

# The role of sugar-backbone heterogeneity and chimeras in the simultaneous emergence of RNA and DNA

Subhendu Bhowmik  and Ramanarayanan Krishnamurthy \*

**Hypotheses of the origins of RNA and DNA are generally centred on the prebiotic synthesis of a pristine system (pre-RNA or RNA), which gives rise to its descendent. However, a lack of specificity in the synthesis of genetic polymers would probably result in chimeric sequences; the roles and fate of such sequences are unknown. Here, we show that chimeras, exemplified by mixed threose nucleic acid (TNA)–RNA and RNA–DNA oligonucleotides, preferentially bind to, and act as templates for, homogeneous TNA, RNA and DNA ligands. The chimeric templates can act as a catalyst that mediates the ligation of oligomers to give homogeneous backbone sequences, and the regeneration of the chimeric templates potentiates a scenario for a possible cross-catalytic cycle with amplification. This process provides a proof-of-principle demonstration of a heterogeneity-to-homogeneity scenario and also gives credence to the idea that DNA could appear concurrently with RNA, instead of being its later descendent.**

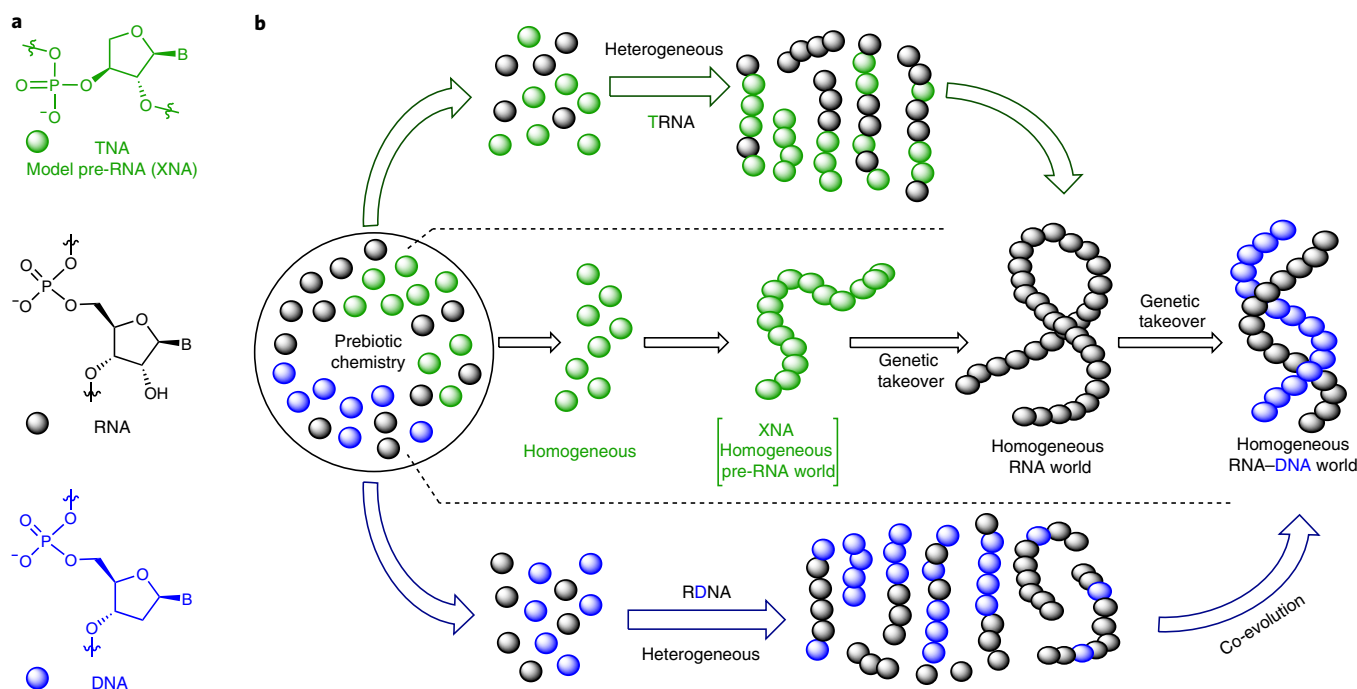
The RNA world hypothesis proposes the emergence of self-replicating and catalytic RNA that later gives rise to proteins and DNA (Fig. 1b, middle)<sup>1,2</sup>. Models posit the existence of a genetic polymer—whether RNA or its precursor—with a homogeneous backbone that transitions to its homogeneous backbone successor<sup>1,3–10</sup>. This transition is proposed to occur despite the difficulties<sup>2,11–14</sup> associated with the generation of the pristine oligomers using prebiotic chemistry<sup>15,16</sup>, and the challenge of replacing one genetic polymer with another<sup>2,17–21</sup> in the absence of any sophisticated discrimination mechanism during the transition in a prebiological world<sup>13,22</sup>. However, there is a growing realization<sup>23–25</sup> that most prebiotic pathways<sup>26,27</sup> would lead to nucleic acid oligomers that consist of mixed backbone units<sup>14,17,19,28</sup>. In this context, RNA that contains a mixture of 2',5' and 3',5' linkages<sup>18,19,29</sup>, and chimeric RNA–DNA systems<sup>17,21</sup>, have been investigated (and it was shown that these types of backbone heterogeneity compromise aptamer function<sup>17–19</sup>), and we have shown that RNA–DNA chimeras consistently form weaker duplexes<sup>14</sup>. Although chimeric RNA–DNA genomes are known in extant biology<sup>30</sup> and such chimeras containing non-heritable backbone heterogeneity were postulated to be useful in the emergence of functional nucleic acids<sup>17,19</sup>, questions were raised about their role as enhanced templates for replication<sup>17,31</sup> to generate polymers with homogeneous backbones<sup>14</sup>. For pre-RNA to RNA transitions, Orgel has speculated two extreme possibilities using threose nucleic acid (TNA) (Fig. 1a)<sup>32</sup> as an example: (1) an all-TNA organism that converts to an all-RNA organism and (2) a gradual replacement of TNA residues by RNA residues within the oligomeric system<sup>33</sup>. The second scenario leads to a continuous pathway from TNA to RNA via chimeric sequences<sup>33</sup>. We proposed a heterogeneity-to-homogeneity scenario<sup>34</sup> for the emergence of RNA and DNA<sup>13,14</sup>, and argued that, based on certain criteria such as the stability and functional advantages inherent to homogeneous backbone polymers, their emergence would be a natural consequence even when starting from a mixture of its constituent building blocks (Fig. 1b, top and bottom)<sup>13</sup>. A demonstration that chimeric TNA–RNA (TRNA) sequences (Fig. 1b, top) or RNA–DNA

