ChemPubSoc Europe

DOI: 10.1002/cbic.201700122

CHEMBIOCHEM Full Papers

# Nonenzymatic Oligomerization of 3′,5′-Cyclic CMP Induced by Proton and UV Irradiation Hints at a Nonfastidious Origin of RNA

Giovanna Costanzo,[a] Alessandra Giorgi,[b] Anita Scipioni,[c] Anna Maria Timperio,[d] Carmine Mancone,[e, f] Marco Tripodi,[e, f] Michail Kapralov,[g] Eugene Krasavin,[g] Holger Kruse,[h] Jiří Šponer,[h, i] Judit E. Šponer,*[h] Václav Ranc,[i] Michal Otyepka,[i] Samanta Pino,[j] and Ernesto Di Mauro*[j]

We report that 3′,5′-cyclic CMP undergoes nonenzymatic di- and trimerization at 20 °C under dry conditions upon proton or UV irradiation. The reaction involves stacking of the cyclic monomers and subsequent polymerization through serial transphosphorylations between the stacked monomers. Proton- and UV-induced oligomerization of 3′,5′-cyclic CMP demonstrates that pyrimidines—similar to purines—might also have taken part in the spontaneous generation of RNA under plausible prebiotic conditions as well as in an extraterrestrial context. The observed polymerization of naturally occurring 3′,5′-cyclic nucleotides supports the possibility that the extant genetic nucleic acids might have originated by way of a straight Occamian path, starting from simple reactions between plausibly preactivated monomers.

[a] Dr. G. Costanzo
Istituto di Biologia e Patologia Molecolari, CNR
Piazzale Aldo Moro 5, 00185 Rome (Italy)

[b] Dr. A. Giorgi
Dipartimento di Scienze Biochimiche, "Sapienza" Università di Roma
Piazzale Aldo Moro 5, 00185 Rome (Italy)

[c] Prof. A. Scipioni
Dipartimento di Chimica, "Sapienza" Università di Roma
Piazzale Aldo Moro 5, 00185 Rome (Italy)

[d] Prof. A. M. Timperio
Department of Ecology and Biology, "La Tuscia" University
Viale dell'Università snc, 01100 Viterbo (Italy)

[e] Dr. C. Mancone, Prof. M. Tripodi
Department of Cellular Biotechnologies and Hematology
Sapienza University of Rome
Via Regina Elena 324, 00161 Roma (Italy)

[f] Dr. C. Mancone, Prof. M. Tripodi
National Institute for Infectious Diseases "L. Spallanzani", IRCCS
Via Portuense 292, 00149 Rome (Italy)

[g] Dr. M. Kapralov, Prof. E. Krasavin
Joint Institute for Nuclear Research, Laboratory of Radiation Biology
141980 Dubna (Russia)

[h] Dr. H. Kruse, Prof. J. Šponer, Dr. J. E. Šponer
Institute of Biophysics, Academy of Sciences of the Czech Republic
Královopolská 135, 61265 Brno (Czech Republic)
E-mail: judit@ncbr.muni.cz

[i] Prof. J. Šponer, Dr. V. Ranc, Prof. M. Otyepka
Regional Centre of Advanced Technologies and Materials
Department of Physical Chemistry
Faculty of Science, Palacky University
17. Listopadu, 771 46 Olomouc (Czech Republic)

[j] Dr. S. Pino, Prof. E. Di Mauro
Department of Biology and Biotechnology
Laboratory affiliated to Istituto Pasteur Italia-Fondazione Cenci Bolognetti
Sapienza University of Rome
Piazzale Aldo Moro 5, 00185 Rome (Italy)
E-mail: ernesto.dimauro@uniroma1.it

Supporting information and ORCID identification numbers for the authors of this article can be found under https://doi.org/10.1002/cbic.201700122.

## Introduction

Extant biological DNA and RNA syntheses are based on template copying by highly evolved polymerases. High levels of adaptation and refinement have been achieved in studies on the in vitro evolution of polymerizing enzymes,[1–3] and it has been shown that genetic information can be stored in and recovered from alternative genetic polymers based on nucleic acid architectures not found in nature.[4]

High-yielding analytically powerful nonenzymatic polymerization methodologies have also been developed,[5–13] enabling in vitro studies on molecular evolution and supporting the Darwinian logics pioneered by Spiegelman.[14] Complex chemistries involved in the nonenzymatic polymerization of high-energy monomers to nucleic acids are well described in the literature.[15] But: where is the starting point of all this?

The formation of phosphodiester bonds in extant polymers is based on the nucleophilic attack of a sugar hydroxy group on the phosphorus atom of an activated nucleotide. The efficiency of the polymerization reaction is determined by the extent of phosphate activation and the properties of the leaving group. High-yielding template-directed RNA syntheses are commonly performed with 5′-phosphorimidazolide nucleosides[9–16] or one of the numerous variants thereof,[17,18] because imidazole and its derivatives are good leaving groups. These are all high-energy compounds that polymerize efficiently, but the likelihood of their prebiotic availability and accumulation is inversely proportional to their intrinsic stability and the elaborate nature of the chemistry necessary for their synthesis. Whereas high energy entails instability, the elaborate chemistry reduces the likelihood of formation. In keeping with these arguments, evidence for the presence of phosphorimidazolides or even triphosphate nucleotides in prebiotic scenarios is lack-

ing.[19,20] In brief, the prebiotic ur-generation of RNA remains undeciphered.

Nucleoside 2′,3′- and 3′,5′-cyclic phosphates were explored four decades ago as potential substrates for RNA polymerization. Attention was mostly devoted to the 2′,3′-cyclic form[21] both for direct polymerization and for the formation of phosphodiester bonds connecting preformed oligonucleotides.[22] These studies were not developed further, and 3′,5′-cyclic purine nucleotides were analyzed only in two connected studies[23,24] that indicated that 3′,5′-cyclic GMP and, to a lesser extent, 3′,5′-cyclic AMP[23] afford short oligomers in water. More recent studies[25,26,27] in this area are discussed below.

3′,5′-Cyclic nucleotides can form under plausible prebiotic conditions,[28,29] as reviewed and discussed in ref. [30]. Heating nucleosides in formamide for circa $10^2$ hours in the presence of phosphates (i.e., hydroxylapatite or $KH_2PO_4$) causes phosphorylation at every possible position in the ribose (2′, 3′, or 5′), followed by cyclization on prolonged treatment. These reactions occur under the same physical-chemical conditions under which nucleobases are abiotically generated,[31–33] thus suggesting that 3′,5′-cyclic nucleotides might have been present in the prebiotic mix.[34] A reaction in which the cleavage of the internal phosphodiester bond of a cyclic nucleotide was coupled with the formation of a new phosphodiester linkage to yield a linear polymer would be thermodynamically favored: the free energy change of phosphodiester bond formation has been calculated to be $-5.5\ kcal\,mol^{-1}$,[35] whereas the enthalpies of hydrolysis of various 3′,5′- and 2′,3′-cyclic nucleotides range from $-7.7$ to $-14.1\ kcal\,mol^{-1}$.[36] Consequently, if the substrate molecules are appropriately positioned, their polymerization is not thermodynamically hampered.

