Check for updates

ChemSystemsChem

Research Article
doi.org/10.1002/syst.202200026

Chemistry Europe
European Chemical Societies Publishing

www.chemsystemschem.org

RNA Oligomerisation without Added Catalyst from 2′,3′-Cyclic Nucleotides by Drying at Air-Water Interfaces**

Avinash Vicholous Dass+,[a] Sreekar Wunnava+,[a] Juliette Langlais+,[a] Beatriz von der Esch,[b] Maik Krusche,[a] Lennard Ufer,[a] Nico Chrisam,[a] Romeo C. A. Dubini,[d] Florian Gartner,[f] Severin Angerpointner,[f] Christina F. Dirscherl,[a] Petra Rovó,[d, e] Christof B. Mast,[a] Judit E. Šponer,[g] Christian Ochsenfeld,[b, c] Erwin Frey,[f] and Dieter Braun*[a]

For the emergence of life, the abiotic synthesis of RNA from its monomers is a central step. We found that in alkaline, drying conditions in bulk and at heated air-water interfaces, 2′,3′-cyclic nucleotides oligomerised without additional catalyst, forming up to 10-mers within a day. The oligomerisation proceeded at a pH range of 7–12, at temperatures between 40–80 °C and was marginally enhanced by K^+ ions. Among the canonical ribonucleotides, cGMP oligomerised most efficiently. Quantification was performed using HPLC coupled to ESI-TOF by fitting the isotope distribution to the mass spectra. Our study suggests a oligomerisation mechanism where cGMP aids the incorporation of the relatively unreactive nucleotides C, A and U. The 2′,3′-cyclic ribonucleotides are byproducts of prebiotic phosphorylation, nucleotide syntheses and RNA hydrolysis, indicating direct recycling pathways. The simple reaction condition offers a plausible entry point for RNA to the evolution of life on early Earth.

Introduction

The central and multifunctional role of RNA within biology points towards RNA as a chief informational biopolymer for the onset of molecular evolution.[1] Polymerisation involving more than a single type of canonical nucleotide, generating a varied pool of RNA strands, has not been achieved under aqueous conditions.[2–6] Chemical activation strategies are deployed to trigger RNA polymerisation[3,7,8] and template-directed primer extension of sequences.[9,10] In the earliest self-replicating systems, the formation of complementary strands for replication and transfer of genetic information by non-enzymatic processes is believed to be important and homopolymers are not considered very useful as genes.[11] Short RNA strands, especially from dimers[11] to tetramers[12,13] have been shown to enhance the copying of mixed-sequence templates in comparison to monomers. Thus, it is necessary to have a oligomerisation mechanism that is able to generate short mixed-sequences that later function as primers and templates for copying of longer sequences.

We base this study on 2′,3′-cyclic mononucleotides (cNMP) which (a) possess an intrinsically activated phosphate; (b) are products of several prebiotic phosphorylation and nucleotide syntheses;[14–18] and (c) are products of neutral to alkaline chemical and enzymatic hydrolyses of RNA.[19–23] In comparison, the dry oligomerisation of 3′,5′-cGMP[24–26] did not foster the oligomerisation of the other ribonucleotides.[26] Orgel and coworkers, reported conditions for 2′,3′-cAMP oligomerisation by drying for 40 days with a 5-fold excess of ethane-1,2-diamine and yields up to 0.67% of 14-mers.[4,6] Other catalysts such as imidazole or urea required temperatures up to 85 °C and offered lower yields.[4,6]

[a] Dr. A. V. Dass,+ S. Wunnava,+ J. Langlais,+ M. Krusche, L. Ufer, N. Chrisam, C. F. Dirscherl, Dr. C. B. Mast, Prof. Dr. D. Braun
Faculty of Physics, Systems Biophysics
Ludwig-Maximilians-Universität München
Amalienstraße 54, 80799, Munich, Germany
E-mail: dieter.braun@lmu.de
[b] B. von der Esch, Prof. Dr. C. Ochsenfeld
Chair of Theoretical Chemistry, Department of Chemistry,
Ludwig-Maximilians-Universität München
Butentandtstraße 5–13, 81377, Munich, Germany
[c] Prof. Dr. C. Ochsenfeld
Max Planck Institute for Solid State Research
Heisenbergstr. 1, 70569 Stuttgart, Germany
[d] R. C. A. Dubini, Dr. P. Rovó
Faculty of Chemistry and Pharmacy,
Ludwig-Maximilians-Universität München
Butentandtstraße 5–13, 81377, Munich, Germany
[e] Dr. P. Rovó
Institute of Science and Technology Austria
c/o NMR Facility, Am Campus 1, 3400 Klosterneuburg, Austria
[f] Dr. F. Gartner, S. Angerpointner, Prof. Dr. E. Frey
Faculty of Physics, Statistical and biological physics,
Ludwig-Maximilians-Universität München,
Theresienstraße. 37, D-80333, Munich, Germany
[g] Dr. J. E. Šponer
Institute of Biophysics Academy of Sciences of the Czech Republic
Královopolská 135, 61265 Brno, Czech Republic

[+] These authors contributed equally to this work.

[**] A previous version of this manuscript has been deposited on a preprint server (http://doi.org/10.26434/chemrxiv-2022-zwh2 t-v2).

Supporting information for this article is available on the WWW under https://doi.org/10.1002/syst.202200026

© 2022 The Authors. ChemSystemsChem published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution Non-Commercial NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. ChemSystemsChem published by Wiley-VCH GmbH

We found that 2′,3′-cGMP oligomerised spontaneously under alkaline (pH 7-12) drying conditions (40-80 °C), within a day. The other canonical cNMP were relatively inert under similar conditions. Our observations of cAMP and cUMP forming up to trimers are consistent with literature.[27,28] The oligomerisation is demonstrated in the presence of bulk water at the air-water interface, within a microfluidic thermal chamber. The chamber mimics conditions of a heated, water filled volcanic rock pore that includes a gas bubble.

In an oligomerisation mixture of cNMP, we observed oligomers rich in G nucleotides, but with C, A and U incorporated at lower concentrations. Computational and modelling results suggest that the oligomers of cGMP form a self-assembled scaffold in the dry state, which could incorporate the nucleotides C, A and U to form short mixed-sequence oligomers.