(RDNA) sequences (Fig. 1b, bottom) can enable the non-enzymatic emergence of a homogeneous backbone oligonucleotide (RNA or DNA) starting from mixtures of chimeric sequences would provide support to the heterogeneity-to-homogeneity scenario<sup>13</sup>.

## Results

**TRNA chimeric sequences function as templates for RNA ligands.** We selected TNA<sup>32</sup>—a Watson–Crick base-pairing system able to cross pair with RNA<sup>32,35</sup>—as a model pre-RNA polymer<sup>13</sup>, based on the prebiotic availability of the sugars<sup>27,36–39</sup> (Fig. 1a). We investigated TRNA chimeric sequences that exhibited peculiar base-pairing properties even though TNA formed strong and stable duplexes with complementary RNA strands (Supplementary Tables 1 and 2)<sup>32</sup>. First, in general, TRNA formed weaker duplexes compared to the unmodified strands. Second, based on which sugar (threose or ribose) unit contained a purine (A) or pyrimidine (T), TRNA demonstrated unpredictable duplex stabilities (Fig. 2a). Unexpectedly, TRNA non-self-complementary strands that showed a weak affinity for each other (Fig. 2a, entry 7) formed stronger duplexes with the corresponding complementary RNA (or TNA) sequences (Fig. 2a, entries 6 and 8), a behaviour that was general for sequences that contained all four nucleobases (Supplementary Table 3 and Supplementary Figs. 7–13).

