of a peptide to an RNA strand prevents the loss of the precious oligomer through diffusion and favors complex formation. Therefore, covalently linked ribonucleopeptides could have provided evolutionary benefits.[15,16] Even today, growing polypeptides are covalently linked to RNA during translation as peptidyl tRNAs. Ideally, an experimental set-up used to search for primordial peptidyl RNAs should run without the need for separate chemical steps, such as the ex situ activation of nucleotides or amino acids,[17,18] or the esterification of tRNA model compounds.[19,20] It should start from free biomolecules and should run without human intervention. Furthermore, the experimental system should have the ability to encode and transmit genetic information, so that advances in the capabilities of the system[21] can be preserved.

We have recently reported an experimental system in which peptides that are covalently linked to oligoribonucleotides form from free amino acids and ribonucleotides (Figure 1).[22] The peptides are *N*-linked, which is why we now prefer to call them "peptido RNAs". The peptido RNAs were found under conditions that induce the de novo formation of RNA chains, as well as genetic copying of

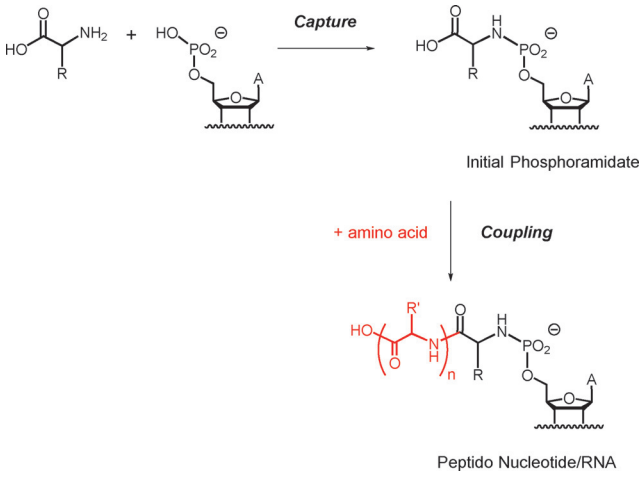

Figure 1. RNA-primed growth of peptide chains involves covalent capture and subsequent amide coupling.

[*] Dipl.-Chem. (FH) H. Griesser, Dr. P. Tremmel, Dr. E. Kervio, C. Pfeffer, Prof. Dr. C. Richert
Institut für Organische Chemie, Universität Stuttgart
70569 Stuttgart (Germany)
E-mail: lehrstuhl-2@oc.uni-stuttgart.de
Prof. Dr. U. E. Steiner
Department of Chemistry, University of Konstanz
78457 Konstanz (Germany)

Supporting information for this article can be found under: http://dx.doi.org/10.1002/anie.201610650.

existing RNA sequences, that is, in a system capable of encoding genetic information.[23] The salt-rich aqueous condensation buffer that fosters these processes contains a water-soluble carbodiimide as the chemical driving force and 1-ethylimidazole as an organocatalyst. Alkylated adenine also shows organocatalytic activity.[23] The spontaneous condensation processes observed in the ice-cold aqueous solution include the formation of cofactors, such as NAD^+, FAD, and ATP, and control experiments showed that the system tolerates phospholipids. It was unclear, however, whether ribonucleotide-linked peptide chain growth was faster than oligomerization of amino acids alone. Herein, we show that the capture of an amino acid by the phosphate of a ribonucleotide accelerates peptide formation over the ribonucleotide-free oligomerization of the amino acid. Further, we demonstrate that peptido RNA also grows on RNA strands, both by itself, and simultaneously with RNA chain growth at the 3′-terminus. This shows that there is a chemistry that links amino acids and RNA that could have favored RNA-promoted peptide coupling long before aminoacyl tRNAs synthetases began to charge tRNAs.

The reaction sequence shown in Figure 1 involves two distinct phosphoramidate species. The first is the product of amino acid capture, the second is a peptido RNA that results from coupling of the first. We had previously found that covalent capture occurs both in solution containing *N*-ethyl-*N*′-(3-dimethylaminopropyl) carbodiimide (EDC) or cyanamide, as well as with ribonucleotides pre-activated as OAt esters.[22] It was unclear, however, whether coupling would occur with an activating agent other than EDC. What activation agents are prebiotically most plausible is an open issue that will not be easy to settle, as it will be extremely difficult to find fossils that answer this question conclusively.