Results

Polymerisation of cGMP

An aqueous solution of the sodium salt of 2′,3′-cGMP (20 mM) was dried for 18 hours at 40 °C in the presence of 40 mM KCl. Since the monomers are monosodium salts, there was an equal concentration of Na^+ ions when in solution (20 mM). All the reported concentrations throughout the article are calculated for a volume of 100 µL. The total concentration of each n-mer (oligomer) is a sum of oligomers containing the linear-phosphate (-P) and the cyclic-phosphate (-cP) on the n-mer terminus. Both endings are well discriminated by HPLC as the n-mer-cP is eluted before a n-mer-P of the same length (S2d). Typically, about 90% of the n-mers consisted of -P endings (S5d). Due to propensity of purines to form non-covalent aggregates in mass spectrometry detection,[29] a combination of HPLC and ESI-TOF techniques were used for detection of oligonucleotides. The non-covalent stacked n-mers (eg. two 4-mers) are discriminated from covalent n-mers (eg. an 8-mer) due to the higher mass of the stacked n-mers by one H_2O in the MS and the corresponding HPLC retention times of n-mers under denaturing HPLC conditions.[30–32]

The denaturing conditions of the HPLC column at 60 °C efficiently resolved synthetic oligoG n-mers without signs of aggregation, as shown in Figure 1c. It must be noted that an n-mer-cP and a cyclised n-mer of the same length would have the same mass, but are unlikely not to be discriminated by the HPLC retention times. The presence of n-mer-cP is established from the ^{31}P NMR peak at ~20 ppm in Figure 1d. Oligomers from 2- to 15-mers (S8a) were detected by HPLC-MS for cGMP oligomerisation. For quantification, only 2- to 10-mers were considered throughout the study.

The error bars can be estimated based on plots of cGMP oligomerisation (5 replicates) in S8a, with a mean standard deviation of 2.95 µM between independent runs of the experiment. The error bars are not indicated in the figures as they would appear insignificant on the log scale. For quantification, the HPLC retention times of the oligomer standards of G were

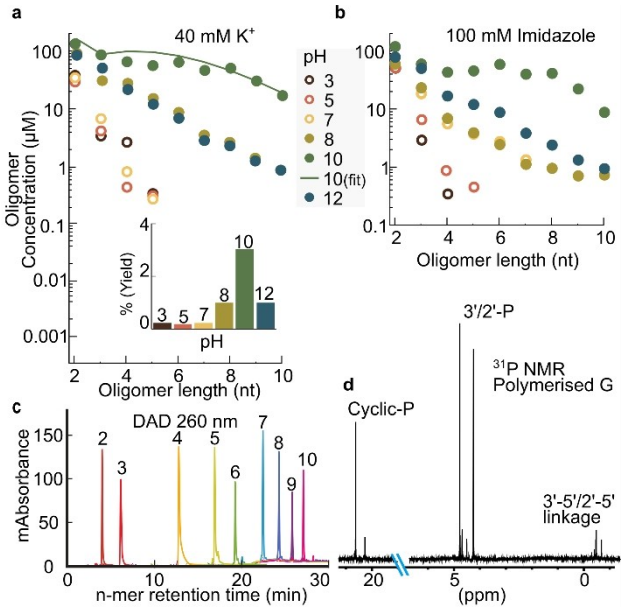

Figure 1. Oligomerisation of Guanosine-2′,3′-cyclic monophosphate (cGMP·Na). A 20 mM cGMP·Na solution was heat-dried with 40 mM K^+ at 40 °C for 18 hours, under ambient pressure in 100 µL volume. (a) Polymerisation was screened over a range of pH 3–12. The reported concentrations were the sum of terminal cyclic (-cP) and linear phosphate (-P) containing oligomers. Oligomers without terminal phosphates were not detected. Polymerisation was optimal at pH 10 with total oligomer yields of ~3.5% (inset). The solid line shows results of the polymerisation model based on stacked assembly (S19). (b) pH screen with 100 mM imidazole under similar conditions. No significant increase in oligomerisation was found by adding imidazole. (c) Diode array detector (DAD) absorbance at 260 nm for 50 µM oligoG standards (-P endings) and 100 µM KCl used for confirming HPLC separation and the determination of retention time for quantification with ion counts. (d) ^{31}P proton-decoupled NMR spectrum (10% D_2O, pH 10), of oligomerised G sample: the signals corresponding to phosphodiester linkages for both 3′-5′ and 2′-5′ are between −0.8 and −1.1 ppm.

first optimised on an RP C-18 HPLC column coupled to ESI-TOF. Figure 1c shows the HPLC chromatogram of 2- to 10-mers for oligoG standards (with 1 eqv. of KCl) with their respective retention times. We found efficient separation and no evidence for the formation of aggregates. The ion counts of the n-mer with their HPLC retention times are shown in S2b. By comparing the ion counts, we confirmed the high efficiency of the post-polymerisation ethanol precipitation protocol and its negligible influence (S3). However, the precipitation was used to remove excess monomers which would otherwise saturate the HPLC column, yielding a robust method for the quantification of the complex oligonucleotide mixtures (S2c).

The calculated isotope probabilities of the n-mers in the various charge states were fitted to the raw mass spectra using a self-written LabView program. This allowed us to identify salt adducts formed in the mass spectrometer and to fit overlapping isotope patterns. The retention times of the oligoG standards were used to obtain time-brackets to sum the mass spectra. Further details on the calibration used for the quantification within the program and the functional modes of the program are elaborated in S1–S6. Based on preliminary enzymatic digestion experiments, we estimated that the formed G

© 2022 The Authors. ChemSystemsChem published by Wiley-VCH GmbH

25704206, 2023, 1, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/syst.202200026 by Cochrane Germany, Wiley Online Library on [29/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

oligomers were linked by 2′-5′ and 3′-5′ phosphodiester linkages in about 1:1 ratio (S21–S23). The linkage type in the oligomerisation was also confirmed by ^{31}P NMR (Figure 1d) and the peaks were assigned based on the literature values.[33,34]

Figures 1a and 1b compare the effect of pH on the lengths and concentrations of the n-mers formed by drying with K^+ (Cl^-) and imidazole respectively. We determined the optimal reaction temperature to be 40 °C (S5, S8b). Imidazole and its derivatives are used in the literature as nucleotide activation agents for templated primer extension reactions,[9] as a buffering agent and a catalyst for oligomerisation.[4] The addition of imidazole did not enhance the length and concentration of n-mers in comparison to oligomerisation with K^+.

