Acid-Catalyzed RNA-Oligomerization from 3′,5′-cGMP

Sreekar Wunnava,[a] Christina F. Dirscherl,[a] Jakub Výravský,[b, c] Aleš Kovařík,[d] Roman Matyášek,[d] Jiří Šponer,[d] Dieter Braun,*[a] and Judit E. Šponer*[d]

Abstract: The assembly of ancient informational polymers from nucleotide precursors is the central challenge of life's origin on our planet. Among the possible solutions, dry polymerization of 3′,5′-cyclic guanosine monophosphate (3′,5′-cGMP) has been proposed as a candidate to create oligonucleotides of 15–20 units in length. However, the reported sensitivity of the reaction to the presence of cations raised questions of whether this chemistry could be relevant in a geological context. The experiments in this study show that the presence of cations is not restrictive as long as the reaction is conducted in an acidic environment, in contrast to previous reports that suggested optimal conditions at pH 9.

Introduction

One of the major obstacles to the establishment of an RNA world is the progression from nucleotide monomers to oligonucleotide sequences. Several possible prebiotic reactions have been proposed by using high-energy phosphates such as carbodiimide derivatives or phosphorimidazolides for non-enzymatic polymerization of ribonucleotides.[1–8] However, due to their reactive nature, these intermediates/precursors are prone to hydrolysis and require feeding or in-situ recycling.[1,9] Another class of oligonucleotide precursors is cyclic phosphate-containing nucleotides which, due to the cyclization of the phosphate moiety, can act as mildly activated monomers. Both nucleoside 2′,3′- and 3′,5′-cyclic phosphates have been used as substrates for polymerization. 2′,3′-Cyclic nucleotides have been reported to polymerize moderately in both templated and template-free ways to yield short oligonucleotides in an aqueous and dry state, respectively.[10,11] On the other hand 3′,5′-cyclic nucleotides (with guanine, adenine, and cytosine) have been shown to polymerize in a template-free manner under dry conditions and rely on the stacked arrangement of the monomers.[12–15] Especially for 3′,5′-cyclic guanosine monophosphate, investigations have shown that aggregation processes associated with the low solubility of the free-acid form could have played a role in the emergence of the first oligonucleotide sequences on the early Earth.[16] Polymerization of cGMP-H has been reported to proceed in a dry state[14–16] as well as in a close-to-saturated aqueous solution[14,16–18] with optimum conditions being 80 °C and pH 9.[18] Under these conditions, a base-catalyzed anionic ring-opening polymerization chemistry has been proposed to take place.[14] Quantum chemical calculations[14] have shown that a stacked supramolecular architecture reminiscent of the one found in the crystal structures[19,20] of Na-salt (hereafter, cGMP-Na) and free-acid form (hereafter H-form or cGMP-H, the form in which all acidic groups are fully protonated) of 3′,5′-cGMP may provide favorable steric conditions for this reaction. However, while cGMP-H polymerizes on drying, cGMP-Na does not.[14,18] The previous report on the base-catalyzed mechanism explained this with the high propensity of Na^+ ions to bind to the anionic phosphate oxygen inactivating the nucleophile and thus silencing the polymerization reaction.[14]

Results and Discussion

When cGMP-H is polymerized in the presence of NaCl, only a large excess (~50x) of NaCl stops the reaction (see Figure 5c in ref. [14]). We re-evaluated the effect of Na^+ ions on cGMP-H polymerization like the one described in ref. [14]. Solutions of cGMP-H containing 1 μmol of the monomer were dried at 80 °C in the presence of different amounts of NaCl for 20 h in a vacuum evaporator. Figure 1 shows the electrophoretic profile (for qualitative assessment) and the HPLC ESI-TOF MS quantification of oligonucleotides in the prepared samples. Our current results agree with those from ref. [14] and show that only a large excess of NaCl leads to deterioration of the oligomerization process. Nonetheless, lower Na^+ concentrations do not

[a] S. Wunnava, C. F. Dirscherl, Prof. D. Braun
Department of Physics, NanoSystems Initiative Munich and Center for Nanoscience
Ludwig-Maximilians-Universität München
Amalienstrasse 54, 80799 Munich (Germany)
E-mail: dieter.braun@lmu.de