This manuscript describes experiments conducted to investigate the conditions that permit di- and trimerization of 3′,5′-cyclic CMP (hereafter abbreviated as 3′,5′-cCMP) in the absence of enzymes and templates. These conditions are compared with those required for the oligomerization of 3′,5′-cyclic GMP (hereafter abbreviated 3′,5′-cGMP).

## Results and Discussion

The polymerization of 3′,5′-cGMP at moderately high (60–80 °C) temperatures in solution (water or formamide) and under dry conditions[37] has been described in detail.[23–26,38] In a more recent study, we reported that prolonged heating of dry 3′,5′-cAMP also yields short oligonucleotide sequences.[27] The covalent nature of the oligomerization products has been confirmed by a variety of analysis methods, such as PAGE,[24,26] MALDI-ToF,[24,26,27] MALDI-ToF/ToF,[27] NMR,[24] and enzymatic analysis.[24]

On the basis of our previous experiences acquired when investigating oligomerization of 3′,5′-cAMP and 3′,5′-cGMP we decided to use a combination of two mass-spectrometric techniques differing in the method of ionization—that is, MALDI-ToF MS and ESI-MS—for the detection of short oligoC sequences in the variously treated 3′,5′-cCMP samples. Whereas the MALDI-ToF technique enables detection of longer oligomers, ESI is restricted to the identification of shorter ones, up to trimers (i.e., species with $m/z$ below 1000 Da). Nevertheless, as we show below, besides being quantitative, ESI detection has another clear advantage over MALDI-ToF in detecting covalent dimers and trimers: the observed signal is less disturbed by the formation of noncovalent adducts that often hampers straightforward interpretation of MALDI-ToF spectra.[39,40]

In our previous studies we have shown that 3′,5′-cyclic purine nucleotides oligomerize at 85 °C both in the dry state and in aqueous or formamide solutions.[24,26] This anionic ring-opening oligomerization is enabled by the self-assembling of the nucleotide precursors into a stacked supramolecular architecture, which ensures optimum steric conditions for the transphosphorylation reactions. Thus, formation of the phosphodiester linkages, which typically proceeds at 85 °C, is made possible by the enhanced stability of purine–purine stacks. Heat-induced oligomerizations conducted under the same conditions did not lead to detectable amount of oligomers when starting from dry samples of 3′,5′-cCMP. We assumed that this is the consequence of the obviously lower stability of pyrimidine–pyrimidine stacks, which could interfere with the persistence of stacked supramolecular architectures at elevated temperatures. Therefore, we looked for other energy sources that could be used to trigger the oligomerization at room temperature.

### Oligomerization of 3′,5′-cCMP at room temperature upon proton irradiation

Proton irradiation at 20 °C resulted in the formation of dimers and trimers.

The uppermost panel of Figure 1 shows the overall MALDI pattern of the material treated with protons. The most intense band of the spectrum is that assigned to noncovalent dimers (cC,cC) at $m/z = 611.26$, which clearly shows the strong tendency of 3′,5′-cCMPs to evaporate from the MALDI matrix as dimers. Expansion of the dimer region suggests formation of the pCpC 629 molecule ($305+305+18+1$; for an explanation see Scheme S1 in the Supporting Information), which is formed by the dimerization of 3′,5′-cCMP. Similarly, formation of covalent trimeric pCpCpC 934 species ($305+305+305+18+1$, Scheme S1) can be expected on the basis of the measured spectra. MALDI-ToF/ToF fragmentation analysis of pCpC dimers formed by proton irradiation of 3′,5′-cCMP at 20 °C showed only the open CMP form (Scheme S1) at $m/z = 324.09$, whereas the signal corresponding to the decisively important pCp fragment (Scheme S1) was completely absent. This was the reason why we also analyzed the sample by ESI-MS. Note that diphosphorylated nucleosides can form solely through fragmentation of covalent oligomers: therefore, their presence among the fragmentation products unambiguously indicates covalent bond formation.

The ESI-MS profiles of the untreated sample and of the irradiated one are shown in Figure 2.

In the overall ESI-MS profile (Figure 2) the absence of the 322 ($=304+18$) signal corresponding to cCMP + water noncovalent dimer or hydrolyzed cCMP is apparent. Further, in contrast with the MALDI-ToF spectrum, the intensity of the signal

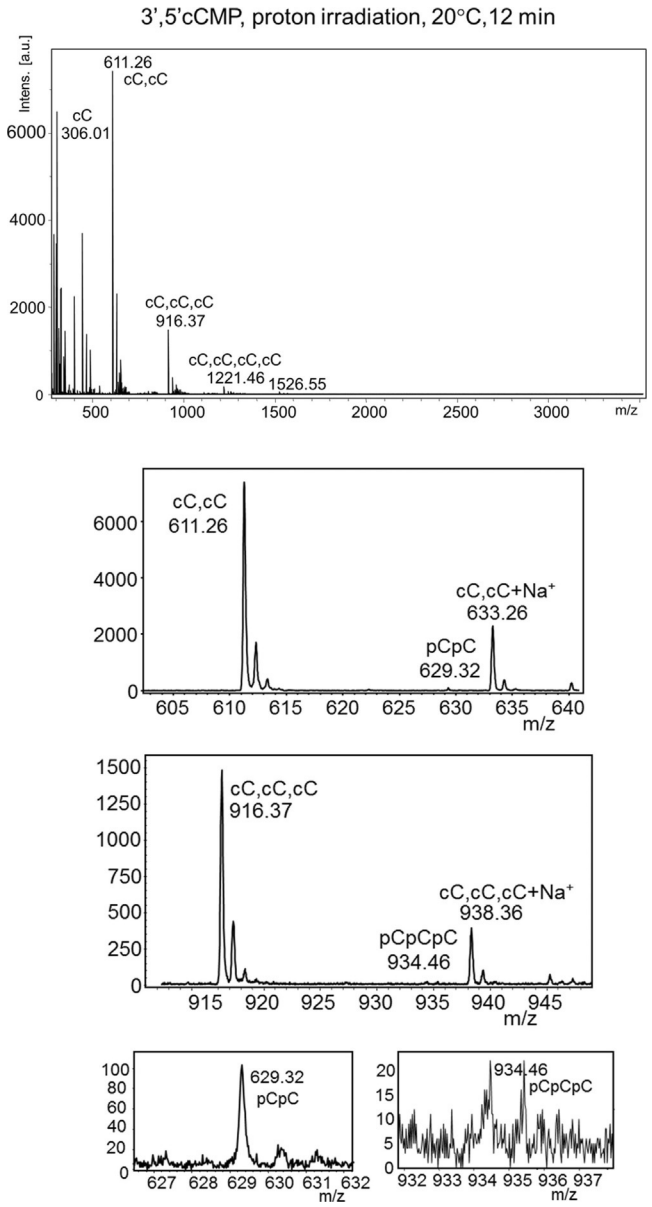



**Figure 1.** MALDI-ToF MS spectra measured for the oligomerization products of 3′,5′-cCMP (proton irradiation, 12 min, 20 °C). The uppermost spectrum shows the overall profile. Expansions of the dimer and trimer regions are presented below. Detection was made in the positive-ion mode. cC stands for 3′,5′-cCMP; for pCpC and pCpCpC, see Scheme S1.

corresponding to the noncovalent dimer at 609 is much lower than that of the dominant monomer peak at 304. This shows that, in the current case, signals detected by ESI-MS are evidently much less disturbed by formation of noncovalent adducts than those detected by MALDI-ToF. Thus, it is reasonable to assume that the signal at 627 (detected as negative ion) corresponds to covalent dimers (pCpC) rather than to a noncovalent adduct of cCMP and hydrolyzed cCMP. Indeed, fragmentation of the 629 signal (in positive mode), unlike the MALDI fragmentation, resulted both in the open CMP (324) and in pCp (i.e., CMP + 80 = 404; Figure 3 and Scheme S1); this is a clear indication that covalent dimers do form in the irradiation experiments. Let us note that the signal corresponding to covalent dimers is completely absent from the ESI spectrum of the untreated sample (Figure 2, insert).