Polymerisation from cNMP

We also tested the polymerisation tendencies of cAMP, cUMP and cCMP under the same heat-drying conditions and found that these monomers did not polymerise to the same lengths and concentrations as cGMP. Figure 2a shows that the polymerisation trend decreases in the order cGMP > cUMP > cAMP > cCMP. The dominance of G-polymerisation prompted us to investigate the copolymerisation of these moderately reactive mononucleotides under the influence of the well oligomerising cGMP. We found that a mixture of two or four different monomers was capable of generating mixed sequence oligomers, where the majority of the mixed oligomers were rich in G. We probed if the oligomerisation of a G and C mixture could reach levels where hybridisation between strands could be possible. Thus, we oligomerised a binary mixture of cGMP and cCMP (20 mM each), under heat-drying conditions (40 °C) in the presence of 40 mM KCl. Comparing quantities of C_2 in Figure 2a and 2b, the concentration of C_2 is enhanced 2 fold and C_3 became detectable; besides the fact that mixed GC oligomers are formed (Figure 2b). The detailed sequence composition for GC mixed polymerisation is seen in Figure 2e, showing that the G_2 to G_{10} contribute to the bulk of the oligomers formed in the polymerisation mixture. Up to two C's were incorporated into oligomers $\leq$4-mers, one C is incorporated into 5-mers and none were detectable beyond them. A similar analysis of GA and GU binary mixtures is available in S9a, b.

GC mixed polymerisation was favoured at temperatures ranging from 40 °C to 80 °C (Figure 2c), similar to cGMP (S5b, S8b). It must be noted that in reactions at 30 °C for 18 hours, the drying was incomplete within the polypropylene tubes used for the experiment and the reaction kinetics in the dry state was reduced. Higher temperatures on the other hand possibly contributed to the degradation of the monomers (S4c) and the formed oligomers as seen in the trace comparisons under 80, 60 and 40 °C in Figure 2c.

Specific cations also influenced cNMP oligomerisation. We found that K^+ ions yielded higher concentrations and lengths of the oligomers in comparison to Na^+ ions at the same concentrations. The presence of Mg^{2+} ions in the reaction mixture inhibited polymerisation (Figure 2d). The dependence

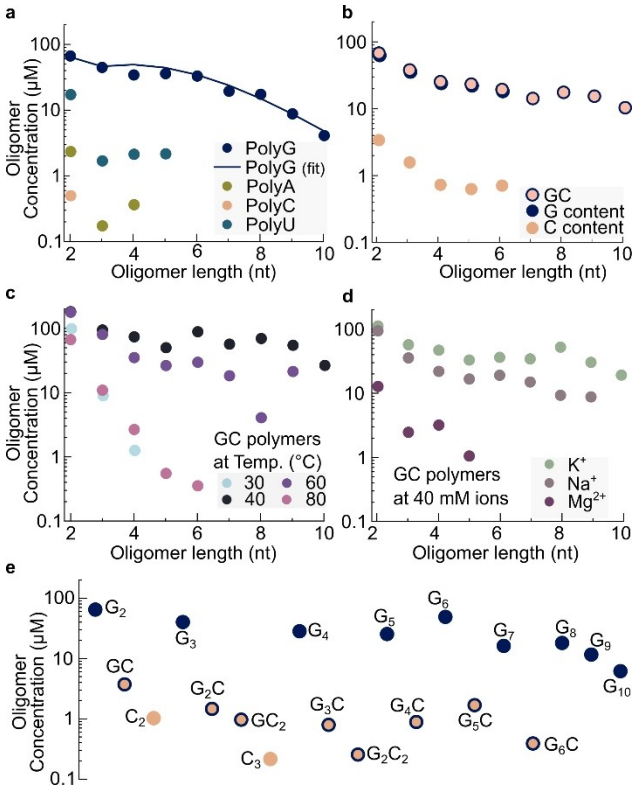

Figure 2. Oligomerisation of mixed Nucleotide 2′,3′-cyclic monophosphate (cNMP). (a) Homooligomers of cGMP·Na, cAMP·Na, cCMP·Na, cUMP·Na were individually produced from a 20 mM solution at 40 °C for 18 hours. Oligomers of G were formed in far higher concentrations than polyC, A and U. The solid line shows results of the polymerisation model based on stacked assembly (S18). The approximately 3x lower yield for oligoG compared to Figure 1a is attributed to the lack of K^+ ions. (b) Oligomerisation of cGMP and cCMP at 40 °C with 40 mM K^+. The base C is incorporated into the sequences in the presence of cGMP while only dimers were detected without it. (c) Temperature screening over a range of 30–80 °C for GC oligomerisation. Reduced concentration of n-mers > 3 is observed for 80 °C, possibly due to ring opening of the cyclic phosphate monomers (S4c). (d) The presence of 40 mM K^+ increased the concentration of n-mers while added Mg^{2+} quenched polymerisation (S10a). (e) Sequence composition of cGMP and cCMP mixed oligomers at 40 mM K^+. Oligomers show G-rich n-mers and suggest the presence of all possible combinations in trimer sequences.

of polymerisation on K^+, Na^+ and Mg^{2+} salt concentrations is shown in S10, indicating that 1–3 eqv. of the same cation display similar results, but the type of cation affected the efficiency of oligomerisation.