[b] J. Výravský
Tescan Brno s.r.o.
Libušina třída 1, 62300 Brno (Czech Republic)

[c] J. Výravský
Department of Geological Sciences, Faculty of Science
Masaryk University
Kotlářská 2, 61137 Brno (Czech Republic)

[d] Dr. A. Kovařík, Dr. R. Matyášek, 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

Supporting information for this article is available on the WWW under https://doi.org/10.1002/chem.202103672

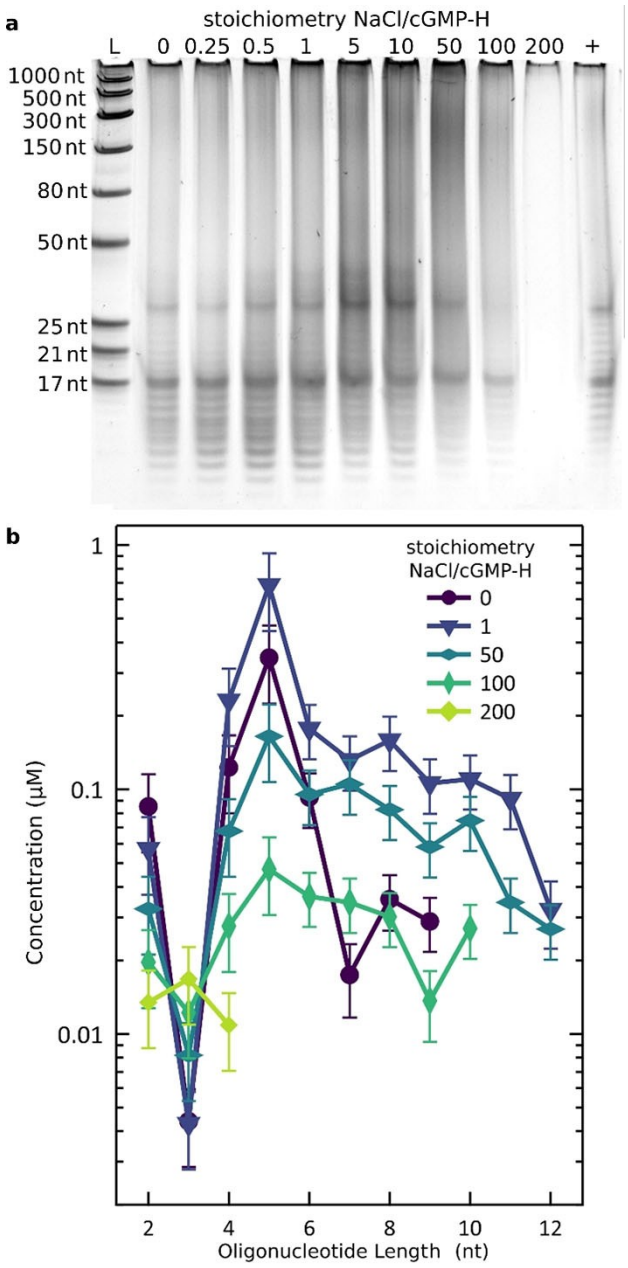

Figure 1. Polymerization of cGMP-H in presence of added Na^+ ions. A solution containing 1 μmol 3′,5′-cGMP-H (pH 3.1) was mixed with NaCl and dried for 20 h at 80 °C under vacuum and redissolved in 100 μL nuclease-free water. a) The polymerization was analyzed qualitatively by gel electrophoresis and SYBR-Gold staining. Polymerization of cGMP is observed in the presence of 0–100 molar equiv. NaCl but inhibited at 200 molar equiv. Lanes: L denotes mixed-sequence ssRNA ladder with sizes given on the left; 0–200 indicate the added micromoles of NaCl; + denotes positive control [dried 3′,5′-cGMP-H (G7504 from Sigma–Aldrich) has preformed oligomers that were ethanol precipitated]. Correspondence of the oligonucleotide sizes in the reference ladder to those of oligoG sequences is presented in Figure S1 in the Supporting Information. b) HPLC ESI-TOF quantification of the selected samples was done after ethanol precipitation. Concentrations are reported for 100 μL injection volume. A UV chromatogram and corresponding mass spectra for tetra- to hexamers isolated from the sample with 0 equiv. NaCl is given in Figure S2 along with the MS spectrum from the void volume of the column showing monomer aggregates.