From the ratio of the ESI signal intensities associated with the covalent dimer (pCpC) at 627 and the monomer (cCMP) at 304 the amount of pCpC formed in the proton irradiation experiment is about 2.6 %.

### Oligomerization of 3′,5′-cCMP at room temperature upon UV irradiation

Di- and trimerization was also observed upon UV irradiation of dry 3′,5′-cCMP samples. On the basis of the MALDI-ToF spectra (Figure 4), formation of dimers and trimers was expected in the UV-treated sample. The covalent nature of the dimeric product formed was confirmed by ESI, because the MALDI-ToF fragmentation—as in the case of the proton-irradiated sample—did not display the signal corresponding to the pCp fragment at *m/z* 404 (Figure 5 and Scheme S1).

Overall, these results show that cyclic pyrimidine nucleotides, similarly to their purine counterparts, also have the potential to generate short oligonucleotide sequences spontaneously as long as a suitably selected energy source is available to trigger the oligomerization.

### Interpretation of the proton irradiation experiments—the mechanism of spontaneous 3′,5′-cCMP oligomerization

It has been shown that thermally induced polymerization of 3′,5′-cGMP proceeds in a basic environment. A comprehensive theoretical model of the reaction mechanism of the spontaneous 3′,5′-cGMP oligomerization suggests that it is initiated by the nucleophilic attack of $OH^-$ ions at the phosphorus atom of 3′,5′-cGMP. The reaction then further propagates due to the self-stacking properties of 3′,5′-cGMP, which naturally provide a suitable orientation of the reactants to support the oligomerization structurally.[26]

Here we suggest that the mechanism of the 3′,5′-cCMP oligomerization might be analogous. Because the solid 3′,5′-cCMP samples used for the polymerization are never fully dehydrated, the incident protons might generate sufficient amounts of $^{\bullet}OH$ radicals, due to radiolysis of water,[41] that might initiate an analogous ring-opening oligomerization as we describe in ref. [26] for $OH^-$ ions and 3′,5′-cGMP (Scheme 1). In our previous study[26] we show that the rate-determining step of the anionic ring-opening polymerization of 3′,5′-cGMPs is the initiation step: the nucleophilic attack of $OH^-$ ions at phosphorus. The chain-extension steps require a noticeably lower activation energy. Thus, in the current work, using a simplified theoretical model representing 1′-deoxyribose 3′,5′-cyclic phosphate, we have compared the activation energy of the rate-determining initiation step for the attack of an $^{\bullet}OH$ radical and for that of an $OH^-$ ion (for methodological details see the section entitled Quantum chemical calculations in the Supporting Information).

Our computations revealed that the activation energy of the radical-induced process is 21.7 kcal mol$^{-1}$ (B3LYP/6–311 + + G** level) whereas that of the anionic process is markedly higher,

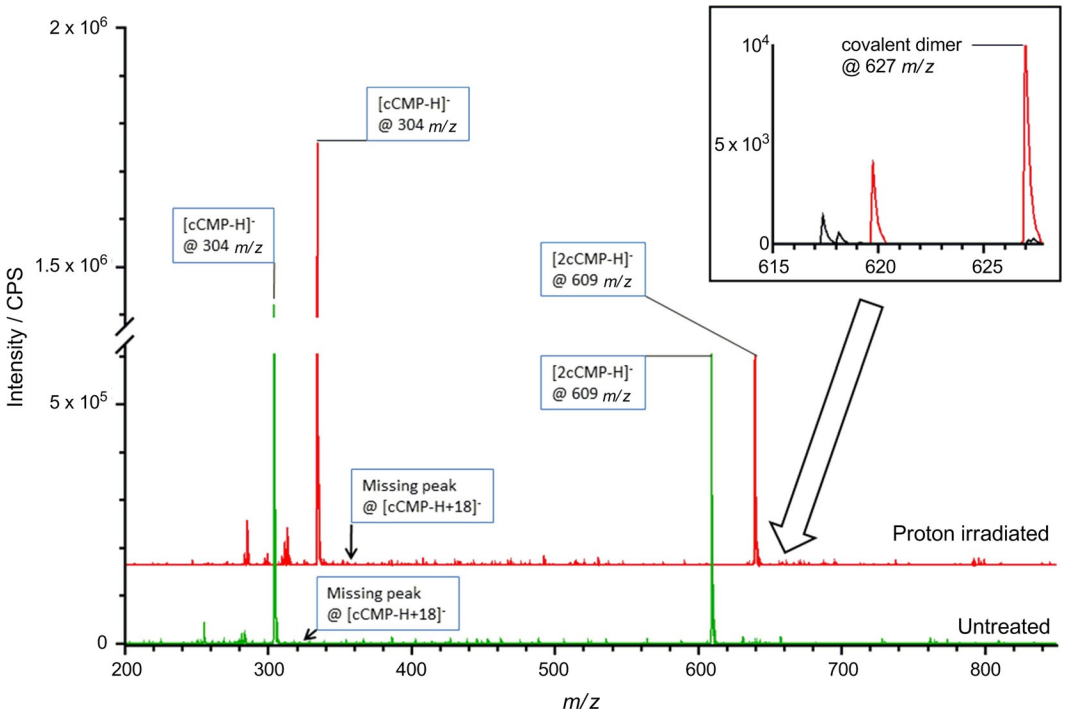



**Figure 2.** ESI-MS of the untreated and proton-irradiated 3′,5′-cCMP. Irradiation was carried out at 20 °C for 12 min. For technical details of sample preparation, see the Experimental Section.

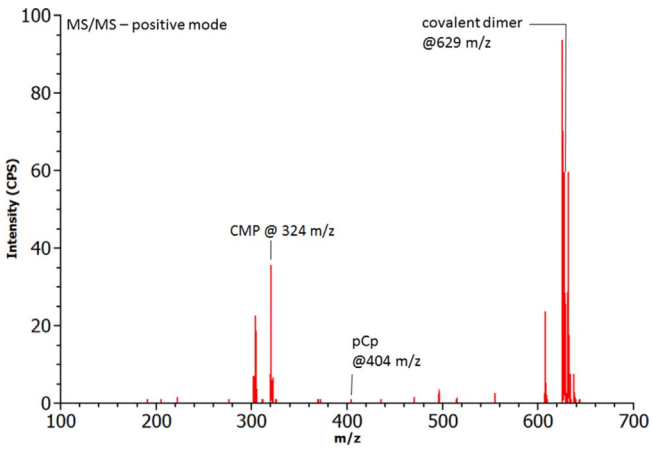



**Figure 3.** Fragmentation of the *m/z* 629 signal found in the ESI-MS spectrum of 3′,5′-cCMP irradiated with protons at 20 °C for 12 min.

amounting to 30.2 kcal mol$^{-1}$. This remarkable difference in the activation energies suggests that high-energy radicals produced by proton irradiation might be capable of inducing polymerization even at room temperature. It is important to note that an adequate theoretical description of the radical-induced chain extension step by using the model from ref. [26] is beyond the capabilities of contemporary quantum chemistry. Nonetheless, the dramatically lower activation energy computed for the attack of the •OH radical in relation to that of $OH^-$ ions qualitatively suggests that the analogous radical-induced chain propagation steps should proceed even more readily than their anionic variants (Scheme 1).