We therefore studied the reaction of glycine with adenosine 5′-monophosphate (AMP) in aqueous buffer containing cyanamide[24] or carbonyl diimidazole (CDI), a derivative of carbonic acid, as activation agent. In either case, both capture and coupling were detected by ^{31}P-NMR at near-neutral pH (Figure 2). Additional data and mass spectra are shown in the Supporting Information. The reactivity is much lower for cyanamide, so that the assay was run for more than two months at 20 °C, and for CDI, the condensation agent was quickly used up, but all three of the condensation agents lead to formation of covalently linked peptide from the free amino acid with a covalent capture product as intermediate. Because EDC contains the cyanamide locked in the productive carbodiimide tautomer, so that assays can be performed on a reasonable time scale, the subsequent experiments were performed with this activation agent. We have no reason to believe that the ethyl and dimethylaminopropyl groups of EDC guide the reaction pathway of amino acids.

Next, we studied whether the covalent capture promotes the formation of peptide bonds over the background reaction. For this, we first monitored the consumption of glycine in condensation buffer in the absence of a ribonucleotide by ^{1}H-NMR (Figure 3). After one day of slow reactions of glycine, when approximately 25 % of the EDC had been consumed, we supplemented the solution with a substoichiometric amount of the phosphoramidate Gly-AMP, that is, the

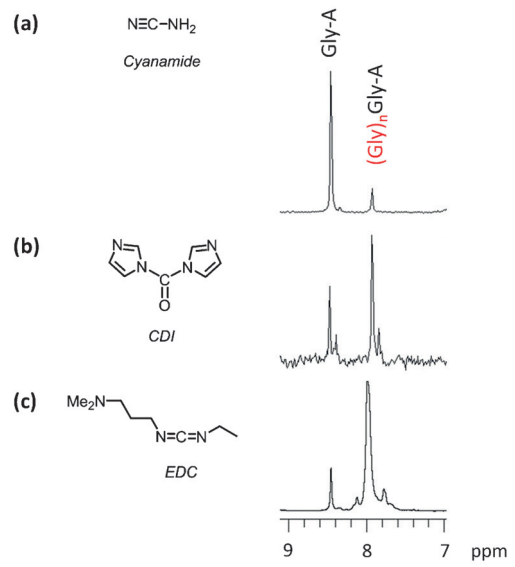

Figure 2. Capture and coupling occur for different activation agents, as determined by ^{31}P-NMR. Structures of activation agents and ^{31}P-NMR spectra of reaction solutions with glycine (0.2–1 M) and AMP (0.2–0.5 M) in buffer (0.5 M HEPES, 0.08 M $MgCl_2$, and 0.15 M 1-ethylimidazole). a) With cyanamide (2 M) after 75 d at 20 °C, pH 6.5; b) with CDI (1 M) after 10 d at 20 °C, pH 7.5; or c) with EDC (0.8 M) as activation agent after 7 d at 0 °C and pH 7.5. See Table S1 in the Supporting Information for further details.