Polymerisation of cNMP in a heated rock pore mimic

Wet-dry cycles in surface-based geological settings are subjected to a drift in salt and pH conditions due to the imbalance caused by the evaporation of pure water and the rehydration of the fluid that contains salt. Wet-dry cycling can also occur in a closed chamber, subjected to a temperature gradient.[35] The water that evaporates on the warm side re-enters the fluid on its cold side. This causes interface shifts and the dew droplet dynamics on the cold side, offering wet-dry cycles under

© 2022 The Authors. ChemSystemsChem published by Wiley-VCH GmbH

25704206, 2023, 1, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/syst.202200026 by Cochrane Germany, Wiley Online Library on [29/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

constant pH and salt conditions. The geological analogues of such a setting would be volcanic rock pores which are partially filled with fluids and are subjected to a thermal gradient. We have previously reported prebiotically important processes such as accumulation, phosphorylation, encapsulation, gelation, strand separation, enzymatic DNA replication and crystallisation within such settings.[35–37]

For the polymerisation within this setting, we started with 20 mM total monomers (5 mM each of cG, cC, cA and cU). After the chamber was loaded with the monomer solution, a thermal gradient was applied which drove continuous wet-dry cycles just above the air-water interface inside the chamber (Figure 3a). Over time, the meniscus of the bulk liquid receded in an oscillatory manner depending on how many dew droplets formed above the interface; and dried material precipitated on the warm side as a consequence (Figure 3b). The dew droplets grew at the cooler side of the chamber by surface-tension driven fusion and made contact with the warm side, rehydrating the dried material and transporting it back into the bulk.[37] This phenomenon was allowed to continue for 18 hours, after which the setup was dismounted and the remaining bulk liquid and the dried flakes (after dissolution) were sampled for analysis.

The pH of the samples at the end the reaction was found to be lowered by a pH unit, indicating the formation of acidic species in the reaction mixture. A likely cause of the pH drop is the acidification by the ring opening of the cyclic phosphate in the mononucleotides and the oligomers (S4e, S5c, d). At higher temperatures, the pH drop was 1.5 to 2 pH units (S4e).

Despite the presence of bulk water, the oligomerisation inside the simulated volcanic-rock pore showed comparable yields as that of the heat-dried conditions. This indicates that the heated interface can access conditions favourable for polymerisation similar to bulk dried polymerisation conditions. The constant feeding of monomers from the bulk fluid could also be an important factor. A length-selective enzymatic DNA replication was reported recently within this setting, indicating the possible continuity of prebiotic chemistry in such a setting.[37]

We observed all the dimer sequence combinations and most of the trimers (Figure 3c). However, the tetramers and pentamers are predominantly sequences rich in G. The length selectivity of the HPLC allowed the detection of longer sequences. However, the isotopic fit to the raw mass spectra provided by our LabView-based analysis showed that longer species with concentrations lower than 0.2 μM were lost in the background noise of the mass spectra. Moreover, different oligomers can have similar masses (eg. Table S3 and S4), so to avoid false positives, sequences with mass overlaps were not included here. This is in addition to the rigid selection criteria, based on fitting of the isotopic distribution (S12) and only considering mass spectra within the optimised n-mer retention times of the HPLC. A full sequence composition analysis for GC and GCAU mixtures with comparison between dry polymerisation and simulated rock-pore polymerisation is provided in S11. In comparison, CAU reaction mixture yielded only dimers (S9c), indicating again the central role of G in the copolymerisation process.

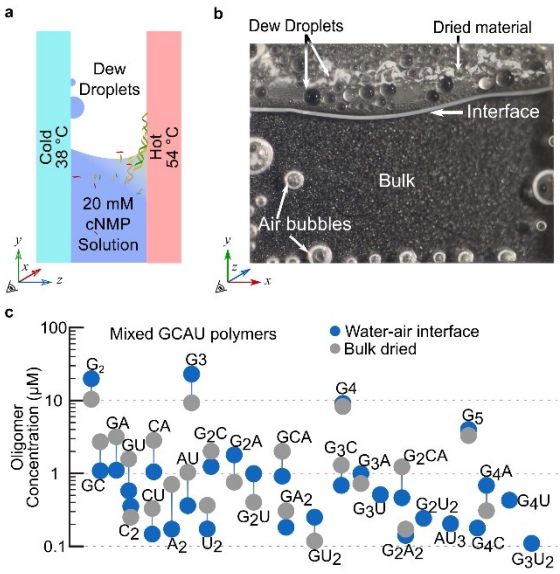

Figure 3. RNA oligomerisation in the vicinity of air inclusions in a heated simulated rock pore. (a) Side view. The chamber is 500 μm in depth and subjected to a heat flow with a temperature gradient of 38–54 °C. (b) Front view. The thermal gradient drives continuous evaporation and recondensation in the air inclusions, triggering accumulation and wet-dry cycles. Molecules accumulated at the interface are dried from a receding interface due to evaporation. Rehydration is provided by dew droplets on the cold side which merge with the bulk solution due to surface tension. (c) Oligomerisation of four canonical monomers: cGMP, cCMP, cAMP and cUMP, 5 mM each, 40 mM KCl at pH 10 for 18 hours. Especially for the longer strands, the oligomerisation in the simulated rock pore shows improved yields over the dry reaction. The physically triggered wet-dry cycling and length selectivity in this environment has been shown to drive efficient replication and selection cycles,[37] making the finding of oligomerisation to provide the raw material for templated ligation very interesting. Moreover, this shows that oligomerisation under simulated geological conditions is possible without the need for arid conditions on early Earth. The trends show a rich set of mixed short sequences when all four nucleotides are mixed together for oligomerisation.

Computational study of the proposed intercalated stacked arrangement

Based on the hypothesis that a stack-assisted geometry is triggering the oligomerisation of 3′,5′-cGMP,[26] we studied the suitability of intercalated stack arrangements for the oligomerisation of 2′,3′-cNMP. We explored the stability of the stack arrangements, and the incorporation of cNMP monomers into polymerised cGMP scaffold, based on minimum energy structures and molecular dynamics simulations (Figure 4a–c and S24–34).

