



Potentially prebiotic isocyanide activation chemistry drives RNA assembly

 Stephanie J. Zhang,^{ib a} Xiwen Jia,^{ib † b} Saurja DasGupta^{ib * cd} and Jack W. Szostak^{ib * ef}

 Cite this: *Chem. Commun.*, 2025, 61, 13920

 Received 15th July 2025,
Accepted 5th August 2025

DOI: 10.1039/d5cc03937a

rsc.li/chemcomm

Nonenzymatic assembly of RNA from chemically activated building blocks, such as phosphorimidazolides, would have been essential for the emergence of ribozymes on the early Earth. We previously showed that ribonucleoside monophosphates can be activated to phosphorimidazolides via a potentially prebiotic phospho-Passerini reaction involving 2-aminoimidazole, 2-methylbutyraldehyde, and methyl isocyanide, and that these activated nucleotides enable template-directed nonenzymatic RNA polymerization in the same reaction mixture. Here, we demonstrate that the same chemistry activates oligoribonucleotides and drives both nonenzymatic and ribozyme-catalyzed RNA ligation within the same reaction environment. By demonstrating a continuous path from prebiotic activation chemistry to RNA template copying by both nonenzymatic and ribozyme-catalyzed ligation, our results provide a more integrated and realistic model for primordial RNA assembly.

The RNA World hypothesis posits that early life relied on RNA both as the carrier of genetic information and as a catalyst, in the form of ribozymes.^{1,2} Ribozymes could have emerged *via* chemical self-assembly pathways such as ligation³ of short oligoribonucleotides. In the early stages, RNA replication may have occurred primarily at the level of short oligonucleotides, while longer functional RNAs did not need to be replicated.⁴ The eventual appearance of ribozyme ligases and polymerases would have enhanced the efficiency of RNA copying pathways, enabling the direct replication of longer RNA sequences, including ribozymes. Before such ribozymes emerged, however,

template-directed RNA copying had to rely on energy-rich activated nucleotides such as phosphorimidazolides.

We previously demonstrated that *in situ* nucleotide activation, using a potentially prebiotic phospho-Passerini chemistry based on methyl isocyanide (MeNC), 2-aminoimidazole (2AI), and 2-methylbutyraldehyde (2MBA), can drive nonenzymatic RNA copying *via* template-directed polymerization.^{5,6} This chemistry, likely to occur in UV-irradiated environments rich in ferrocyanide and nitroprusside, provides a direct link between monomer activation and polymerization and could potentially support continuous RNA copying under prebiotic conditions. High concentration of 2AI inhibits the accumulation of the 5'-5' imidazolium-bridged dinucleotide intermediate required for nonenzymatic RNA polymerization.^{7,8} Nonenzymatic RNA ligation, however, does not rely on this intermediate and proceeds in the presence of excess 2AI.^{7,9} Although slower than polymerization,¹⁰⁻¹² ligation better accommodates structured or A/U-rich templates¹³ and exhibits greater fidelity,¹⁴ while requiring fewer condensation steps.¹⁵ *In situ* activation of short oligonucleotides (≤ 4 nt) has been shown to support polymerization by forming bridged intermediates with monomers.⁹ However, ligation driven by the *in situ* activation of oligonucleotides has not been demonstrated.

The emergence of ribozymes marked a key transition in the origin of life. We previously used *in vitro* selection to identify ligase ribozymes that use 2-aminoimidazole-activated oligoribonucleotides as substrates, thus bridging nonenzymatic and enzymatic RNA ligation.¹⁶ However, the hydrolytic instability of phosphorimidazolide substrates limits their accumulation, posing a major challenge to uninterrupted RNA assembly. Within protocells, substrate activation would support both chemical and ribozyme-catalyzed RNA ligation, and would have been essential for maintaining effective substrate concentrations and mitigating the inhibitory effects of unactivated substrates.¹⁷ Here, we outline a prebiotically relevant scenario in which *in situ* activation of RNA substrates enables both nonenzymatic and ribozyme-catalyzed RNA ligation within a shared reaction environment, thus paving the way for the rapid propagation of RNA sequences (Fig. 1).

^a Department of Pathology, Brigham and Women's Hospital, Boston, Massachusetts 02115, USA

^b Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138, USA

^c Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, Indiana 46556, USA. E-mail: sdasgupta@nd.edu

^d Department of Biological Sciences, University of Notre Dame, Notre Dame, Indiana 46556, USA

^e Department of Chemistry, The University of Chicago, Chicago, Illinois 60637, USA. E-mail: jwszostak@uchicago.edu

^f Howard Hughes Medical Institute, The University of Chicago, Chicago, Illinois 60637, USA

† Present address: Novo Nordisk, Lexington, Massachusetts 02421, USA.