## Stacking of 3′,5′-cCMP and steric conditions for its oligomerization

Like the polymerization of 3′,5′-cGMPs, the oligomerization of 3′,5′-cCMPs proceeds in two steps: stacking of 3′,5′-cCMP and subsequent ring-opening polymerization. The optimal conditions for these two key steps are not necessarily identical.

Indeed, our quantum chemical calculations suggested that 3′,5′-cCMPs might form a stacked architecture similar to the one that supports the polymerization of 3′,5′-cGMPs[24,26] (Figure S1). This is further supported by our CD spectroscopy results (Figure 6A, B) and hyperchromicity measurements detailed below (Figure 6C).

The absorption spectra of aqueous 3′,5′-cCMP solutions exhibit a strong enhancement of the molar extinction coefficient with a sigmoidal trend (Figure 6C). This remarkable concentration-dependent hyperchromic effect provides information on the dissolved compounds' association due to stacking interactions and the structures of the resulting aggregates. According to quantum mechanics, the intensity of a spectroscopic band depends on the transition dipole moment produced by the interaction between the chromophore's electric dipole and the electric field of the electromagnetic radiation. However, neighboring chromophores can interact and perturb each other, changing the intensity of the absorption bands. The interaction between a chromophore and its neighbors can be described in terms of coupling between transition dipole moments that lie in the plane of the aromatic moieties. In parallel stacks, the dipoles are aligned in a reciprocal repelling arrangement that reduces the electric dipole moments with a consequent hypochromism. In contrast, antiparallel alignments, in

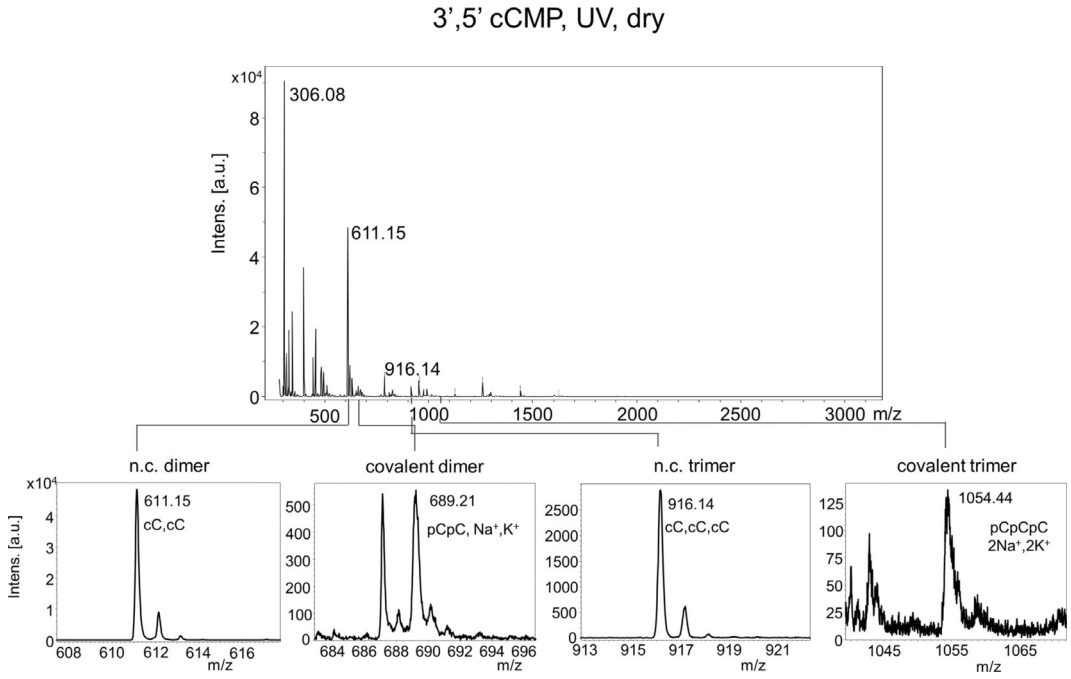



**Figure 4.** MALDI-ToF MS spectrum of the oligomerization products formed by UV irradiation of a dry 3′,5′-cCMP sample (irradiation at 254 nm for 5 min; for additional details, see the Experimental Section). Detection was made in the positive ion mode. cC stands for 3′,5′-cCMP; for pCpC and pCpCpC, see Scheme S1. n.c.: noncovalent.

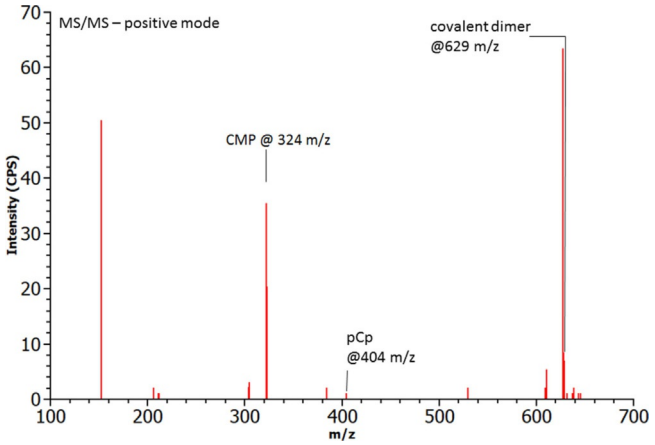
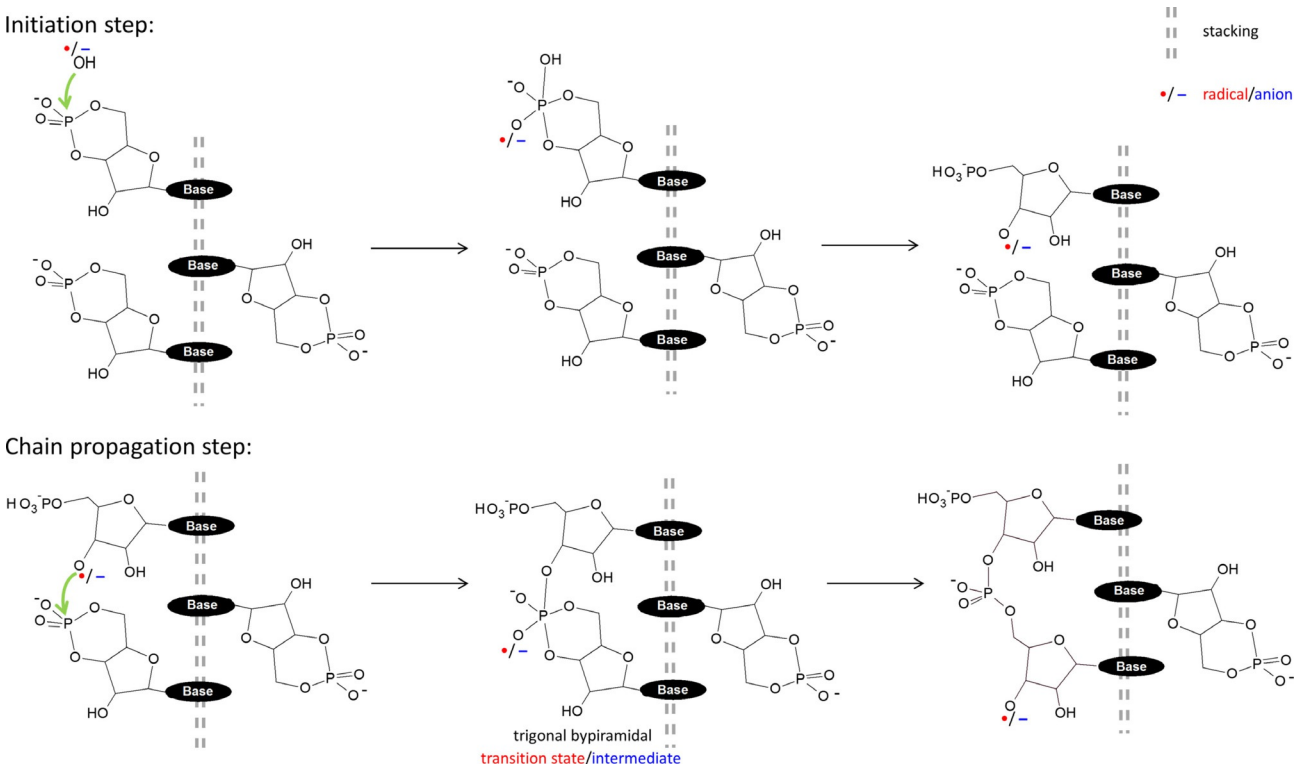