To investigate the suggested intercalated stack arrangement for several possible species, we have computed the stacking interaction energies and evaluated the minimum energy geometries obtained at ωB97 M-V/def2-TZVPD level of theory.[38–41] All systems were studied in the gas phase as well as with implicit solvation (C-PCM).[42] The quantum mechanical

© 2022 The Authors. ChemSystemsChem published by Wiley-VCH GmbH

25704206, 2023, 1, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/syst.202200026 by Cochrane Germany, Wiley Online Library on [29/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

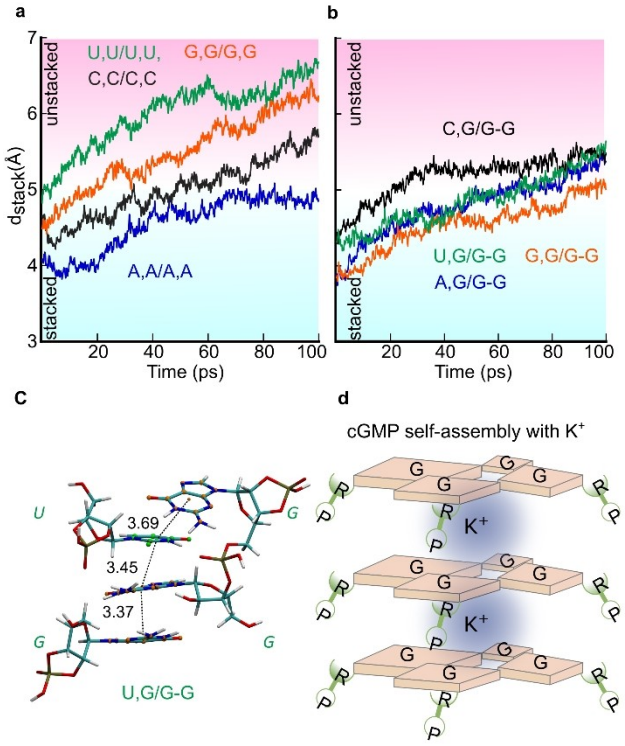

Figure 4. Possible supramolecular assemblies facilitating polymerisation of cNMP. Molecular dynamics simulations suggest polymerisation of A, U and C by intercalating into stacks of oligoG. (a) Distances between bases in the complex of unpolymerised nucleotides show that polymerisation is disfavoured due to drifting away of the complex. (b) The assembly formed when the bases are templated by covalently linked G–G (2′-5′) dimers, forms the most stable complexes and make possible the incorporation of the G, C, A, and U within the n-mers observed in our experiments. (c) Snapshot after energy minimization of stacking interaction between a oligomerised scaffold of cGMP (G–G) to template the cGMP and cUMP monomers (U,G). The dotted lines mark distances between the bases used to evaluate the stability of the complex. (d) However, the self-oligomerisation of oligoG could also be based on stacks of G-tetrads, stabilised by inner K^+ ions, coinciding with the promotion of oligoG formation by K^+ ions. R denotes the ribose of RNA.

computations were performed using FermiONs++[43–45] in combination with Chemshell.[46]

These computations were complemented by GFN-FF molecular dynamics simulations,[47] for the systems encapsulated in an explicit water sphere using xtb.[48] The stacking of homogeneous monomers were tested (N,N/N,N) with N = A, U, G, C and the incorporation of monomers into a dimer and trimer scaffold of G was probed (N,N/G–G or N,N/G–G–G). The 3′-5′ linked G–G and G–G–G accommodate A, U and G monomers into the scaffold providing a stable arrangement for the initiation of polymerisation. For C an alternate arrangement involving hydrogen bonding with a G within the scaffold is observed (S28). We found that a 2′-5′ oligoG scaffold seemed to enhance the alignment (Figure 4a, b, c), confirmed both by static and dynamic computations (S30, S31).

Theoretical model of cGMP polymerisation

Additional evidence supporting a stacked polymerisation mechanism comes from the observed non-exponential length distribution of the oligomer concentrations. This supports the idea that the formation of dimers is the rate limiting step: the concentration drop from monomers to dimers was most significant. For the cGMP oligomerisation in Figure 1a, the 20 mM monomer concentrations drop to 0.15 mM for G_2, then forming a flat concentration plateau, in contrast to the typical exponential length distribution in homogeneous polymerisation.[49]

To test this idea, we fit the concentration distribution of G homooligomers with a stacked polymerisation model (solid line Figure 1a and 2a). The model assumed a three-step polymerisation reaction: i) a monomer of length i and a oligoG scaffold k can stack together with rate ν, ii) the de-stacking rate $\delta_{k,i}$ decreases exponentially with the number of stacked bases $n_{k,i}$, iii) another monomer of length j can stack to the complex. If the stacks persist long enough, the polymerisation reaction ligates the two monomers with rate ρ (see for details S18–S19). The model fits the experimental data, suggesting a rate limiting step for the formation of short oligomers due to the required mutual alignment. It should be noted that it is difficult to distinguish between inter-base stacking or a plausible G-tetrad arrangement suggested based on the enhanced polymerisation observed with K^+ (Figure 4d).

Discussion

Our data suggests that cGMP oligomerises in dry state at moderate temperatures and pH. The oligomerisation occurs over a range of temperatures (40–80 °C) and pH (7–12) and does not require additional catalysts, making this reaction robust. Dissolved gases and salts could adjust the pH of the environment, making RNA formation more probable under early Earth models.[50,51] We also showed polymerisation in the wet-dry cycling environment at a heated air-water interface, adding RNA polymerisation to the pool of prebiotic processes possible within such a setting.[35–37] The tested conditions of wet-dry cycles at an air-water interface or direct drying keep the reaction out of equilibrium. The cyclic monomers undergo polymerisation and ring-opening (Figure 1d), of which the ring-opening is still the dominant product at the tested temperatures (S4). Under the tested conditions, the reaction yielded oligomers up to 15-mers. The formed oligoG incorporated cCMP, cAMP and cUMP monomers, albeit in lower concentration, which did not homooligomerise significantly. As a rough comparison to the yields achieved by Verlander and Orgel with homooligomers of cAMP in the presence of ethane-1,2-diamine, we observed ~0.35% for a 6-mer of oligoG in 18 hours compared to 0.81% for polyA in 40 days.[6]

An important feature of this oligomerisation is that the 2′,3′-cyclic phosphate group, under alkaline pH, is sufficient to trigger oligomerisation without ex-situ or in-situ activation mechanism or added catalysts, and under low salt conditions.