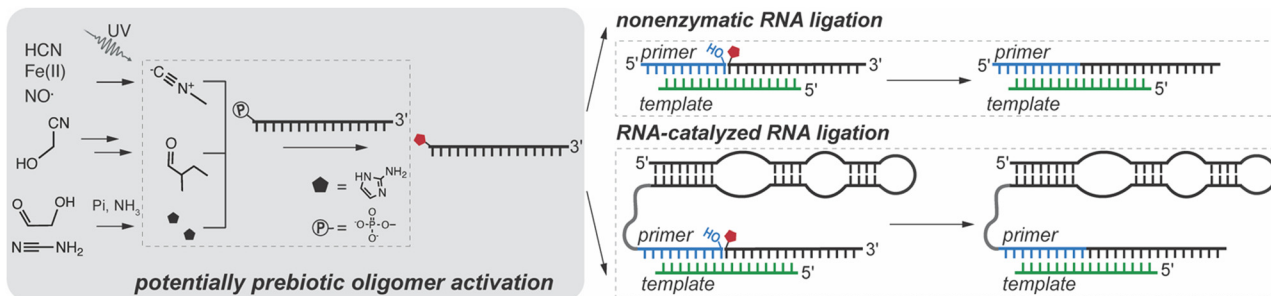


Fig. 1 Prebiotic RNA copying via templated ligation of activated oligonucleotides generated through potentially prebiotic pathways. The left panel shows a potentially prebiotic pathway for oligoribonucleotide activation using methyl isocyanide, 2-methyl butyraldehyde, and 2-aminoimidazole. The right panels illustrate two RNA ligation processes: (1) nonenzymatic ligation, where the activated oligonucleotide is chemically ligated to a primer aligned on a template and (2) ribozyme-catalyzed ligation, where a ribozyme facilitates the templated ligation of an activated oligonucleotide to the primer.

Having previously demonstrated the activation of mononucleotides and short oligonucleotides by phospho-Passerini chemistry (henceforth, isocyanide chemistry), we sought to first optimize the activation of longer oligonucleotides, which generally show reduced activation efficiencies.^{5,6,9} Given that longer oligonucleotides were likely less abundant than shorter oligonucleotides and mononucleotides in prebiotic settings,^{18,19} we began by incubating 10 μM of a 16-nt 5'-monophosphorylated oligoribonucleotide substrate (henceforth, ligator; Table S2, SI) with 50 mM MeNC, 50 mM 2MBA, 50 mM 2AI, and 30 mM Mg^{2+} in 200 mM HEPES at pH 8. High-performance liquid chromatography (HPLC) analysis revealed $25.3 \pm 0.2\%$ conversion of the 5'-monophosphorylated ligator to its 5'-phosphor-2-aminoimidazolide form after 6 hours (Fig. S1A, SI). Extending the reaction to 24 hours did not significantly increase the activation yield, because the reaction had plateaued by 3 hours (Fig. S1B, SI). These conditions served as a baseline for further optimization. Unlike nonenzymatic polymerization, ligation does not require 5'-5' imidazolium-bridged intermediates, which allowed us to raise 2AI concentrations. After optimization, we achieved >90% activation using 250 mM 2AI and either 250 or 500 mM MeNC and 2MBA (Table 1, reactions 6 and 8). Reducing the concentrations of MeNC or 2MBA below 250 mM caused a marked drop in activation yields, underscoring the need for excess aldehyde

and isocyanide, consistent with prior studies on mononucleotide activation.^{5,8,9} Interestingly, further increasing 2AI concentrations by 5- to 10-fold led to worse activation (Table 1, reactions 3 and 4). In prebiotic environments, 2AI levels could have been limited by UV-induced degradation.²⁰ Our results indicate that hundreds of millimolar MeNC and 2MBA, and tens to hundreds of millimolar 2AI could have supported continuous RNA polymerization and ligation by maintaining a diverse pool of activated nucleotides and oligonucleotides (Table 1 and Table S1, SI).