**Figure 5.** Fragmentation of the *m/z* 629 signal found in the ESI-MS spectrum of 3′,5′-cCMP irradiated with UV at 20 °C for 5 min.

which the dipoles are in an attracting geometry, increase the transition dipole moments, producing hyperchromism as a consequence.[42] The experimentally observed hyperchromism suggests chromophore–chromophore interactions and antiparallel stacking of 3′,5′-cCMPs. The concentration dependence of the hyperchromic effect arises because the intensity of such effects varies with the cube of the distance between the chromophores ($r^{-3}$). Solutions with concentrations below around 2 mM obeyed the Beer–Lambert law.

Figure 6C shows that self-assembly through stacking was observed for aqueous solutions of 3′,5′-cCMP at 20 °C, starting at a concentration of about 10 mM. The intensity of the hyperchromicity effect stopped increasing at concentrations above 20 mM, and precipitation was observed above 23 mM. In conclusion, the stacking of 3′,5′-cCMP in pure water is strongly concentration-dependent and occurs in a narrow window of concentrations.

The persistence of the stacking as a function of temperature was measured by CD spectroscopy. CD measurements were performed with a 23 mM solution of 3′,5′-cCMP in water. Figure 6A shows the normalized CD spectra of 3′,5′-cCMP at 20 °C (black solid line) and at 40 °C (black dashed line). The CD spectrum at 60 °C was very similar to that recorded at 40 °C (Figure S3A). Increasing the temperature from 20 to 40 or 60 °C dramatically changed the form and intensity of the CD spectrum. The strong decrease in the CD band intensity at higher temperatures suggests that the 3′,5′-cCMP units interact closely in stacks at room temperature but become unstacked as the temperature rises. This behavior is also reflected in the theoretically computed CD spectra of model systems derived from the stacked 3′,5′-cCMP architecture (Figure S1). Figure 6A compares the computed CD spectra of the whole stacked cCMP hexamer with those computed for simplified models consisting of the central four and two nucleotides of the original hexamer model. The band shape of the computed spectrum for the hexamer (blue curve in Figure 6A) agrees well with that of the experimentally measured spectrum at 20 °C (black solid curve), and the band positions of the computed spectrum exhibit an acceptable systematic blue shift of 0.6 eV relative to the experimentally measured spectrum. Reducing the number of cCMP units in the computational model was used to mimic the gradual degradation of the stacked supramolecular architecture. As shown by the magenta and red curves in Figure 6A, this leads




**Scheme 1.** Proposed mechanisms for the ring-opening polymerization of cyclic nucleotides initiated by $OH^-$ anions[26] and by 'OH radicals. The green arrow refers to the nucleophilic attack of the radical/anionic center formed at O3′ at the next cyclic phosphate unit in the stacked ladder.

to a systematic band splitting and to a decrease in the intensities of the computed spectra, in keeping with the results observed experimentally after increasing the temperature from 20 to 40 °C. The good qualitative agreement between the computed and experimentally measured CD spectra clearly supports the notion that a temperature increase can induce gradual unstacking of the 3′,5′-cCMP aggregates that form readily at 20 °C.

The temperature-dependence of unstacking is revealed in Figure 6B, which shows the normalized CD signal intensity measured at 270 nm and 40 °C as a function of time. The kinetic transition from the stacked to the unstacked state was abrupt, as would be expected for a pseudo-first-order transition. The stacked state of 3′,5′-cCMP was stable at room temperature but lasted for $< 1.3 \times 10^3$ s at 40 °C, and was more rapidly lost at 50 °C (Figure S4). The temperature-dependent transition was completely reversible: when the temperature was reduced to 20 °C again the stacked architecture was restored (Figure S3B).

## Conclusion

How the first RNA molecules could have arisen in the absence of elaborately preactivated precursors, presumably without enzymes and templates, possibly within a single and robust chemical frame, is an open question. Processes analogous to modern laboratory procedures based on the template-directed elongation of oligonucleotide primers through the reaction of activated nucleotides with the primer terminus in the absence of enzymes could not have occurred under prebiotic conditions when no initial template existed. Therefore, as noted previously,[12,43] the question of what might have preceded the first sequences capable of acting as templates for polymerization is not solved.

Here we describe the nonenzymatic formation of RNA di- and trinucleotides from 3′,5′-cCMP and suggest a coherent reaction pathway for this process. The reaction is enabled by the self-assembly of the monomers into stacked aggregates that can undergo transphosphorylations. Because stacking of pyrimidines is disfavored at the temperatures (ca. 80 °C) that had been found to induce oligomerization of purine nucleotides (Figure S3A) other energy sources had to be used to trigger di- and trimerization of 3′,5′-cCMPs. Our results show that proton irradiation at room temperature is particularly well-suited for this purpose. These observations represent a clear proof-of-principle that cyclic pyrimidine nucleotides can spontaneously polymerize. We deem that the noticeably less efficient UV irradiation also leads to di- and trimerization in a similar chemistry.