hinder the polymerization of cGMP-H and hence the silencing effect of Na^+ ions added in large excess cannot explain the inability of cGMP-Na to polymerize and the issue requires further analysis. Furthermore, a base-catalyzed mechanism[14] with an optimum pH of 9[18] fails to explain how the low starting pH of cGMP-H in the experiments reported recently by Costanzo et al.[16] (pH 3.7, measured at room temperature) could still enable the polymerization. Under these low-pH conditions, the equilibrium favors the free-acid form of the molecule which is electrically neutral. Thus, these monomers are not suited to support an anionic, that is, base-catalyzed polymerization. However, being a transesterification reaction, it is very likely that a complementary acid-catalyzed polymerization mechanism might also exist.[21,22]

This motivated us to revisit the role of pH on the polymerization reaction of 3′,5′-cGMP-H. We hypothesized that the inability of cGMP-Na to polymerize upon drying is due to the different starting pH of the solution. While 1 mM solution of 3′,5′-cGMP-H has an acidic pH (ca. 3.1) and polymerizes on drying, a 1 mM cGMP-Na solution has a neutral pH (ca. 6.7–7) and does not polymerize. We confirmed this hypothesis by testing the polymerization of cGMP at different starting pH by the addition of NaOH and HCl to cGMP-H and cGMP-Na, respectively.

Figure 2a shows the dependence of polymerization on the amount of NaOH added to 1 mM cGMP-H solution before drying at 80 °C under vacuum for 20 h. We have found that already equimolar amounts of NaOH are sufficient to extinguish the polymerization reaction (Figure 2a). With the addition of 0.5 or fewer equivalents of NaOH, the resulting solution acts as a buffer with a stable pH of around 3, and, as a result, the polymerization proceeds. Similarly, starting from cGMP-Na, the solution was acidified with different molar equivalents of HCl.

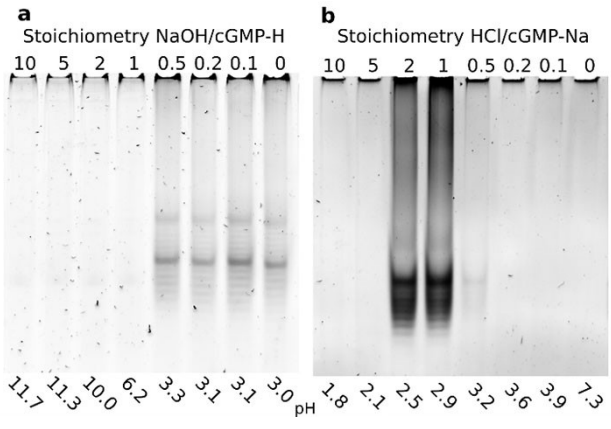

Figure 2. Polymerization of 3′,5′-cGMP depends on the pH of the starting solution. a) 1 mM cGMP-H solution was mixed with NaOH in 0.1–10 molar equiv. The resulting solutions were dried at 80 °C for 20 h under vacuum. Polymerization of cGMP is inhibited in samples with 1 or more molar equiv. of NaOH (i.e., at pH ≥ 6.2). b) 1 mM cGMP-Na solution was dried (80 °C under vacuum, 20 h) after the addition of 0.1–10 molar equiv. of HCl. Polymerization is visible only in samples with 0.5–2.0 molar equiv. of HCl (pH ~ 3). Both gels were stained with SYBR-Gold. Lanes: 0–10 denote equivalents of HCl or NaOH added to the solution; the pH of the resulting solution is given under the lane.

Upon drying at 80 °C under vacuum for 20 h, only the solutions with 0.5–2 equivalents of HCl showed polymerization (Figure 2b). Further acidification below pH 2.5 (starting from 5 molar equiv. of added HCl), however, leads to a loss in oligomers presumably due to depurination reactions. Similar depurination reactions have also been observed for the lipid-assisted polymerization of 5′- nucleotides under dehydrating acidic conditions.[23]

The addition of weaker bases like NH_4OH, triethylamine or tris(hydroxymethyl aminomethane) [hereafter, Tris] has the same inhibitory effect on the polymerization of cGMP-H as the addition of NaOH (Figure S3), thus suggesting that the pH of the cGMP solution before drying must be acidic to allow polymerization. However, the previous report[18] has shown that if Tris·HCl buffer is used instead, the polymerization proceeds even though the starting pH 8.4 was alkaline. The reported experiment[18] involved polymerization in aqueous conditions starting from a close-to-saturated solution of cGMP-H in a relatively small volume (15 μL).