We used the optimal condition of 250 mM 2AI, 500 mM MeNC, and 250 mM 2MBA (Table 1, condition 6) to generate activated substrates for nonenzymatic RNA ligation (Fig. 2). Ligators were subjected to activation conditions for 6 hours and subsequently added to the primer–template duplex for ligation without further purification (Fig. 2A). Ligation yields after 3 hours with isocyanide-activated ligators or purified, pre-activated ligators generated *via* EDC coupling were comparable ($\sim 1.5\%$ vs. $\sim 2\%$) (Fig. S2 and S3, SI), reaching approximately 26% ligation after 24 hours (Fig. 2C). To better simulate prebiotic conditions, where no artificial boundaries

Table 1 Oligonucleotide activation yields under various conditions as determined by HPLC analysis. Reactions were carried out with 10 μM 5'-phosphorylated ligator in 200 mM HEPES at pH 8, 30 mM Mg^{2+} , and varying concentrations of 2AI, 2MBA, and MeNC. Samples were incubated at room temperature for 6 hours and then quenched (the optimal time point as identified in Fig. S1, SI). Conditions yielding >90% activation are shown in bold, with the optimal condition (reaction condition 6) highlighted in italics. Error values represent the standard deviation of the mean, based on at least two replicates ($n \geq 2$)

| Reaction | [2AI] (mM) | [MeNC] (mM) | [2MBA] (mM) | Activation yield (%) |
|----------|------------|-------------|-------------|----------------------------------|
| 1 | 50 | 50 | 50 | 25.3 ± 0.2 |
| 2 | 50 | 250 | 250 | 83.2 ± 1.2 |
| 3 | 250 | 250 | 250 | 67.5 ± 1.5 |
| 4 | 500 | 250 | 250 | 44.7 ± 3.5 |
| 5 | 250 | 50 | 250 | 15.8 ± 1.9 |
| 6 | 250 | 500 | 250 | 95.6 ± 1.9 |
| 7 | 250 | 250 | 50 | 28.4 ± 0.9 |
| 8 | 250 | 250 | 500 | 92.2 ± 2.5 |

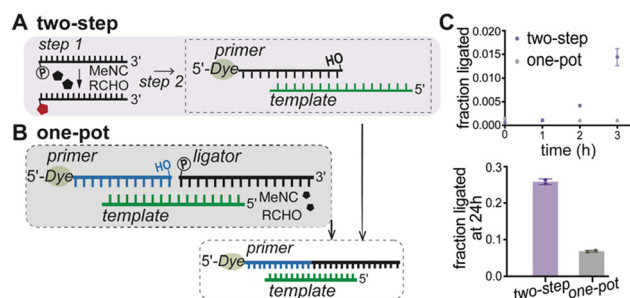


Fig. 2 Nonenzymatic RNA ligation driven by isocyanide activation chemistry. Schematic of nonenzymatic ligation using (A) ligators activated *via* isocyanide activation chemistry prior to the ligation reaction, and (B) ligators activated *in situ* during ligation. (C) Time-course analysis of nonenzymatic ligation with ligators either pre-activated (purple) or activated *in situ* (gray) with isocyanide chemistry. The bottom panel compares ligation yields after 24 hours. Ligation reactions contained 1.2 μM primer, 1.32 μM template, 200 mM HEPES (pH 8), 50 mM Mg^{2+} , and 2.4 μM of either the isocyanide-activated ligator mixture (added after 6 hours of activation) or activated directly within the ligation reaction. Ligation reactions were analyzed by denaturing PAGE (see Fig. S3, SI). The increase in the rate of ligation over time likely results from the gradual increase in the concentration of activated ligators.



would exist between substrate activation and RNA copying, we tested whether oligonucleotide activation and nonenzymatic ligation could proceed in the same pot (Fig. 2B). Upon adding an unactivated ligator to the primer–template complex under activation conditions, only ~7% of the primer was ligated after 24 hours (Fig. 2C and Fig. S3, SI). A previous study has shown that 2AI-activated ligators undergo minimal hydrolysis under similar conditions over similar timescales,¹⁶ suggesting that hydrolysis was not responsible for the reduced yield. Based on previous findings, the reduced ligation yields likely stem from poor activation of the ligator due to steric hindrance at the ligation junction, which makes its 5'-phosphate group less accessible to the activation reagents. If so, a fluctuating environment that allows oligonucleotides to repeatedly bind to and dissociate from templates could lead to increased substrate activation and ligation (see later).