The self-oligomerization of 3′,5′-cCMP is a more problematic process than oligomerization of 3′,5′-cGMP, due to the weaker stacking self-association of the former. However, these difficulties do not hamper the spontaneous emergence of the first RNAs, because different nucleotides do not need to oligomerize with the same efficiency. In a recent study[44] we have suggested that short oligoG sequences could play a key role in orchestrating nonenzymatic RNA synthesis, acting as the "seeds" of the process. Thus, even dimerization or trimerization of

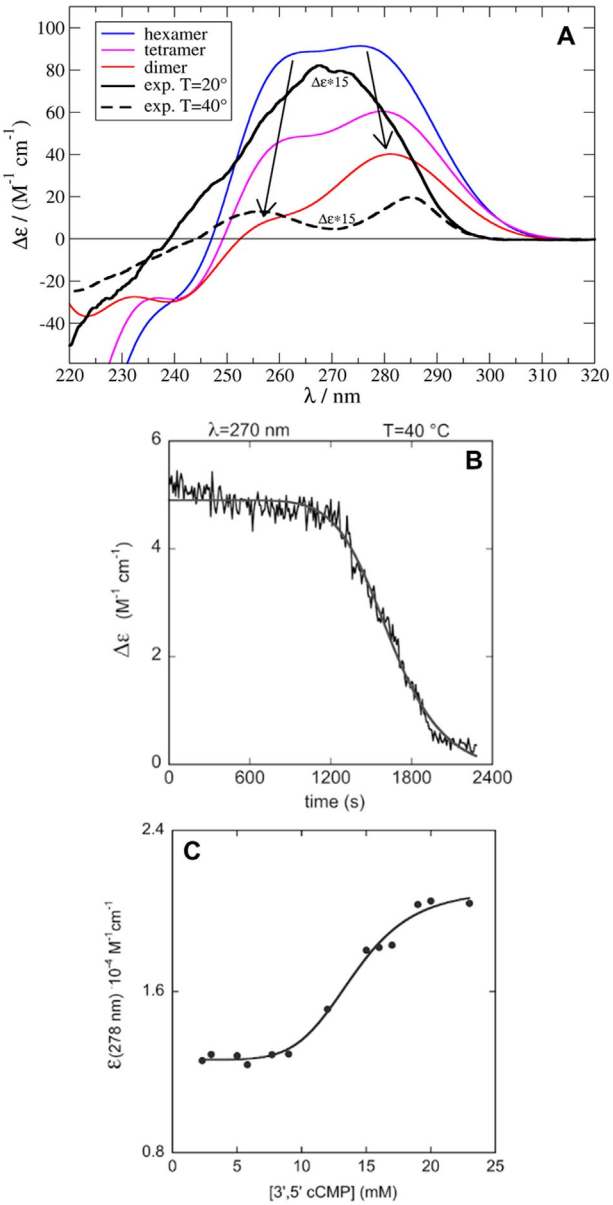



**Figure 6.** CD spectra and hyperchromicity measurements suggest the formation of stacked supramolecular architectures prior to oligomerization. A) Comparison of the experimentally measured (black solid and dashed curves) and computed (blue, magenta, red curves) CD spectra of cyclic CMPs. CD spectra were measured in a 23 mM aqueous solution of 3′,5′-cCMP at 20 (——) and 40 °C (-----). CD spectra were computed at the sTD-DFT level of theory. For details of the calculations see the section entitled Quantum chemical calculations in the Supporting Information. $\Delta\varepsilon$ stands for molar CD. B) Kinetics of unstacking as a function of time at 40 °C: variation of molar CD ($\Delta\varepsilon$) of a 23 mM aqueous solution of 3′,5′-cCMP measured at 270 nm and 40 °C as a function of time. The interpolating line is drawn as a guide to the eye. C) The hyperchromic effect is revealed in the variation of the molar extinction coefficient ($\varepsilon$) with the concentration of 3′,5′-cCMP in water at 20 °C. The interpolating line is drawn as a guide to the eye.

cCMPs could greatly facilitate the generation of the first mixed-sequence RNA chains.

As reported in refs. [28], [29], and [30], 3′,5′-cyclic nucleotides can be conveniently synthesized in formamide solution through phosphorylation of nucleosides. Nucleosides can be obtained by irradiating formamide with slow protons.[45] The proton energy (170 MeV) used in this study is similar to that of protons in solar flares.[46] A recent study by Airapetian et al. also points at the importance of particle radiation delivered by solar flares in shaping the chemical composition of the early terrestrial atmosphere.[47] In addition, the conditions used by us to trigger oligomerization are exactly the same as those used to synthesize nucleosides from formamide in ref. [45].

As discussed in refs. [48] and [49], liquid formamide would have been able to form and accumulate on the early Earth in volcanic environments at temperatures far above the boiling point of water as the thermal dissociation product[50] of ammonium formate. Formamide is part of other prebiotic scenarios, such as the "discontinuous synthesis scenario" of the origin of life, suggested by Benner and co-workers.[51] Further, recent efforts aimed at reconstructing life's origin on the basis of HCN also consider liquid formamide[52] as a plausible context for phosphorylation reactions.

Although UV irradiation is generally considered to be one of the most important energy sources available on the primeval Earth, proton irradiation is relevant to the solar flares to which our planet has continuously been exposed. In addition, proton irradiation is a component of cosmic rays, which extends the scope of the chemistry described here to an extraterrestrial context. We have previously shown that the building blocks of nucleic acids can be formed under the same conditions.[45] This suggests that the formamide route to oligonucleotides might represent a universal concept for the origin of life on Earth and in space. By invoking highly reactive species such as radicals as the initiators of oligomerization, we shift the problematics of nucleotide activation to the selection of a suitable initiator: it seems that by using a sufficiently energy-rich initiator even relatively unstable molecular architectures could serve as the starting points for polymerization.

## Experimental Section

**Materials**: Cytidine 3′,5′-cyclic monophosphate (3′,5′-cCMP) was obtained from BioLog LSI (Bremen, Germany) in its *acid form* (3′,5′-cCMP, $H^+$), as a solution (1 mM). The compound was *custom-made* and specially purified to guarantee 1) the maximum possible purity in terms of the absence of adduct-forming cations (mostly $Na^+$), and 2) the absence of evaporation or precipitation steps at any stage in the synthetic process. The purity of the final product was 99.67% (HPLC at 253 nm), as reported by the provider. MALDI-ToF MS and ESI-MS analyses showed the total absence of oligomerized materials (Figures 2 and S2). Doubly distilled deionized MilliQ water was used throughout.

The *Na-free form* was used for all of the polymerization experiments. The salt (Na) form was used in the hyperchromicity and CD spectroscopy measurements in order to avoid interference from covalently bound molecules that can only be formed from $H^+$-form precursors.[24,26]

### Methods

**Polymerization of dry 3′,5′-cCMP samples**: Dry polymerization was performed by using a pellet, which was typically obtained by concentrating a cyclic nucleotide solution (1 mM, 150 μL, 1.5×

$10^{-7}$ mol) in a Savant evaporator operating in cooling mode under vacuum until the desired dryness was obtained. The pellet was then incubated under the indicated conditions (time, temperature, irradiation), after which the reaction was stopped by quick freezing.