The preferential association of chimeric TRNA sequences with homogeneous RNA (or TNA) sequences (Fig. 2a, entries 6 and 8) implied that chimeric sequences could act selectively as templates for the non-enzymatic ligation of homogeneous sugar backbone ligands, and thereby facilitate the emergence of a homogeneous backbone oligomer (for example, RNA) starting from a mixture of oligonucleotides. To test this proof-of-concept, we employed the widely used water soluble 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC)-mediated ligation conditions<sup>40</sup> for homogeneous RNA ligands templated by TRNA chimeric and RNA templates, and compared it with ligation of the chimeric TRNA ligands (Fig. 2b). The 3'-NH<sub>2</sub>-modified TNA ligand<sup>41</sup> and 3'-NH<sub>2</sub>-deoxynucleotide (7<sup>NH<sub>2</sub></sup>)-terminated RNA ligand<sup>42</sup> were used to conduct the ligation

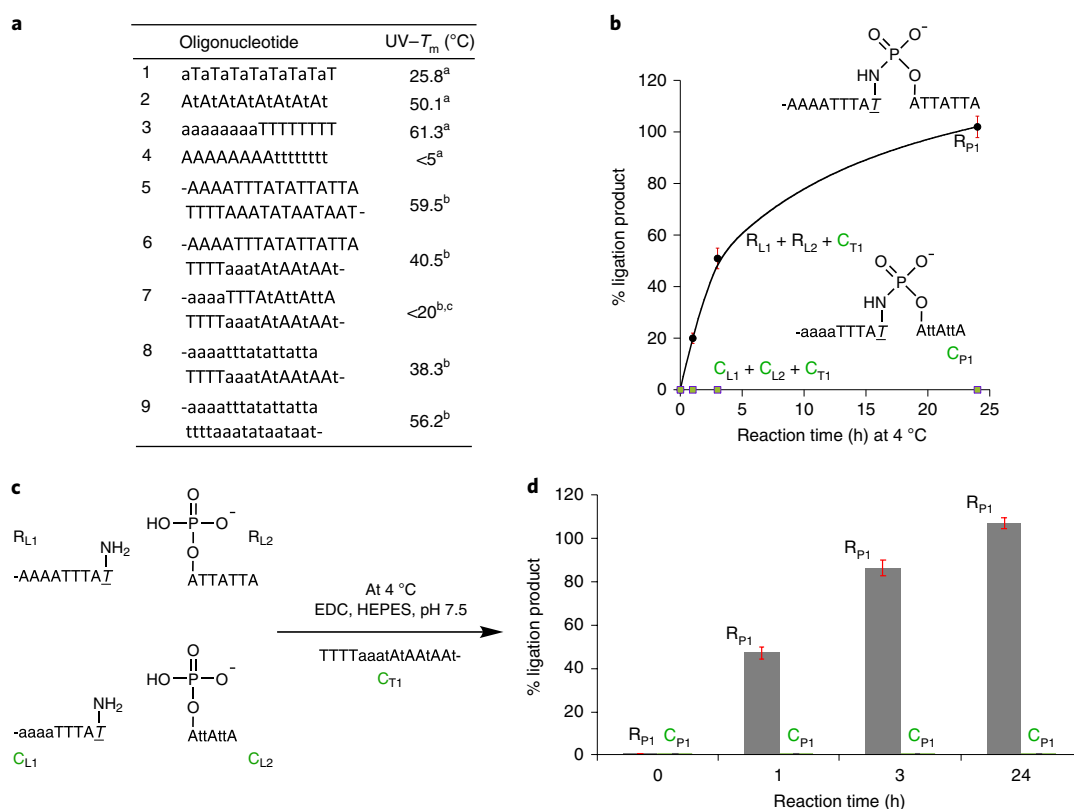


**Fig. 1 | The prebiotic clutter generated heterogeneity-to-homogeneity scenario versus the biology-inspired model of replacing one homogeneous genetic system with its homogeneous genetic successor. a**, Constitutional formula representation of the three oligonucleotide building blocks investigated in this study. **b**, Three possible scenarios for the emergence of RNA and DNA from prebiotic chemistry. Middle: the classical RNA world concept in which the formation of a pristine and homogeneous RNA (or pre-RNA) leads to its homogeneous backbone successor DNA (or RNA). Top: a heterogeneous mixture of TNA (pre-RNA) and RNA that forms chimeric TRNA sequences that transition to homogenous RNA, which then gives rise to DNA. Bottom: a heterogeneous RNA–DNA mixture that progresses and/or co-evolves via chimeric RDNA sequences directly to homogeneous RNA and DNA simultaneously.

reaction within a reasonable time frame, as the corresponding TNA-3'-OH and RNA-3'-OH residues react very slowly (Supplementary Figs. 14–17). The single phosphoramidate linkage was shown to have no special effect on the duplex stability (Supplementary Fig. 12). The reactions were monitored by anion-exchange chromatography and the products were confirmed by comparison with standards and matrix-assisted laser desorption/ionization-time-of-flight (MALDI-TOF) mass spectrometry (Supplementary Figs. 18–28). As expected from a previous study<sup>41</sup>, the efficiency and the rate of ligation reactions paralleled the affinity (and thermal stability) of the templates for the ligands in the order: RNA template with TNA ligands  $\approx$  RNA template with RNA ligands  $\geq$  TRNA chimeric template with RNA ligands  $\gg$  RNA template with TRNA chimeric ligands  $\gg \gg \gg$  TRNA chimeric template with TRNA chimeric ligands (Supplementary Figs. 18–23). Control reactions that lack the template(s) showed no product formation (Supplementary Figs. 25–28). We then examined the ligation behaviour of the mixture of all four ligands in the presence of the chimeric TRNA template (Fig. 2c) and observed, by anion-exchange chromatography, only the formation and growth of the RNA product from homogeneous RNA ligands, with no discernible chimeric TRNA product from heterogeneous TRNA ligands (Fig. 2d and Supplementary Fig. 24). However, MALDI-TOF analysis of the reaction of chimeric TRNA ligands with the chimeric TRNA template at 24 hours did show traces of the chimeric TRNA product (Supplementary Fig. 18). We did not investigate intensively a parallel scenario for the emergence of homogeneous TNA sequences<sup>43</sup> (due to the investment in synthesizing the various TNA 3'-NH<sub>2</sub>-phosphoramidites), although we expect a similar propensity<sup>32</sup> based on the observation that homogeneous TNA ligands were also preferentially ligated by the chimeric TRNA template (Supplementary Fig. 20).

**RDNA chimeric templates ligate complementary RNA and DNA ligands.** The above results inspired us to investigate mixed DNA and RNA chimeric sequences based on (1) our previous studies of RDNA chimeras<sup>44</sup> and the plausible coexistence and co-evolution of RNA and DNA in prebiotic scenarios<sup>17,21,28,44</sup> and (2) the ease of commercial and synthetic availability of diverse RDNA chimeric sequences. We studied a series of RDNA chimeric sequences (Supplementary Table 4), which, again, formed stronger duplexes with complementary homogeneous RNA over the corresponding complementary chimeric RDNA (Supplementary Table 5 and Supplementary Figs. 29–35). To test whether the preferential association of RDNA with RNA would also translate to the selective ligation of RNA ligands (as seen in the TRNA system), we investigated the ligation behaviour of a hexadecamer chimeric RDNA template ( $C_{T2}$  in Fig. 3a) with RNA and RDNA ligands that contained 3'-NH<sub>2</sub> deoxynucleotide units. The ligation of RNA sequences ( $R_{L3}$  and  $R_{L4}$ ) on the chimeric RDNA template ( $C_{T2}$ ) was not only faster than the corresponding ligation of the chimeric RDNA ligands ( $C_{L3}$  and  $C_{L4}$  on  $C_{T2}$  (Fig. 3b)), but was almost equal to the efficiency of RNA ligands  $R_{L3}$  and  $R_{L4}$  (or chimeric  $C_{L3}$  and  $C_{L4}$  ligands (Supplementary Fig. 39)) on an RNA template,  $R_{T2}$  (Supplementary Figs. 36–46).