To examine whether the same activation chemistry could also support ribozyme-catalyzed RNA ligation, we began by assessing the compatibility of ligase ribozyme activity with the conditions of chemical activation. We introduced 500 mM MeNC, 250 mM 2MBA, and 250 mM 2AI into a ribozyme-catalyzed ligation reaction using pre-activated, purified ligators in 200 mM HEPES at pH 8 and at 10 mM Mg²⁺. We observed a decrease in ligase activity in the presence of 2AI (Fig. S4, SI), with a marked reduction in ligation yield above 50 mM 2AI (Fig. S5A, SI). To test if 2AI interferes with ligation by chelating Mg²⁺ and consequently reducing its availability for RNA catalysis, we used a fluorescence assay²¹ based on the Mg²⁺-sensitive dye (Mag-fura-2) to quantify free Mg²⁺. The assay revealed a progressive decrease in free Mg²⁺ concentration as 2AI levels increased, with detectable effects at 2AI concentrations as low as 5 mM (Fig. S6, SI). To counteract this effect, we increased the Mg²⁺ concentration from 10 mM to 200 mM in reactions containing 200 mM 2AI. This adjustment restored ribozyme activity, increasing ligation yield from 4.1 ± 0.1% to 33.3 ± 0.1% after 3 hours (Fig. S5B, SI). As this ribozyme reaches catalytic saturation at 4 mM Mg²⁺,²² the observed improvement is likely due to the restoration of the free Mg²⁺ pool in the presence of 2AI.

Next, we integrated the isocyanide activation chemistry with ribozyme-catalyzed ligation (Fig. 3 and Fig. S7, SI). Ligators were activated for 6 hours and then added to the ribozyme primer–template duplex without any purification, yielding 42 ± 1% ligation after 3 hours, which is comparable to yields obtained with purified ligators pre-activated with EDC (Fig. S5A, SI). Extending the reaction time to 24 hours increased the yield to 54 ± 1% (Fig. 3B and Fig. S7A, SI), demonstrating that oligonucleotides activated by potentially prebiotic chemistry can drive efficient ribozyme-catalyzed ligation. We then tested whether substrate activation and ribozyme-catalyzed ligation can proceed in a shared environment, mimicking a prebiotic scenario with no separation between chemical activation and RNA copying by ribozymes. We added the ribozyme primer–template duplex, the unactivated ligator, and activation reagents in one pot (Fig. 3C). We observed a gradual activation of the ligator over the first 3 hours (Fig. 3D and Fig. S7B, SI), and a ligation yield of 23 ± 1% after 24 hours—roughly half the yield obtained with pre-activated ligators (Fig. 3B and D). The ligation time course displayed biphasic kinetics, with a lag phase corresponding to gradual accumulation of activated ligators, followed by an increase in ligation

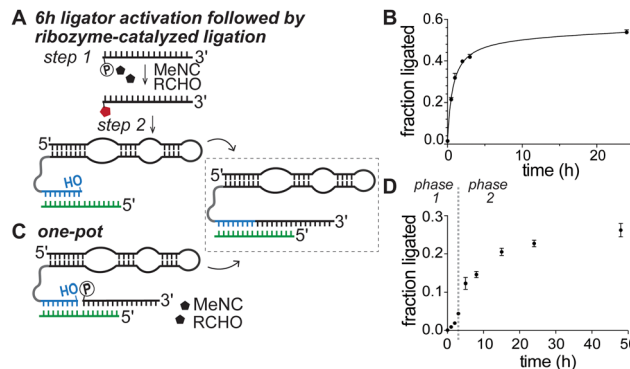


Fig. 3 Isocyanide activation of oligonucleotides drives ribozyme-catalyzed RNA ligation. (A) Ribozyme-catalyzed ligation of pre-activated oligonucleotides. The unactivated ligator was incubated under activation conditions for 6 h and then added to a mixture containing the ribozyme–primer construct and the template without any purification. Ligation reactions contained 1.2 μM ribozyme–primer construct, 1.32 μM template, and 2.4 μM isocyanide-activated ligators in the presence of 200 mM HEPES at pH 8 and 50 mM Mg²⁺. (B) Time course of ribozyme-catalyzed ligation in (A) (see Fig. S7A, SI for the corresponding denaturing PAGE data). (C) *In situ* MeNC activation of oligonucleotides enables ribozyme-catalyzed ligation in one pot. 1.2 μM ribozyme–primer construct and 1.32 μM template were incubated with 2.4 μM unactivated ligators in the presence of 250 mM MeNC, 250 mM 2MBA, 250 mM 2AI, 200 mM HEPES at pH 8, and 50 mM Mg²⁺. (D) Time course of ribozyme-catalyzed ligation in (C) (See Fig. S7B, SI for the corresponding denaturing PAGE data). The plot is biphasic: phase 1 reflects increasing ligation as activated ligators accumulate. In phase 2, ligation proceeds more rapidly with activated ligators and eventually plateaus.