**Conditions for polymerization**: Proton irradiation was performed at the Phasotron facility of the Joint International Nuclear Institute, Dubna, by exposing dry pellets in Eppendorf tubes to protons. The characteristics of the irradiation were: energy = (170 ± 5) MeV, average LET (linear energy transfer) = 0.565 KeV $\mu m^{-1}$, area of uniform proton field = 8 × 8 $cm^2$, the uniform proton field was bounded 10 × 10 $cm^2$ by the collimator system, the proton flux was (1248 ± 109) proton $cm^{-2}$, dose rate = 0.5 Gy $min^{-1}$ total adsorbed dose = 6 Gy. The exposure was performed for 12 min at 20 °C.

UV irradiation was performed by exposing the sample to the light generated by a Vilber Lourmat lamp (VL-4.LC, 4 W, 254 nm wavelength, no emission below 200 nm), for 5 min at 20 °C. The intensity of the resulting light at 15 cm is 265 $\mu W\,cm^{-2}$.

**MALDI-ToF mass spectrometry**: The polymerization product (typically about 100 ng) was mixed (at a ratio of 1:1, 1:3, 1:5, or 1:10) with an aqueous 3-hydroxypyridine-2-carboxylic acid matrix solution (20 mg $mL^{-1}$) and analyzed with an AutoFlex II instrument (Bruker Daltonics, Bremen, Germany), equipped with a 337 nm nitrogen laser and operating in reflector positive mode. Calibration was performed by using the default calibration procedure and five standard spots (Mass Standards kit for Calibration P/N 4333604, AB Sciex, Framingham, MA, USA).

**MALDI-ToF/ToF mass spectrometry**: Mass spectra were obtained between 50 and 5000 Da, with 4000 laser shot intensity (Nd:YAG laser at 355 nm, at least 50 Shots/Sub-Spectrum for 2000 Total Shots/Spectrum) in reflectron positive mode with an 4800 MALDI-ToF/ToF instrument (AB Sciex). For each sample, a data-dependent acquisition method was created to select intense peaks, excluding those due to the matrix. MS/MS spectra were acquired in positive mode with 5000 laser shot intensity (Nd:YAG laser at 355 nm, at least 50 Shots/Sub-Spectrum for 2000 Total Shots/Spectrum) with use of atmospheric gas as the collision gas. Spectra were processed and analyzed by use of 4000 Series Explorer (AB Sciex).

**ESI mass spectrometry**: Mass spectra were obtained with a Bruker amaZon SL mass spectrometer by a direct-infusion method. Full-scan spectra were obtained in the negative ion mode to avoid spectral interferences. Consecutive fragmentation experiments were performed in the positive ion mode to increase the absolute analytical signal of the parent ion and corresponding fragments.

**Absorption spectroscopy**: Ultraviolet absorption spectra were measured with a JASCO V-550 spectrophotometer equipped with a Peltier device for temperature control by using quartz cells of 0.1, 0.02, 0.01, and 0.001 cm optical path length. The absorption spectra were recorded at 20 °C. All solutions for absorption and CD spectroscopy were prepared by suitable dilution of 3′,5′-cCMP in doubly distilled water.

**Electronic circular dichroism (ECD) spectroscopy**: ECD spectra were recorded in a quartz cell of 0.01 cm optical path length with a JASCO J 715 spectropolarimeter equipped with a Peltier device for temperature control. Measurements were performed in the 330–220 nm spectral range, in which cyclic CMP has the specific π–π* transitions at 20 and 40 °C. The presented ECD spectra each represent four averaged scans obtained with an instrument scanning speed of 50 nm $min^{-1}$, a response time of 1 s, and a resolution of 1 nm.

**Quantum chemical calculations**: A model of the stacked 3′,5′-cCMP architecture consisting of six nucleotides was computed with the Gaussian09 program package[53] by using the approach described in ref. [26].

Vertical singlet excitation energies and the corresponding rotatory strengths (dipole length formalism) were calculated at the sTD-DFT level[54] by using DFT-D2[55] optimized geometries. Further details of the computations are given in the section on quantum chemical calculations in the Supporting Information.

## Acknowledgements

*Thanks to Silvia Lopizzo for helpful contributions. This work is supported by the Italian Space Agency (ASI) project "Esobiologia e Ambienti estremi: dalla chimica delle Molecola alla Biologia degli estremofili" Number 2014-026-R.0 (CUP: F 92I14000030005) and by the Joint Institute for Nuclear Research (Dubna, Russia) project "Study of the biological action of high-energy heavy charged particles of different energies" 04-9-1077-2014/2016. Financial support from the project GAČR 17-05076S is gratefully acknowledged. Partial institutional funding from the project LO1305 (NPU program) of the Ministry of Education, Youth and Sports of the Czech Republic is also acknowledged.*

## Conflict of Interest

*The authors declare no conflict of interest.*

**Keywords:** cyclic CMPs · nucleotides · polymerization · proton and UV irradiation · RNA

[1] C. Cozens, V. B. Pinheiro, A. Vaisman, R. Woodgate, P. Holliger, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 8067 – 8072.
[2] J. Attwater, A. Wochner, P. Holliger, *Nat. Chem.* **2013**, *5*, 1011 – 1018.
[3] H. Mutschler, A. Wochner, P. Holliger, *Nat. Chem.* **2015**, *7*, 502 – 508.
[4] V. B. Pinheiro, A. I. Taylor, C. Cozens, M. Abramov, M. Renders, S. Zhang, J. C. Chaput, J. Wengel, S.-Y. Peak-Chew, S. H. McLaughlin, P. Herdewijn, P. Holliger, *Science* **2012**, *336*, 341 – 344.
[5] T. Wu, L. E. Orgel, *J. Am. Chem. Soc.* **1992**, *114*, 7963 – 7969.
[6] J. P. Ferris, *Orig. Life Evol. Biosph.* **2002**, *32*, 311 – 332.
[7] W. Huang, J. P. Ferris, *J. Am. Chem. Soc.* **2006**, *128*, 8914 – 8919.
[8] S. S. Mansy, J. P. Schrum, M. Krishnamurthy, S. Tobe, D. A. Treco, J. W. Szostak, *Nature* **2008**, *454*, 122 – 125.
[9] J. P. Schrum, A. Ricardo, M. Krishnamurthy, J. C. Blain, J. W. Szostak, *J. Am. Chem. Soc.* **2009**, *131*, 14560 – 14570.
[10] O. Taran, O. Thoennessen, K. Achilles, G. von Kiedrowski, *J. Syst. Chem.* **2010**, *1*, 9.
[11] C. Deck, M. Jauker, C. Richert, *Nat. Chem.* **2011**, *3*, 603 – 608.
[12] S. Egetenmeyer, C. Richert, *Chem. Eur. J.* **2011**, *17*, 11813 – 11827.
[13] A. Kaiser, S. Spies, T. Lommel, C. Richert, *Angew. Chem. Int. Ed.* **2012**, *51*, 8299 – 8303; *Angew. Chem.* **2012**, *124*, 8424 – 8428.
[14] D. R. Mills, F. R. Kramer, C. Dobkin, T. Nishihara, S. Spiegelman, *Proc. Natl. Acad. Sci. USA* **1975**, *72*, 4252 – 4256.
[15] J. F. Atkins, R. F. Gesteland, T. R. Cech in *RNA Worlds: From Life's Origin to Diversity in Gene Regulation*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, **2011**.
[16] A. Hill, Jr., L. Orgel, T. Wu, *Orig. Life Evol. Biosph.* **1993**, *23*, 285 – 290.
[17] J. W. Szostak, *J. Syst. Chem.* **2012**, *3*, 2.
[18] S. L. Beaucage, M. H. Caruthers, *Tetrahedron Lett.* **1981**, *22*, 1859 – 1862.
[19] F. H. Westheimer, *Science* **1987**, *235*, 1173 – 1178.
[20] L. E. Orgel, *Crit. Rev. Biochem. Mol. Biol.* **2004**, *39*, 99 – 123.
[21] M. S. Verlander, R. Lohrmann, L. E. Orgel, *J. Mol. Evol.* **1973**, *2*, 303 – 316.