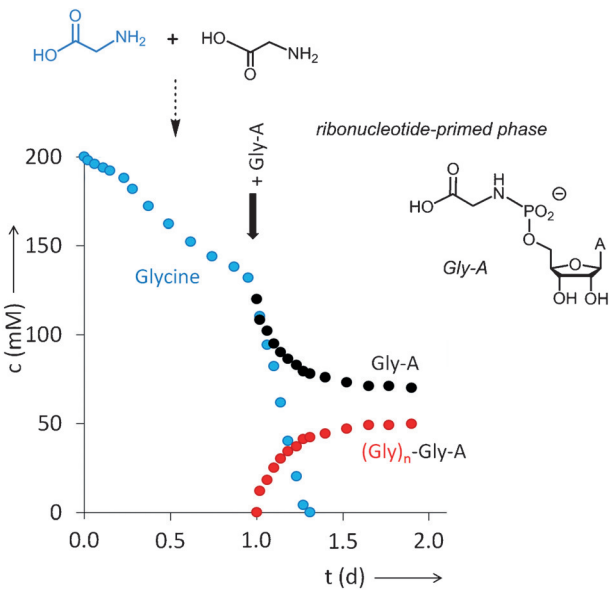

Figure 3. Priming by ribonucleotide capture accelerates peptide coupling. Glycine (0.2 M) was allowed to react in condensation buffer (0.5 M HEPES, 80 mM $MgCl_2$, 0.15 M 1-ethylimidazole, and 0.8 M EDC) at 0 °C, and after 1 d, Gly-AMP (0.12 M) was added, leading to accelerated consumption of the amino acid, as determined by ^{1}H NMR. See the Supporting Information for details.

product of amino acid capture. Now, a strong acceleration was observed, and the remaining free glycine was consumed within eight hours. As the glycine pool was depleted, the Gly-AMP signal leveled off. The more rapid decay of the glycine signal was indicative of chain growth, with several glycine

residues being attached to one initial phosphoramidate. Such chain growth is a multistep process with complex kinetics, but simple monoexponential fits indicate an approximately 10-fold faster consumption of glycine in the second phase of the kinetic experiment.[25] Inorganic phosphate does not have the same effect. It largely precipitates from condensation buffer, and when the Mg^{2+} concentration is lowered to 1 mM to prevent this, little or no peptide formation is detectable, while AMP still gives robust chain growth (Supporting Information).

We wished to better understand ribonucleotide priming. Initially, we assumed that, like acylation,[26] the phosphorylation of the amino group prevents diketopiperazine formation. It also affects the reactivity of the carboxylic acid moiety, possibly improving activation and coupling. Because it is not easy to monitor peptide chain growth quantitatively in condensation buffer by ^{1}H NMR due to signal overlap, we developed a methodology for the ^{13}C-NMR channel. This methodology uses glycine that is ^{13}C-isotope-labeled at position 1 for much-improved signal, and Gd-DTPA (a compound used clinically in MRI) as a water-soluble relaxation enhancer to obtain quantifiable signals. Control experiments with the latter added either at the beginning or at the end of the assays confirmed that the stable paramagnetic complex does not affect product distribution detectably. That near-equal signal intensities for carbon nuclei in different relaxation environments result was confirmed by measuring spectra for histidine, an amino acid similar in size to glycylglycine and containing a variety of different carbon centers. When inverse-gated ^{13}C-spectra were run with a recycle delay of 2 s at low millimolar Gd-DTPA concentration, all of the signal intensities of histidine were within 2% of each other (Figures S5 and S6).

With the ^{13}C-based methodology, we were now in a position to track the fate of glycine in the absence and presence of AMP. For glycine alone, 65% of the free amino acid remained after 24 h at 0°C (Figure 4). No diketopiperazine was detectable,[27] as also confirmed by spiking with authentic material. Instead, guanidine formation (Gly-EDC, 18% of glycine) and formation of the two different *N*-acylureas (Gly-EDU, 8% after 1 d) were largely responsible for glycine consumption. At best, 9% of the glycine carbonyl signals after 1 d were due to peptide coupling products. It is likely that what little peptide chain growth occurred is primed by guanidine formation. In stark contrast, in the assay solution containing AMP, almost all of the glycine was consumed after 24 h, with only 4% free Gly left. Besides capture product Gly-A (13%) and some guanidine (9%), peptido nucleotides with chain lengths of up to five residues were identified as products of glycine consumption (70%).

We then asked whether peptide chain growth was limited to ribonucleotide-based forms, or could also be primed by 5′-terminal phosphates of RNA strands. For this, we performed assays with the RNA primer 5′-pGUAUGCUGG-3′, featuring a 5′-phosphate, as expected for oligomerization products of NMPs bound to a longer RNA strand. Under all of the conditions tested thus far, peptide chain growth was detectable. Figure 5 shows MALDI mass spectra that were acquired under conditions that allow for quantitatively monitoring