On repeating the experiment as described in ref. [18], oligomer formation was indeed detected (Figure 3a). We propose the following arguments for this anomalous observation. First, due to such a small starting volume and high concentration of monomers, the reaction most likely proceeded in a semi-dry state as the liquid evaporated. Second, the pH of Tris·HCl buffer falls rapidly with increasing temperature.[24] Thus, heating from 20 to 80 °C might easily revert a weakly basic buffer to a weakly acidic one.

Finally, in the experiment in ref. [18], the concentration of cGMP-H (a weak acid) was 10 mM in 20 mM Tris·HCl buffer.

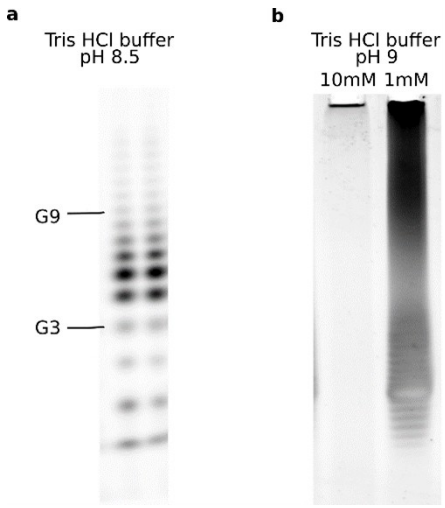

Figure 3. Polymerization of 3′,5′-cGMP in a Tris·HCl buffer. a) Two parallel samples were prepared according to ref. [18]: 150 μL 1 mM cGMP-H was dried and dissolved in 15 μL of 20 mM Tris·HCl pH 8.5 buffer. The samples were polymerized for 5 h at 85 °C in closed tubes. The polymerization products were radioactively labeled with γ^{32}P-ATP and separated by electrophoresis on a denaturing gel. G3 and G9 denote band positions of G_3 and G_9 oligonucleotide standards. b) A 1 mM solution of cGMP-H was mixed with Tris·HCl buffer to give final buffer concentrations of 10 and 1 mM at pH 9 before being dried at 80 °C for 20 h under vacuum. Polymers were observed only at 1 mM buffer concentration. The gel was stained with SYBR Gold.

Under such concentration ratios, the buffering capacity of Tris·HCl can be questioned. We tested this by drying a 1 mM cGMP-H solution in Tris·HCl buffer pH 9 at 1 and 10 mM buffer concentration at 80 °C under vacuum for 20 h. Polymerization was only observed (Figure 3b) for the solution with 1 mM buffer concentration. In light of the above arguments, it is likely that under the experimental conditions of ref. [18], the polymerization occurred from a weakly acidic solution under drying conditions. Coupled with a high concentration of cGMP-H in the solution and evaporation of the solvent, the monomers become concentrated in an acidic environment, enhancing the polymerization.

Earlier quantum chemical calculations by some of us[14] suggested that the crystal structure of cGMP (Na- as well as H-form)[19,20] provides favorable steric conditions for the trans-phosphorylation reactions leading to oligonucleotide formation. Using the theoretically derived optimum nucleotide arrangement from ref. [14], we found the potential acid-catalyzed polymerization mechanism described in Figure S4. It is thus reasonable to assume that the ability of the monomers to crystallize under given conditions might dramatically influence the experimentally observed polymerization rates. The very first step of this process is nucleation, a process in which intermolecular contacts are established between the solute molecules at the expense of disrupting interactions with the surrounding solvent environment. This process is strongly dependent on the protonation state of the nucleotide monomers and thus the pH of the chemical environment.

While cGMP-Na has a solubility of 136 mM (supplier specified value 50 mg/mL), the H-form material has a much lower solubility (predicted value 10.9 mM[25]). Unlike the sodium salt, which dissociates completely in water giving the nucleotide a negative charge, the free acid form, being a weak acid, only partially dissociates. Thus, in an aqueous solution of the free-acid form material, a large part of solute molecules is neutral (Figure 4a) and more prone to aggregation due to hydrophobic stacking interactions.