© 2022 The Authors. ChemSystemsChem published by Wiley-VCH GmbH

25704206, 2023, 1, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/syst.202200026 by Cochrane Germany, Wiley Online Library on [29/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

25704206, 2023, 1, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/syst.202200026 by Cochrane Germany, Wiley Online Library on [29/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

The finding that the oligomerisation starts without added catalysts – and that the reaction site is not yet blocked by a catalyst – is a very good starting point for Darwinian evolution to speed up this reaction rate. Low salt conditions are interesting for RNA evolution since they notably help strand separation and reduce RNA degradation.[36]

cNMP oligomerisation is found to be a relatively clean reaction under the tested conditions. In comparison, *in situ* EDC activation yields side products, especially at high temperatures.[52] We did not detect any major side products with ESI-TOF, other than the salt adducts of sodium and potassium.

The abiotic formation and recyclability of the cNMP monomers is feasible, as they are known to be produced under several phosphorylation conditions, nucleotide syntheses and are common degradation products of RNA.[19–23] Thus, with the likelihood of finding catalytic boosts for this found reaction mode, a cycle of reactions involving polymerisation, oligomer extension, polymer hydrolysis and reactivation of monomers under early Earth conditions becomes conceivable. Furthermore, recombination and templated ligation involving 2′,3′-cyclic ending oligomers[53] have been observed.

For our studies, we compared two monovalent ions (K^+, Na^+) and one divalent cation (Mg^{2+}). They were chosen for their relevance in contemporary cytosolic media, their abundance on the early Earth[54] and for the role of Mg^{2+} in ribozyme activity.[55] Polymerisation is enhanced in the presence of K^+ in comparison to Na^+ ions. The inhibition by Mg^{2+} ions possibly occurs by a combination of base catalysis mechanism, the deactivation of -cP ends of the reactant, products and enhanced oligomer hydrolysis. Despite its role in ribozyme functionality, at high concentrations Mg^{2+} inhibits RNA replication by creating strong RNA duplexes, limiting thermal denaturation and enhancing temperature dependent hydrolysis.[56] It is also known that the presence of ~1.5 mM Mg^{2+} is sufficient to inhibit the membrane self-assembly of fatty acids and this has been considered an incompatible aspect for the co-emergence of RNA and fatty acid membranes.[57,58] However, under the discussed reaction conditions of cNMP oligomerisation, RNA formation and encapsulation with fatty acids might be conceivable within freshwater locations on the primordial Earth. Moreover, we have shown that efficient strand separation can be achieved by low sodium concentrations, triggered by microscale water cycles within heated rock pores.[36,37]

Our very preliminary digestion studies and ^{31}P NMR results suggest a considerable backbone heterogeneity (2′-5′ and 3′-5′) within the oligomers. However, a full quantitative treatment is beyond the scope of this study. It has been demonstrated that the presence of 2′-5′ linkages allow efficient strand separation by reducing the melting temperature (*Tm*) of oligomers, which is pertinent in the case of G-rich sequences that are observed in this oligomerisation.[59] Lowering of *Tm* is critical to replication of sequences.[59,60] These studies also show that the presence of 2′-5′ linkages allow the folding of RNA into three-dimensional structures, similar to native linkages and do not hinder the evolution of functional RNAs, such as ribozymes. The susceptibility towards enhanced hydrolysis of the 2′-5′ over the 3′-5′ linkages could select the latter in wet-dry cycling conditions, similar to the reported backbone selection of RNA and DNA.[59,61,62]

Mechanistically, molecular dynamics studies indicated that cGMP oligomerisation could be due to the formation of intercalated stacks of cGMP as a consequence of hydrophobic interactions between the guanine bases. On attaining a stable intermolecular arrangement, the 5′-OH of a nucleotide can attack the cyclic phosphate of the neighbouring nucleotide. This could allow the formation of oligomeric G-scaffolds (Figure 4c, d, S25). However, the formation of tetrad stacks over one another with a central K^+ ion between the stacks could also promote oligomerisation (Figure 4d).

The notion of multi-molecular assemblies is supported by the presence of slow-diffusing species observed in ^{1}H, ^{31}P diffusion ordered spectroscopy (DOSY) of cGMP-KCl solution (S15, S16). Reports in literature point to self-assembly of 5′-GMP and 3′-GMP into helical stacks.[63] The presence of several slow-diffusing species indicate a range of molecular environments, making it impossible to identify a single type of self-assembly by NMR. It has also been reported that G-quadruplex structures could be stable up to a pH of ~10.8 at ambient temperatures.[64,65]

The formation of dimers appears to be a limiting step in the oligomerisation. Such a threshold behaviour is known to be an optimal control strategy for self-assembly processes.[66] With this, monomers remain available in high concentration, leading to long-tailed, non-exponential polymer distributions. This limits the total efficiency of the polymerisation but favours the formation of the oligomers, important for downstream reactions such as templated replication.

It should be noted that an efficient generation of very long and random RNA sequences would make hybridisation and replication inconceivable. At this point, the generated G-rich sequences might not seem optimal for hybridisation and replication. However, a biased pool of short oligomers (10- to 15-mers) further constrains the sequence space, favouring selectivity and making templated replication plausible.[11,67,68] We think that the findings are a first step to provide oligonucleotides for templated ligation and the emergence of an evolutionary dynamics with RNA.

Conclusion

We report the oligomerisation of canonical nucleotides that produced RNA of mixed sequences under drying conditions in bulk and at heated air-water interface. A wide range of temperatures (40–80 °C) and pH (7–12) promoted oligomerisation. Best yields were reported by mild heating (40 °C) of monomers at low salt concentrations and under alkaline drying conditions (pH 10). The reaction proceeded best at 1–2 equivalents of K^+ and Na^+, while Mg^{2+} ions inhibited it. In an equal mixture of four nucleotides, equal incorporation of all four was not observed and the mixed sequences were dominated by G. However, 2′,3′-cGMP fostered the incorporation of the otherwise scarcely reactive C, A, and U, generating short, mixed

© 2022 The Authors. ChemSystemsChem published by Wiley-VCH GmbH

sequences. This reaction under the tested temperatures, pH and salt conditions provide a novel route to fresh water oligomerisation towards short RNA strands, an important intermediate step towards providing the raw materials for an RNA-based emergence of life.