that closely resembled the kinetics observed with pre-activated ligators. To test if reduced ligation was due to inhibition by template-bound unactivated ligators, we incubated the ribozyme–primer construct and template with varying ratios of purified activated and unactivated ligators. We observed that increasing the fraction of unactivated ligators resulted in a marked decrease in ligation, suggesting that unactivated ligators competitively inhibit ligation (Fig. S8, SI). This reduced efficiency of the one-pot ribozyme-catalyzed ligation reaction is comparable to that observed for non-enzymatic ligation (Fig. 2C and 3C).

We hypothesized that activation was inefficient in one-pot reactions due to the limited accessibility of the activation reagents to the 5'-phosphate group of the unactivated ligator, likely sequestered at the ligation junction when bound to the template. To test this possibility, we incubated the ribozyme–primer construct with the unactivated ligator under activation conditions but withheld the template for the first 3 hours (Fig. S9, SI). In the absence of template binding and the formation of a tight ligation junction, the ligator 5'-phosphate group should remain more accessible, allowing more efficient activation. As expected, this approach improved ligation yield to 33.2 ± 0.5% in 24 h. In a prebiotic scenario, where shorter ligators would be more abundant, the rapid dissociation of unactivated ligators from their templates is expected to favor ribozyme-catalyzed ligation of ligators that have been activated off-template. However, weak ligator–template interactions may have a detrimental effect on ribozyme-catalyzed RNA ligation. To test this possibility, we repeated the ligation reaction with templates shortened by 1–3 nucleotides from their 5' end (Fig. S10, SI). Reducing base-pairing



between the template and the ligator from 7 to 5 base pairs had only a modest effect on the ligation yield. Therefore, we suggest that environments with fluctuating temperatures, salt or pH, which allow longer oligos with weak interactions with their templates to dissociate, would favor substrate activation and subsequent ligation by the ribozyme when they are again bound to the template.

Our findings indicate that a high concentration of reagents, especially, isocyanide is crucial for efficient oligonucleotide activation. The continuous (or periodic) photochemical generation of methyl isocyanide upon UV irradiation (Fig. 1) could enhance ligation by re-activating hydrolyzed unactivated ligators.⁴ Such a mechanism could sustain cycles of RNA activation and copying. We performed ribozyme-catalyzed ligation, where 5'-monophosphorylated (*i.e.*, unactivated) ligators, MeNC, 2MBA, and 2AI were added daily over four consecutive days. This setup mimics a prebiotic scenario in which both oligonucleotides and activation reagents are intermittently replenished due to geochemical or photochemical fluctuations. Repeated additions increased ligation yield from 18% on day 1 to 43.7% by day 4 (Fig. S11, SI). Successive additions of activation reagents on the early Earth due to geochemical fluctuations could have enabled efficient oligonucleotide activation, supporting multiple cycles of ribozyme-catalyzed RNA copying.

Recent research has identified a one-pot reaction where 3'-monophosphorylated oligonucleotides are activated by diamidophosphate and imidazole to produce oligonucleotide 2',3'-cyclic phosphates. This activation was shown to facilitate hairpin ribozyme-catalyzed RNA ligation.²³ However, hairpin ribozyme-catalyzed ligations were performed under ice-eutectic conditions to shift the equilibrium in favor of ligation over cleavage. The cyclic phosphate pathway produces ~30% ligation after 30 days, whereas ligation with 5'-phosphorimidazolides reaches similar yields within 24–48 hours. Moreover, ligation with 2',3'-cyclic phosphate ligators is extremely slow without enzymes, whereas 5'-phosphorimidazolides support both nonenzymatic and ribozyme-catalyzed ligation in the same environment. By demonstrating the activation of oligonucleotide substrates and their nonenzymatic and enzymatic ligation under a unified set of conditions, this work offers a more complete experimental model of prebiotic RNA assembly.