Similar behavior has been reported for aromatic carboxylic acids where it has been revealed that stacking might be a more important driving force for nucleation than H-bonding and electrostatic interactions.[26] In contrast to their free acid form, the solubility of the sodium salts of aromatic carboxylic acids is much higher and upon quick evaporation of the solvent, these materials often form amorphous phases rather than crystalline materials.[27]

This is the case for cGMP-Na as well which forms an amorphous material upon drying. In contrast, the H-form material forms well-developed crystals under the same drying conditions (Figure 4b). The transition between the H- and Na-form materials is smooth: precipitates are observed upon acidification of the Na-form material (Figure 4c). This indicates that charge neutralization of the nucleotide monomers helps the stacking-assisted nucleation process and hence crystallization. Monomer assembly in well-developed crystals of 3′,5′-cGMP then provides optimum steric conditions for the polymerization reaction.

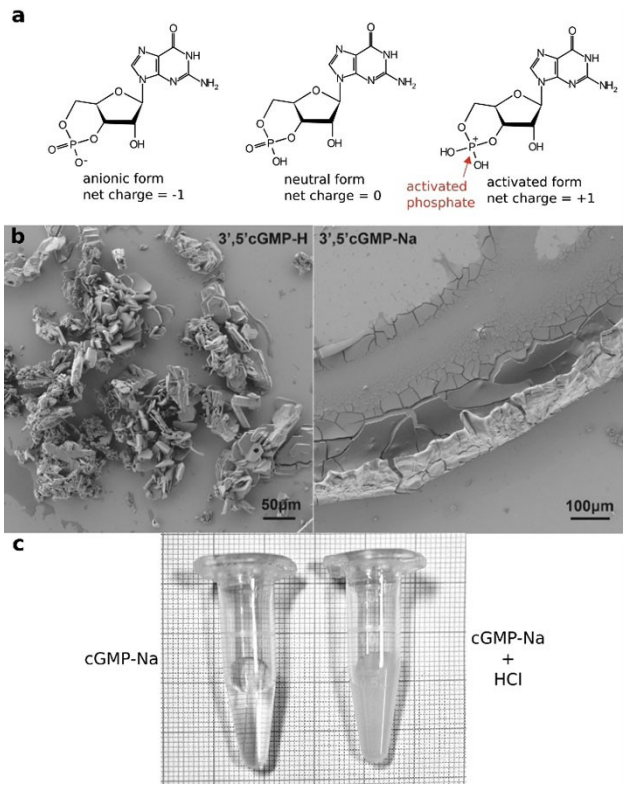

Figure 4. a) Protonation states of 3′,5′-cGMP. The phosphate group can have either a −1 charge, a +1 charge, or no charge. b) Electron microscope images of H- and Na-form 3′,5′-cGMP dried under the same conditions. Samples were deposited on glass microscope slides by dropwise evaporation of 1 mM aqueous solutions at 80 °C. While the H-form material is dominated by large (several tens of µm long) crystals, 3′,5′-cGMP-Na forms an amorphous mass. c) 3′,5′-cGMP-Na is more soluble than the H form. 3′,5′-cGMP-Na is soluble at a concentration of 50 mM (left). Upon addition of equivalent moles of HCl, the solubility of 3′,5′-cGMP is significantly reduced, and the solution becomes turbid (right). The low solubility of the H-form material is due to its ability to form stacking-assisted aggregate structures similar to aromatic carboxylic acids.[26]

Conclusion

To summarize, we have demonstrated that the nonenzymatic oligomerization of 3′,5′-cGMP in a dry state depends critically on the pH of the solution it dries in. We show that acidic conditions aid polymerization in two ways. Primarily, they enable the formation of stacked supramolecular architectures made of neutral free-acid-form monomers. This process could serve as an efficient selection force for the accumulation of 3′,5′-cGMP-H from a complex prebiotic broth. Furthermore, under acidic conditions, the monomer can be activated by binding an additional proton to the P=O group of the neutral cyclic phosphate, which, when coupled to the crystalline structure, enables polymerization. Overall, our study suggests that protonation in a drying acidic environment could have served as a simple strategy for activation on the primordial Earth and might provide the long-sought clue for the question of how the first oligonucleotide sequences formed from cyclic nucleotide precursors.