Acknowledgements

We would like to thank Ulrich Gerland, Tobias Göppel, Joachim Rosenberger and Bernhard Altaner for their helpful remarks and discussions; Thomas Matreux, Alexandra Kühnlein, Noël Yeh Martin and Maximilian Weingart for comments on the manuscript. The authors thank J. Kussmann (LMU Munich) for providing a development version of the FermiONs++ program package. Financial support was provided by the European Research Council (ERC Evotrap, grant no. 787356, the Simons Foundation (grant no. 327125), the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 364653263 – TRR 235 (CRC 235), the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2094 – 390783311, and the Center for NanoScience. 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: RNA · Polymerisation · Prebiotic chemistry · Non-equilibrium · Air-water interfaces

[1] H. S. Bernhardt, *Biol. Direct* **2012**, *7*, 1–10.
[2] 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*, 1–14.
[3] J. P. Ferris, A. R. Hill, R. Liu, L. E. Orgel, *Nature* **1996**, *381*, 59–61.
[4] M. S. Verlander, R. Lohrmann, L. E. Orgel, *J. Mol. Evol.* **1973**, *2*, 303–316.
[5] C. V. Mungi, N. V. Bapat, Y. Hongo, S. Rajamani, *Life* **2019**, *9*, 1–11.
[6] M. S. Verlander, L. E. Orgel, *J. Mol. Evol.* **1974**, *3*, 115–120.
[7] A. Luther, R. Brandsch, G. Von Kiedrowski, *Nature* **1998**, *396*, 245–248.
[8] S. J. Zhang, D. Duzdevich, D. Ding, J. W. Szostak, *Proc. Nat. Acad. Sci.* **2022**, *119*, 2021.09.07.459201.
[9] 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.
[10] E. Kervio, M. Sosson, C. Richert, *Nucleic Acids Res.* **2016**, *44*, 5504–5514.
[11] C. Richert, G. Leveau, D. Pfeffer, B. Altaner, E. Kervio, U. Gerland, F. Welsch, *Angew. Chem. Int. Ed.* **2022**, *202203067*, 1–6.
[12] L. Li, N. Prywes, C. P. Tam, D. K. Oflaherty, V. S. Lelyveld, E. C. Izgu, A. Pal, J. W. Szostak, *J. Am. Chem. Soc.* **2017**, *139*, 1810–1813.
[13] D. K. O'Flaherty, N. P. Kamat, F. N. Mirza, L. Li, N. Prywes, J. W. Szostak, P. Sheeringa, *J. Am. Chem. Soc.* **1997**, *140*, 1–5.
[14] E. I. Jiménez, C. Gibard, R. Krishnamurthy, *Angew. Chem. Int. Ed.* **2020**, *60, 19, 1077–10783*, DOI 10.1002/anie.202015910.
[15] Z. Liu, L. F. Wu, J. Xu, C. Bonfio, D. A. Russell, J. D. Sutherland, *Nat. Chem.* **2020**, *12*, 3–10.
[16] H. J. Kim, S. A. Benner, *Astrobiology* **2021**, *21*, 298–306.
[17] M. W. Powner, B. Gerland, J. D. Sutherland, *Nature* **2009**, *459*, 239–242.
[18] Y. Yamagata, H. Inoue, K. Inomata, *Origins Life Evol. Biospheres* **1995**, *25*, 47–52.
[19] R. Breslow, *Acc. Chem. Res.* **1991**, *24*, 317–324.
[20] Y. Li, R. R. Breaker, *J. Am. Chem. Soc.* **1999**, *121*, 5364–5372.
[21] H. Peng, B. Latifi, S. Müller, A. Lupták, I. A. Chen, *RSC Chem. Biol.* **2021**, *2*, 1370–1383.
[22] A. M. Pyle, *Science* **1993**, *261*, 709–714.
[23] S. I. Nakano, D. M. Chadalavada, P. C. Bevilacqua, *Science* **2000**, *287*, 1493–1497.
[24] M. Morasch, C. B. Mast, J. K. Langer, P. Schilcher, D. Braun, *ChemBioChem* **2014**, *15*, 879–883.
[25] J. E. Šponer, J. Šponer, A. Giorgi, E. Di Mauro, S. Pino, G. Costanzo, *J. Phys. Chem. B* **2015**, *119*, 2979–2989.
[26] S. Wunnava, C. F. Dirscherl, J. Výravský, A. Kovařík, R. Matyášek, J. Šponer, D. Braun, J. E. Šponer, *Chem. A Eur. J.* **2021**, *27, 70, 17581–17585*, DOI 10.1002/chem.202103672.
[27] C. M. Tapiero, J. Nagyvary, *Nature* **1971**, *231*, 42–43.
[28] S. Dagar, S. Sarkar, S. Rajamani, *RNA* **2020**, *26*, 756–769.
[29] J. E. Šponer, J. Šponer, A. Kovařík, O. Šedo, Z. Zdráhal, G. Costanzo, E. Di Mauro, *Life* **2021**, *11*, 1–13.
[30] A. Premstaller, P. J. Oefner, *LCGC Eur.* **2002**, *15, 7, 410–422*.
[31] A. Premstaller, P. J. Oefner, *Denaturing High-Performance Liquid Chromatography*, Humana Press, New Jersey, **n.d**.
[32] P. B. Danielson, R. Kristinsson, R. J. Shelton, G. S. LaBerge, *Expert Rev. Mol. Diagn.* **2005**, *5*, 53–63.
[33] H. R. Palmer, J. J. Bedford, J. P. Leader, R. A. J. Smith, *J. Biol. Chem.* **2000**, *275*, 27708–27711.
[34] S. Motsch, D. Pfeffer, C. Richert, *ChemBioChem* **2020**, *21, 2013–2018*.
[35] M. Morasch, J. Liu, C. F. Dirscherl, A. Ianeselli, A. Kühnlein, K. Le Vay, P. Schwintek, S. Islam, M. K. Corpinot, B. Scheu, D. B. Dingwell, P. Schwille, H. Mutschler, M. W. Powner, C. B. Mast, D. Braun, *Nat. Chem.* **2019**, *11*, 779–788.
[36] A. Ianeselli, C. B. Mast, D. Braun, *Angew. Chem. Int. Ed.* **2019**, *58*, 13155–13160; *Angew. Chem.* **2019**, *131*, 13289–13294.
[37] A. Ianeselli, M. Atienza, P. W. Kudella, U. Gerland, C. B. Mast, D. Braun, *Nat. Phys.* **2022**, *18, 579–585*, DOI 10.1038/s41567-022-01516-z.
[38] N. Mardirossian, M. Head-Gordon, *J. Chem. Phys.* **2016**, *144*, DOI 10.1063/1.4952647.
[39] O. A. Vydrov, T. Van Voorhis, *J. Chem. Phys.* **2010**, *133*, DOI 10.1063/1.3521275.
[40] F. Weigend, *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057–1065.
[41] F. Weigend, R. Ahlrichs, *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.
[42] M. Cossi, N. Rega, G. Scalmani, V. Barone, *J. Comput. Chem.* **2003**, *24*, 669–681.
[43] J. Kussmann, C. Ochsenfeld, *J. Chem. Phys.* **2013**, *138*, DOI 10.1063/1.4796441.
[44] J. Kussmann, C. Ochsenfeld, *J. Chem. Theory Comput.* **2015**, *11*, 918–922.
[45] J. Kussmann, C. Ochsenfeld, *J. Chem. Theory Comput.* **2017**, *13*, 3153–3159.
[46] P. Sherwood, A. H. De Vries, M. F. Guest, G. Schreckenbach, C. R. A. Catlow, S. A. French, A. A. Sokol, S. T. Bromley, W. Thiel, A. J. Turner, S. Billeter, F. Terstegen, S. Thiel, J. Kendrick, S. C. Rogers, J. Casci, M. Watson, F. King, E. Karlsen, M. Sjøvoll, A. Fahmi, A. Schäfer, C. Lennartz, *J. Mol. Struct.* **2003**, *632*, 1–28.
[47] S. Spicher, S. Grimme, *Angew. Chem.* **2020**, *132*, 15795–15803; *Angew. Chem. Int. Ed.* **2020**, *59*, 15665–15673.
[48] C. Bannwarth, E. Caldeweyher, S. Ehlert, A. Hansen, P. Pracht, J. Seibert, S. Spicher, S. Grimme, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2021**, *11*, 1–49.
[49] C. B. Mast, S. Schink, U. Gerland, D. Braun, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 8030–8035.
[50] J. F. Kasting, M. T. Howard, *Philos. Trans. R. Soc. London Ser. B* **2006**, *361*, 1733–1741.
[51] J. Krissansen-Totton, G. N. Arney, D. C. Catling, *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 4105–4110.
[52] E. Edeleva, A. Salditt, J. Stamp, P. Schwintek, J. Boekhoven, D. Braun, *Chem. Sci.* **2019**, *10*, 5807–5814.
[53] A. V. Lutay, E. L. Chernolovskaya, M. A. Zenkova, V. V. Vlasov, *Dokl. Biochem. Biophys.* **2005**, *401*, 163–166.
[54] S. Maurer, *Life* **2017**, *7*, DOI 10.3390/life7040044.
[55] T. S. Lee, C. S. López, G. M. Giambaşu, M. Martick, W. G. Scott, D. M. York, *J. Am. Chem. Soc.* **2008**, *130*, 3053–3064.