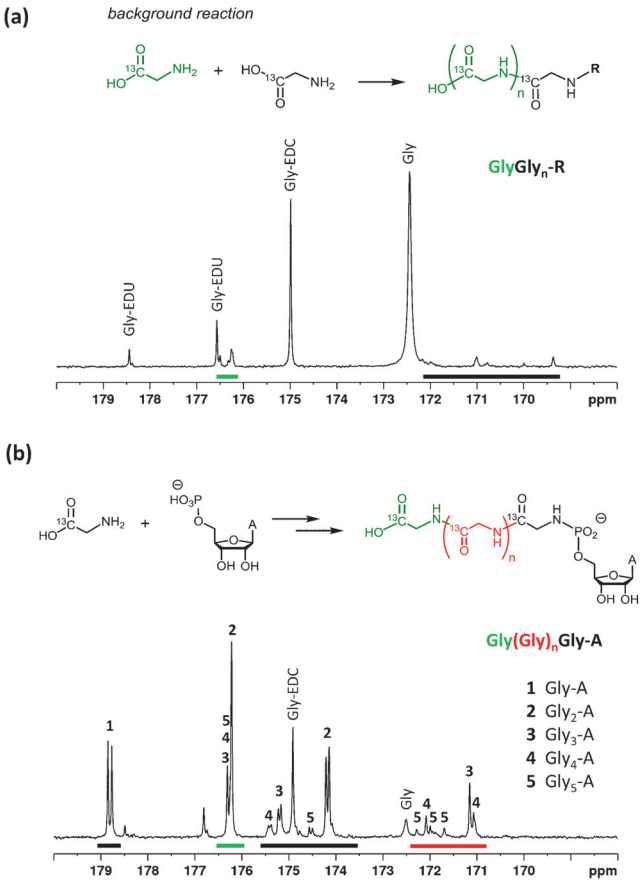

Figure 4. Peptide formation is more efficient in the ribonucleotide-captured form, as detected by ^{13}C NMR of assays involving C1 ^{13}C-labeled glycine (0.2 M) in condensation buffer, a) without and b) with AMP (0.2 M). Shown are ^{13}C-NMR spectra (100 MHz, carbonyl region) of assay solutions after 24 h at 0°C at 1.8 mM Gd-DTPA as relaxation enhancer and inverse gated proton decoupling to minimize the NOE effect of conventional decoupling. In (a), both the signals for remaining glycine (Gly) and the products of reactions with EDC (guanidium adduct Gly-EDC and the two different N-acylureas (Gly-EDU) are labeled. Other peak assignments, based on two-dimensional spectra, are given as color bars below the spectra, with green for C-terminal carboxylates, red for internal amide carbons, and black for amide carbons of glycine residues directly linked to AMP. The latter appear as doublets, because of coupling to the 5′-phosphorous nucleus. See Figure S7 for an earlier time point.

primer extension.[28,29] For 0.1 M glycine in full-strength condensation buffer, peaks for peptido RNAs with up to six glycine residues were readily discernible after 3 d (Figure 5a). For the side-chain-bearing amino acids serine and threonine, we increased the amino acid concentration and lowered the EDC concentration to 0.4 M. Again, significant chain growth was detected (Figure 5b,c).

Next, we performed an assay with both an amino acid and the ribonucleotide CMP to test for the compatibility of peptide synthesis and RNA copying. Indeed, the hoped-for simultaneous extension of the primer at both of its termini was observed (Figure 5d). Extension at the 3′-terminus by a nucleotide and peptide production at the 5′-terminus (as peptido chains) manifested themselves in the peak pattern of the mass spectrum acquired after 2 d. Apparently, peptido