While this framework offers insight into how prebiotic activation chemistry could support RNA ligation, several limitations remain. First, although MeNC can be generated under simulated prebiotic conditions, its geochemical availability, stability, and accumulation on the early Earth are not well established. Second, while the relatively high concentrations of MeNC, 2AI, and aldehyde are useful for modeling activation pathways, their prebiotic relevance is uncertain.^{5,6} However, our observation that repeated cycles of reagent and oligonucleotide addition sustain ligation over multiple days suggests that dynamic environments, such as wet-dry cycles, could allow continuous substrate turnover. Finally, we find that template-bound oligonucleotides are less efficiently activated, which implies that environmental fluctuations that promote strand

dissociation and reannealing were critical in prebiotic settings. Such cycles could allow oligonucleotides to be activated in their unbound state and subsequently undergo template-directed ligation under conditions favorable to hybridization.

Ribozymes that facilitate monomer and oligomer activation would make substrate activation and reactivation much more efficient. The integration of prebiotic activation chemistries into model protocells, such as lipid vesicles or coacervates,^{24,25} may reveal how compartmentalization could have influenced early biochemical processes.

S. J. Z., X. J., and S. D. designed the research; S. J. Z. and X. J. performed the research; S. J. Z., X. J., S. D., and J. W. S. analyzed the data; and S. J. Z., S. D., and J. W. S. wrote the paper.

J. W. S. is an Investigator of the Howard Hughes Medical Institute. This work was supported in part by grants from the Simons Foundation (290363) and the NSF (2325198) to J. W. S., and by University of Notre Dame Startup Funds to S. D. S. J. Z. is supported by a postdoctoral fellowship (F32 AG087642) from the National Institutes of Health.

Conflicts of interest

The authors have no conflicts to declare.

Data availability

Supporting data including experimental methods, supporting figures, and tables referenced in the main text have been included as part of the SI. See DOI: <https://doi.org/10.1039/d5cc03937a>

References

- 1 W. Gilbert, *Nature*, 1986, **319**, 618.
- 2 J. W. Szostak, *et al.*, *Nature*, 2001, **409**, 387–390.
- 3 L. J. Zhou, *et al.*, *J. Am. Chem. Soc.*, 2020, **142**, 15961–15965.
- 4 B. J. Weimann, *et al.*, *Science*, 1968, **161**, 387.
- 5 S. J. Zhang, *et al.*, *J. Am. Chem. Soc.*, 2020, **142**, 14810–14813.
- 6 A. Mariani, *et al.*, *J. Am. Chem. Soc.*, 2018, **140**, 8657–8661.
- 7 T. Walton and J. W. Szostak, *Biochemistry*, 2017, **56**, 5739–5747.
- 8 S. J. Zhang, *et al.*, *Proc. Natl. Acad. Sci. U. S. A.*, 2022, **119**, e2116429119.
- 9 D. Ding, *et al.*, *Nucleic Acids Res.*, 2023, **51**, 6528–6539.
- 10 R. Rohatgi, *et al.*, *J. Am. Chem. Soc.*, 1996, **118**, 3340–3344.
- 11 R. Rohatgi, *et al.*, *J. Am. Chem. Soc.*, 1996, **118**, 3332–3339.
- 12 N. Prywes, *et al.*, *eLife*, 2016, **5**, e17756.
- 13 G. F. Joyce, *et al.*, *Proc. Natl. Acad. Sci. U. S. A.*, 1987, **84**, 4398–4402.
- 14 K. D. James and A. D. Ellington, *Chem. Biol.*, 1997, **4**, 595–605.
- 15 J. W. Szostak, *Philos. Trans. R. Soc., B*, 2011, **366**, 2894–2901.
- 16 T. Walton, *et al.*, *Proc. Natl. Acad. Sci. U. S. A.*, 2021, **118**, 5741–5748.
- 17 J. W. Szostak, *J. Syst. Chem.*, 2012, **3**, 2.
- 18 L. J. Zhou, *et al.*, *RNA*, 2021, **27**, 1–11.
- 19 J. W. Szostak, *Angew. Chem., Int. Ed.*, 2017, **56**, 11037–11043.
- 20 Z. R. Todd, *et al.*, *Chem. Commun.*, 2019, **55**, 10388–10391.
- 21 R. Saha, *et al.*, *Nat. Commun.*, 2018, **9**, 2313.
- 22 S. DasGupta, *et al.*, *Chem. Eur. J.*, 2023, **29**, e202301376.
- 23 H. Lin, *et al.*, *Angew. Chem., Int. Ed.*, 2022, **61**, e202113625.
- 24 D. Deamer, *Life*, 2017, **7**, 5.
- 25 A. Agrawal, *et al.*, *Sci. Adv.*, 2024, **10**, eadn9657.