© 2022 The Authors. ChemSystemsChem published by Wiley-VCH GmbH

25704206, 2023, 1, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/syst.202200026 by Cochrane Germany, Wiley Online Library on [29/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

[56] A. Salditt, L. M. R. Keil, D. P. Horning, C. B. Mast, G. F. Joyce, D. Braun, *Phys. Rev. Lett.* **2020**, *125*, 48104.
[57] I. A. Chen, K. Salehi-Ashtiani, J. W. Szostak, *J. Am. Chem. Soc.* **2005**, *127*, 13213–13219.
[58] K. Adamala, J. W. Szostak, *Science* **2013**, *342*, 1098–1100.
[59] A. E. Engelhart, M. W. Powner, J. W. Szostak, *Nat. Chem.* **2013**, *5*, 390–394.
[60] J. Sheng, L. Li, A. E. Engelhart, J. Gan, J. Wang, J. W. Szostak, *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 3050–3055.
[61] S. Bhowmik, R. Krishnamurthy, *Nat. Chem.* **2019**, *11*, 1009–1018.
[62] A. Mariani, J. D. Sutherland, *Angew. Chem. Int. Ed.* **2017**, *56*, 6563–6566; *Angew. Chem.* **2017**, *129*, 6663–6666.
[63] M. Gellert, M. N. Lipsett, D. R. Davies, *Proc. Natl. Acad. Sci. USA* **1962**, *48*, 2013–2018.
[64] Y. Y. Yan, J. H. Tan, Y. J. Lu, S. C. Yan, K. Y. Wong, D. Li, L. Q. Gu, Z. S. Huang, *Biochim. Biophys. Acta Gen. Subj.* **2013**, *1830*, 4935–4942.
[65] R. Del Villar-Guerra, R. D. Gray, J. B. Chaires, *Curr. Protoc. Nucleic Acid Chem.* **2017**, *2017*, 17.8.1–17.8.16.
[66] F. M. Gartner, I. R. Graf, E. Frey, *Proc. Natl. Acad. Sci. USA* **2022**, *119*, DOI 10.1073/pnas.2116373119.
[67] P. W. Kudella, A. V. Tkachenko, A. Salditt, S. Maslov, D. Braun, *Proc. Natl. Acad. Sci. USA* **2021**, *118*, DOI 10.1073/pnas.2018830118.
[68] S. Toyabe, D. Braun, *Phys. Rev. X* **2019**, *9*, 11056.

Manuscript received: July 29, 2022
Accepted manuscript online: September 19, 2022
Version of record online: October 11, 2022

© 2022 The Authors. ChemSystemsChem published by Wiley-VCH GmbH

25704206, 2023, 1, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/syst.202200026 by Cochrane Germany, Wiley Online Library on [29/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License