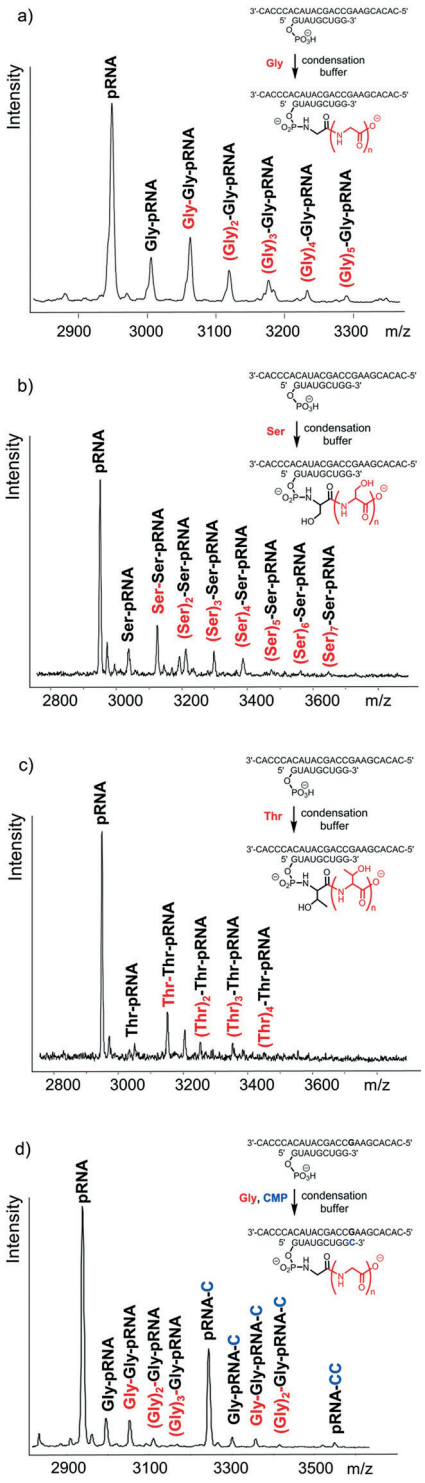

Figure 5. Peptide chain growth on RNA strands: reaction schemes and MALDI-TOF mass spectra showing formation of a peptido chain at the 5′-terminus of the 5′-phosphate-terminated RNA primer (50 μM) in duplex with a longer RNA strand in condensation buffer (0.5 M HEPES, 0.08 M $MgCl_2$, 0.15 M 1-ethylimidazole at pH 7.5 and 0 °C). Assay with a) glycine (0.1 M) and EDC (0.8 M) after 3 d; b) 0.4 M serine and 0.4 M EDC after 9 d; c) 0.4 M threonine and 0.4 M EDC after 9 d; and d) both glycine (0.1 M) and cytidine 5′-monophosphate (CMP, 40 mM) at 0.8 M EDC after 2 d. See Table S2 for a list of peak intensities.

RNA formation can occur under conditions that allow for the emergence of more complex RNA machineries through formation of genetic material and genetic copying of this material.

The peak patterns in Figure 5 show less capture product containing the first amino acid than either the remaining primer or the dipeptide-containing product. This confirms that once the first amino acid is captured, peptide coupling is rapid, so that the reactive initial phosphoramidate is depleted compared to the other species. Subsequent chain growth appears to produce a peak pattern expected for similar reaction rates. No run-away reaction rapidly exhausts the condensation agent, preventing the formation of other products, such as copies of an RNA template or cofactors. Such well-behaved, mutually compatible or synergistic processes are characteristic of living systems.

In conclusion, our results show that the formation of peptide chains that are covalently linked to ribonucleotides or RNA via amino acid capture and peptide coupling is a robust process that occurs with different activation agents. The covalent capture induces a rate acceleration that makes the difference between unsuccessful and successful peptide synthesis. Both ribonucleotides and the terminal phosphates of RNA chains induce the effect. Taken together, this suggests that RNA had the ability to capture and couple amino acids long before either ribozymes or enzymes were formed. The rich structure space of peptido RNAs forming with other proteinogenic amino acids and ribonucleotides is reported in the accompanying paper.[30]

Acknowledgements

The authors thank Maren Bechthold, Sebastian Motsch, and Marilyne Sosson for sharing results, Drs. Birgit Claasen and Delphine Garnier for NMR spectra, and Tara Krahl for expert technical assistance. Grants by DFG (RI 1063/8-2 and RI 1063/16-1) and EU COST action CM1304 support this research.

Conflict of interest

The authors declare no conflict of interest.