[22] D. A. Usher, D. Yee, *J. Mol. Evol.* **1979**, *13*, 287 – 293.
[23] G. Costanzo, S. Pino, F. Ciciriello, E. Di Mauro, *J. Biol. Chem.* **2009**, *284*, 33206 – 33216.
[24] G. Costanzo, R. Saladino, G. Botta, A. Giorgi, A. Scipioni, S. Pino, E. Di Mauro, *ChemBioChem* **2012**, *13*, 999 – 1008.
[25] M. Morasch, C. B. Mast, J. K. Langer, P. Schilcher, D. Braun, *ChemBioChem* **2014**, *15*, 879 – 883.
[26] J. E. Šponer, J. Šponer, A. Giorgi, E. Di Mauro, S. Pino, G. Costanzo, *J. Phys. Chem. B* **2015**, *119*, 2979 – 2989.
[27] G. Costanzo, S. Pino, A. M. Timperio, J. E. Šponer, J. Šponer, O. Nováková, O. Šedo, Z. Zdráhal, E. Di Mauro, *PLoS ONE* **2016**, *11*, e0165723.
[28] G. Costanzo, R. Saladino, C. Crestini, F. Ciciriello, E. Di Mauro, *J. Biol. Chem.* **2007**, *282*, 16729 – 16735.
[29] R. Saladino, C. Crestini, F. Ciciriello, S. Pino, G. Costanzo, E. Di Mauro, *Res. Microbiol.* **2009**, *160*, 441 – 448.
[30] R. Saladino, G. Botta, S. Pino, G. Costanzo, E. Di Mauro, *Chem. Soc. Rev.* **2012**, *41*, 5526 – 5565.
[31] "On the Prebiotic Synthesis of Nucleobases, Nucleotides, Oligonucleotides, pre-RNA and pre-DNA Molecules", R. Saladino, C. Crestini, V. Neri, F. Ciciriello, G. Costanzo, E. Di Mauro in *Topics in Current Chemistry, Vol: 259: "Prebiotic Chemistry"* (Ed. P. Walde), Springer, Berlin, **2005**.
[32] R. Saladino, C. Crestini, V. Neri, F. Ciciriello, G. Costanzo, E. Di Mauro, *ChemBioChem* **2006**, *7*, 1707 – 1714.
[33] R. Saladino, G. Botta, M. Delfino, E. Di Mauro, *Chem. Eur. J.* **2013**, *19*, 16916 – 16922.
[34] Other prebiotically relevant phosphorylation strategies have been reported in, for example, the following works: a) K. Ozawa, A. Nemoto, E. I. Imai, H. Honda, K. Hatori, K. Matsuno, *Orig. Life Evol. Biosph.* **2004**, *34*, 465 – 471; b) C. Cheng, C. Fan, R. Wan, C. Tong, Z. Miao, J. Chen, Y. Zhao, *Orig. Life Evol. Biosph.* **2002**, *32*, 219 – 224.
[35] K. S. Dickson, C. M. Burns, J. P. Richardson, *J. Biol. Chem.* **2000**, *275*, 15828 – 15831.
[36] S. A. Rudolph, E. M. Johnson, P. Greengard, *J. Biol. Chem.* **1971**, *246*, 1271 – 1273.
[37] As illustrated in a series of studies, dry conditions have long been considered a plausible prebiotic milieu for the formation of simple oligonucleotide sequences on the early Earth. See, for example, refs. [21], [22], and [25], as well as S. A. Benner, H.-J. Kim, M. A. Carrigan, *Acc. Chem. Res.* **2012**, *45*, 2025 – 2034.
[38] S. Pino, J. E. Šponer, G. Costanzo, R. Saladino, E. Di Mauro, *Life* **2015**, *5*, 372.
[39] B. Burcar, L. Cassidy, E. Moriarty, P. Joshi, K. Coari, L. McGown, *Orig. Life Evol. Biosph.* **2013**, *43*, 247 – 261.
[40] B. T. Burcar, M. Jawed, H. Shah, L. B. McGown, *Orig. Life Evol. Biosph.* **2015**, *45*, 31 – 40.
[41] R. J. Woods, A. K. Pikaev, *Applied Radiation Chemistry, Radiation Processing*, Wiley, New York, **1993**, p. 167.
[42] C. R. Cantor, P. H. Schimmel, *Techniques for the Study of Biological Structure and Function, Part II*, W. H. Freeman, San Francisco, **1980**.
[43] W. K. Johnston, P. J. Unrau, M. S. Lawrence, M. E. Glasner, D. P. Bartel, *Science* **2001**, *292*, 1319 – 1325.
[44] J. E. Šponer, J. Šponer, E. Di Mauro, *Wiley Interdiscip. Rev.: RNA* **2017**, *8*, e1400.
[45] R. Saladino, E. Carota, G. Botta, M. Kapralov, G. N. Timoshenko, A. Y. Rozanov, E. Krasavin, E. Di Mauro, *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E2746 – E2755.
[46] F. Ferrari, E. Szuszkiewicz, *Astrobiology* **2009**, *9*, 413 – 436.
[47] V. S. Airapetian, A. Glocer, G. Gronoff, E. Hebrard, W. Danchi, *Nature Geosci.* **2016**, *9*, 452 – 455.
[48] J. E. Šponer, J. Šponer, O. Nováková, V. Brabec, O. Šedo, Z. Zdráhal, G. Costanzo, S. Pino, R. Saladino, E. Di Mauro, *Chem. Eur. J.* **2016**, *22*, 3572 – 3586.
[49] G. Cassone, J. Šponer, F. Saija, E. Di Mauro, A. Marco Saitta, J. E. Šponer, *Phys. Chem. Chem. Phys.* **2017**, *19*, 1817 – 1825.
[50] O. Kröcher, M. Elsener, E. Jacob, *Appl. Catal. B* **2009**, *88*, 66 – 82.
[51] M. Neveu, H. J. Kim, S. A. Benner, *Astrobiology* **2013**, *13*, 391 – 403.
[52] J. Xu, M. Tsanakopoulou, C. J. Magnani, R. Szabla, J. E. Šponer, J. Šponer, R. W. Góra, J. D. Sutherland, *Nat. Chem.* **2017**, *9*, 303 – 309.
[53] *Gaussian 09*, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, N. J. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, Gaussian, Inc., Wallingford, CT, USA, **2009**.
[54] C. Bannwarth, S. Grimme, *Comput. Theor. Chem.* **2014**, *1040 – 1041*, 45 – 53.
[55] S. Grimme, *J. Comput. Chem.* **2006**, *27*, 1787 – 1799.

Manuscript received: March 3, 2017
Accepted manuscript online: May 4, 2017
Version of record online: June 30, 2017