Experimental Section

Polymerization of 3′,5′-cGMP: Monomer solutions were dried such that the total amount of dried cGMP (free-acid, Biolog G001H and Na-salt, Sigma G6129) was 1 µmol. To test the effect of various ions and pH, test ions were added to the monomer solution in required stoichiometry prior to drying. The solutions were then dried in a centrifugal vacuum evaporator (Genevac EZ-2 Elite) at 80 °C for 20 h. The drying time was found to be ca. 2 h for 1 mL solution, thus the 20 h of vacuum drying includes ca. 2 h of drying followed by ca. 18 h of incubation in the dry state. The dry products were then dissolved in 100 µL of nuclease-free water such that the final cGMP concentration would be 10 mM. All reactions were done in unbuffered solutions unless specific buffer interactions were being studied. This was done to limit the interactions of buffers and other ions apart from the test ion with cGMP. The resulting solutions were solubilized by intermittent heating at 60 °C and vigorous vortexing and immediately used (without precipitation) for electrophoresis on a polyacrylamide gel.

Gel electrophoresis: Denaturing polyacrylamide gels were made at ~25% acrylamide concentration from a 40% acrylamide/bis-acrylamide (29:1) stock solution (Carl Roth A515.1) and contained 50 wt% urea and 1x TBE (from 10x, Carl Roth 3050.1). Gel staining was done with 1x SYBR Gold (from 10000x, Invitrogen, S11494) in 1x TBE (from 10x, Carl Roth 3061.1) for 5 min and rinsed with 1x TBE.

The stained gel was visualized in BIORAD ChemiDoc Gel imaging system.

Ethanol precipitation: To the dissolved samples, 20 µg of glycogen (Sigma G8751) and 500 mM ammonium acetate (Sigma A7262) was added. To this, 3 volumes of cold 100% ethanol (Carl Roth 9065.4) were added and the samples were incubated at 4 °C for 18 h. The samples were centrifuged at 15000 rpm for 30 min at 4 °C and the pellet was washed with cold 70% ethanol and centrifuged at 15000 rpm for 30 min at 4 °C. The resulting pellets were air-dried and dissolved in the required volume of nuclease-free water for downstream analysis.

HPLC-ESI TOF analysis: Agilent 1260 Infinity II LC System coupled with a 6230B Time of Flight was used for HPLC-MS analysis. Agilent AdvanceBio Oligo C18 column (4.6x150 mm, 2.7 µm) was used as the stationary phase and the elution was done under a gradient of methanol (Merk 1060352500) with 200 mM hexafluoroisopropanol (Carl Roth 2473.3) and 8 mM triethylamine (Carl Roth X875.1) as ion-pairing agents. Time-of-flight mass spectrometry was done in negative ion mode.

Scanning electron microscopic (SEM) analysis: The SEM imaging of the dried 3′,5′-cGMP, H- and Na- form samples was performed on the TESCAN CLARA SEM in high-vacuum mode with 2kV acceleration voltage, 10 pA beam current and 5 mm working distance.

Acknowledgements

The authors would like to acknowledge funding by the Advanced Grant (EvoTrap #787356) PE3, ERC-2017-ADG from the European Research Council and the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through the CRC 235 Emergence of Life (Project ID 364653263). Open Access funding enabled and organized by Projekt DEAL.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Keywords: cGMP · nonenzymatic polymerization · nucleotides · prebiotic chemistry