Keywords: amino acids · nucleotides · peptides · prebiotic chemistry · RNA

How to cite: *Angew. Chem. Int. Ed.* **2017**, *56*, 1219–1223
Angew. Chem. **2017**, *129*, 1239–1243

[1] H. F. Noller, *Science* **2005**, *309*, 1508–1514.
[2] F. H. C. Crick, S. Brenner, A. Klug, G. Pieczenik, *Origins Life Evol. Biospheres* **1976**, *7*, 389–397.
[3] M. Delarue, *Curr. Opin. Struct. Biol.* **1995**, *5*, 48–55.
[4] For a recent review, see: D. W. Morgens, *J. Mol. Evol.* **2013**, *77*, 185–196.
[5] L. E. Orgel, *J. Mol. Evol.* **1989**, *29*, 465–474.
[6] M. Yarus, J. J. Widmann, R. Knight, *J. Mol. Evol.* **2009**, *69*, 406–429.

[7] P. Schimmel, B. Henderson, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 11283 – 11286.
[8] J. D. Sutherland, J. M. Blackburn, *Chem. Biol.* **1997**, *4*, 481 – 488.
[9] W. Gilbert, *Nature* **1986**, *319*, 618.
[10] J. G. Forsythe, S.-S. Yu, I. Mamajanov, M. A. Grover, R. Krishnamurthy, F. M. Fernandez, N. V. Hud, *Angew. Chem. Int. Ed.* **2015**, *54*, 9871 – 9875; *Angew. Chem.* **2015**, *127*, 10009 – 10013.
[11] K. Altwegg et al., *Sci. Adv.* **2016**, *2*, e1600285.
[12] M. W. Powner, B. Gerland, J. D. Sutherland, *Nature* **2009**, *459*, 239 – 242.
[13] S. Becker, I. Thoma, A. Deutsch, T. Gehrke, P. Mayer, H. Zipse, T. Carell, *Science* **2016**, *352*, 833 – 836.
[14] R. Pascal, L. Boiteau, A. Commeyras, *Top. Curr. Chem.* **2005**, *259*, 69 – 122.
[15] M. Di Giulio, *J. Mol. Evol.* **1997**, *45*, 571 – 578.
[16] T. Cavalier-Smith, *J. Mol. Evol.* **2001**, *53*, 555 – 595.
[17] A. L. Weber, L. E. Orgel, *J. Mol. Evol.* **1978**, *11*, 9 – 16.
[18] J.-P. Biron, A. L. Parkes, R. Pascal, J. D. Sutherland, *Angew. Chem. Int. Ed.* **2005**, *44*, 6731 – 6734; *Angew. Chem.* **2005**, *117*, 6889 – 6892.
[19] A. L. Weber, L. E. Orgel, *J. Mol. Evol.* **1978**, *11*, 189 – 198.
[20] R. M. Turk, M. Illangasekare, M. Yarus, *J. Am. Chem. Soc.* **2011**, *133*, 6044 – 6050.
[21] C. de Duve, *Nature* **2005**, *433*, 581 – 582.
[22] M. Jauker, H. Griesser, C. Richert, *Angew. Chem. Int. Ed.* **2015**, *54*, 14564 – 14569; *Angew. Chem.* **2015**, *127*, 14772 – 14777.
[23] M. Jauker, H. Griesser, C. Richert, *Angew. Chem. Int. Ed.* **2015**, *54*, 14559 – 14563; *Angew. Chem.* **2015**, *127*, 14767 – 14771.
[24] A. Schimpl, R. M. Lemmon, M. Calvin, *Science* **1965**, *147*, 149 – 150.
[25] Little free peptide forms when glycine oligomerizes (Figure 4), so that the true rate acceleration is at least 70-fold.
[26] J.-C. Cavadore, A. Previero, *Bull. Soc. Chim. Biol.* **1969**, *51*, 1245 – 1253.
[27] D. Beaufils, S. Jepaul, Z. Liu, L. Boiteau, R. Pascal, *Origins Life Evol. Biospheres* **2016**, *46*, 19 – 30.
[28] E. Kervio, A. Hochgesand, U. Steiner, C. Richert, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 12074 – 12079.
[29] E. Kervio, M. Sosson, C. Richert, *Nucleic Acids Res.* **2016**, *44*, 5504 – 5514.
[30] H. Griesser, M. Bechthold, P. Tremmel, E. Kervio, C. Richert, *Angew. Chem. Int. Ed.* DOI: 10.1002/anie.201610651; *Angew. Chem.* DOI: 10.1002/ange.201610651.

Manuscript received: October 31, 2016
Revised: November 30, 2016
Final Article published: December 21, 2016