The duplex formation in octameric homogeneous and chimeric sequences that contain all five canonical nucleosides again showed a preferential association of the homogeneous backbone sequences with complementary chimeric templates (Supplementary Table 5). Based on this, we investigated the ligation reaction mediated by the chimeric template  $C_{T4}$  with RNA and the chimeric ligands shown in Fig. 3c. The results revealed a temperature-dependent ligation behaviour that was not observed in the hexadecameric AU system (Supplementary Figs. 50–52). Although at lower temperatures (4°C) there was little difference between the rate of ligation



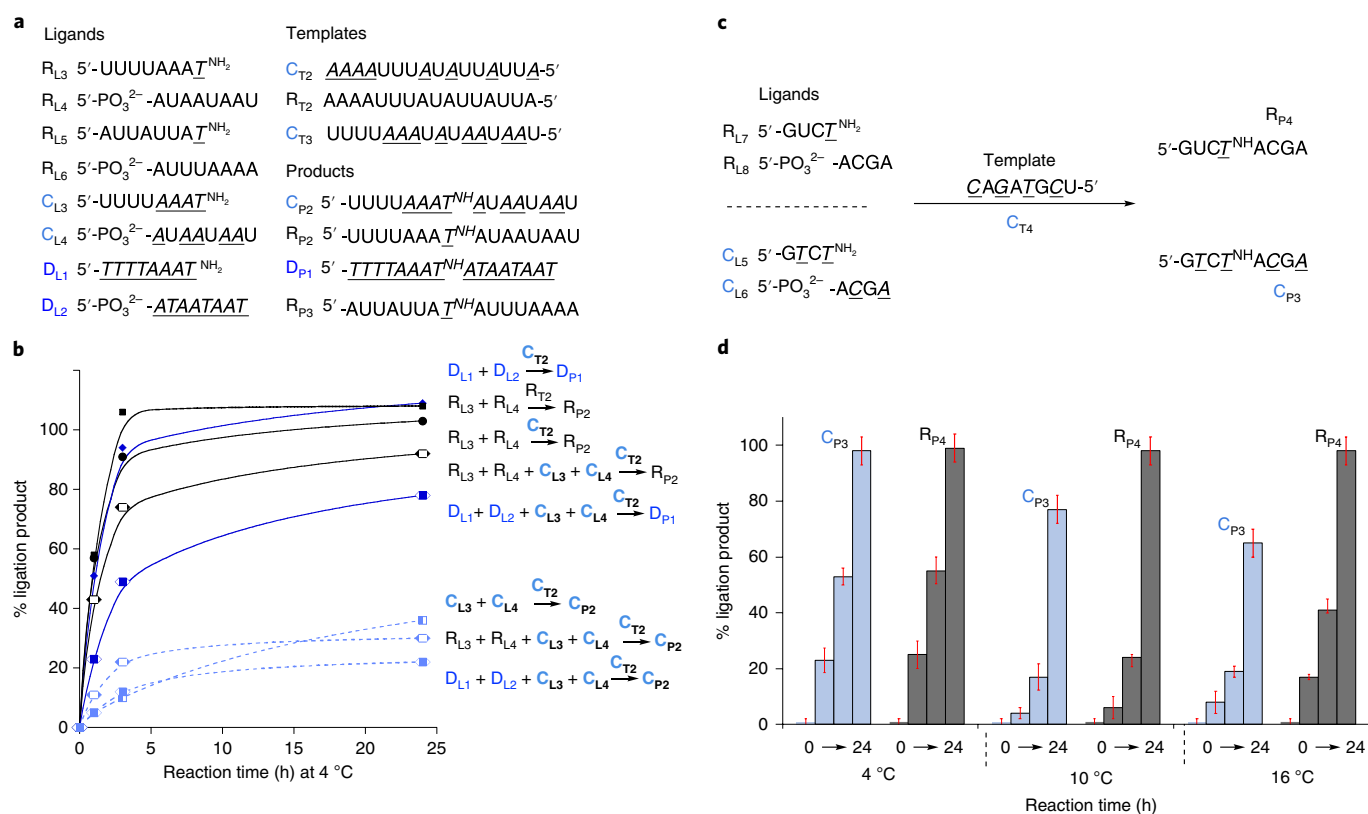
**Fig. 2 | The preferential association with, and ligation of homogeneous ligands by, a chimeric TRNA template over chimeric ligands.** **a**, Thermal stability of TRNA chimeric duplexes in 1 M NaCl, 10 mM  $\text{Na}_2\text{HPO}_4$ , 100  $\mu\text{M}$  EDTA, pH 7.2. <sup>a</sup>5  $\mu\text{M}$  duplex concentration. <sup>b</sup>2  $\mu\text{M}$  duplex concentration. <sup>c</sup>Entry 7, no clear sigmoidal transition in the ultraviolet-thermal melt (UV- $T_m$ ) was observed. **b**, Comparison of the rate of ligation reaction at 4  $^{\circ}\text{C}$  of homogeneous RNA ( $R_{L1}$  and  $R_{L2}$ ) and heterogeneous TRNA ( $C_{L1}$  and  $C_{L2}$ ) ligands on a heterogeneous TRNA template,  $C_{T1}$ . **c**, EDC-mediated ligation reaction at 4  $^{\circ}\text{C}$  of a mixture of homogeneous RNA ( $R_{L1}$  and  $R_{L2}$ ) and chimeric TNA-RNA ( $C_{L1}$  and  $C_{L2}$ ) ligands using a TRNA chimeric sequence ( $C_{T1}$ ) as the template. **d**, Comparison of the amounts of products  $R_{P1}$  and  $C_{P1}$  produced in the reaction mixture in **c** (Supplementary Fig. 24 gives the conditions). A, T = RNA,  $\underline{\text{T}}$  = DNA; a, t = TNA. The line in **b** is drawn as a guide to indicate the trend and is not a mathematical curve fitting. Percentage yields are calculated with respect to the template  $C_{T1}$ . Experiments were run in triplicate and the error range was less than  $\pm 5\%$ ; error bars represent s.d.