[1] M. Jauker, H. Griesser, C. Richert, *Angew. Chem. Int. Ed.* **2015**, *54*, 14559–14563; *Angew. Chem.* **2015**, *127*, 14767–14771.
[2] B. J. Weimann, R. Lohrmann, L. E. Orgel, H. Schneider-Bernloehr, J. E. Sulston, *Science* **1968**, *161*, 387–387.
[3] M. Dorr, P. Loffler, P.-A. Monnard, *Curr. Org. Synth.* **2013**, *9*, 735–763.
[4] L. E. Orgel, R. Lohrmann, *Acc. Chem. Res.* **1974**, *7*, 368–377.
[5] P. Wright, D. Lloyd, W. Rapp, A. Andrus, *Tetrahedron Lett.* **1993**, *34*, 3373–3376.
[6] R. Lohrmann, P. K. Bridson, L. E. Orgel, *Science* **1980**, *208*, 1464–1465.
[7] T. Walton, W. Zhang, L. Li, C. P. Tam, J. W. Szostak, *Angew. Chem. Int. Ed.* **2019**, *58*, 10812–10819; *Angew. Chem.* **2019**, *131*, 10926–10933.
[8] A. Mariani, D. A. Russell, T. Javelle, J. D. Sutherland, *J. Am. Chem. Soc.* **2018**, *140*, 8657–8661.
[9] J. D. Ibanez, A. P. Kimball, J. Oró, *Science* **1971**, *173*, 444–446.
[10] M. S. Verlander, R. Lohrmann, L. E. Orgel, *J. Mol. Evol.* **1973**, *2*, 303–316.
[11] M. Renz, R. Lohrmann, L. E. Orgel, *Biochim. Biophys. Acta Nucleic Acids Protein Synth.* **1971**, *240*, 463–471.
[12] 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.
[13] G. Costanzo, A. Giorgi, A. Scipioni, A. M. Timperio, C. Mancone, M. Tripodi, M. Kapralov, E. Krasavin, H. Kruse, J. Šponer, J. E. Šponer, V. Ranc, M. Otyepka, S. Pino, E. Di Mauro, *ChemBioChem* **2017**, *18*, 1535–1543.
[14] J. E. Šponer, J. Šponer, A. Giorgi, E. Di Mauro, S. Pino, G. Costanzo, *J. Phys. Chem. B* **2015**, *119*, 2979–2989.
[15] M. Morasch, C. B. Mast, J. K. Langer, P. Schilcher, D. Braun, *ChemBioChem* **2014**, *15*, 879–883.
[16] G. Costanzo, J. E. Šponer, J. Šponer, A. Cirigliano, P. Benedetti, V. Giliberti, R. Polito, E. Di Mauro, *ChemSystemsChem* **2021**, *3*, e202000011.
[17] G. Costanzo, S. Pino, F. Ciciriello, E. Di Mauro, *J. Biol. Chem.* **2009**, *284*, 33206–33216.
[18] G. Costanzo, R. Saladino, G. Botta, A. Giorgi, A. Scipioni, S. Pino, E. Di Mauro, *ChemBioChem* **2012**, *13*, 999–1008.
[19] A. K. Chwang, M. Sundaralingam, *Acta Crystallogr. Sect. B Struct. Crystallogr. Cryst. Chem.* **1974**, *30*, 1233–1240.
[20] K. A. Ślepokura, *Acta Crystallogr. Sect. C Struct. Chem.* **2016**, *72*, 465–479.
[21] J. Otera, *Chem. Rev.* **1993**, *93*, 1449–1470.
[22] F. A. Carey, R. J. Sundberg, *Advanced Organic Chemistry*, Springer, Boston, **2007**.
[23] C. V. Mungi, N. V. Bapat, Y. Hongo, S. Rajamani, *Life* **2019**, *9*, 1–11.
[24] K. Reineke, A. Mathys, D. Knorr, *Int. J. Food Prop.* **2011**, *14*, 870–881.
[25] I. V. Tetko, J. Gasteiger, R. Todeschini, A. Mauri, D. Livingstone, P. Ertl, V. A. Palyulin, E. V. Radchenko, N. S. Zefirov, A. S. Makarenko, V. Y. Tanchuk, V. V. Prokopenko, *J. Comput.-Aided Mol. Des.* **2005**, *19*, 453–463.
[26] A. J. Cruz-Cabeza, R. J. Davey, S. S. Sachithananthan, R. Smith, S. K. Tang, T. Vetter, Y. Xiao, *Chem. Commun.* **2017**, *53*, 7905–7908.
[27] C. Butterhof, T. Martin, W. Milius, J. Breu, *Z. Anorg. Allg. Chem.* **2013**, *639*, 2816–2821.

Manuscript received: October 12, 2021
Accepted manuscript online: November 2, 2021
Version of record online: November 23, 2021