between the two systems, the rate of ligation of chimeric ligands and the amounts of products formed at higher temperatures (10 and 16  $^{\circ}\text{C}$ ) differed considerably with preference for the ligation product from homogeneous ligands on the chimeric template (Fig. 3d and Supplementary Fig. 52). This indicates that temperature could also control and modulate the overall dynamics and distribution of the end-products.

The trend of preferential association correlating with the ligation capacity of  $C_{T2}$  also extended to DNA ligands ( $D_{L1}$ ,  $D_{L2}$ ), in place of RNA ligands, which gives rise to the homogeneous DNA product  $D_{P1}$  (Fig. 3b), and was valid even when starting from a pool of mixed  $R_{L3}+R_{L4}+C_{L3}+C_{L4}$  ligands or  $D_{L1}+D_{L2}+C_{L3}+C_{L4}$  ligands (Fig. 3b and Supplementary Figs. 47–49). When all the ligands ( $R_{L3}$ ,  $R_{L4}$ ,  $D_{L1}$ ,  $D_{L2}$ ,  $C_{L3}$  and  $C_{L4}$ ) were added to the chimeric RDNA template  $C_{T2}$  in a single pot, three major ligation products,  $R_{P2}$  (38%),  $D_{P1}$  (20%) and an RNA–DNA cross-ligation product ( $RD_{P1}$ , 75%) were formed at 24 hours; no chimeric product from  $C_{L3}+C_{L4}$  was detected (Supplementary Figs. 53 and 54). The nature of the cross-ligation product was confirmed with appropriate control experiments and shown to be the result of  $D_{L1}$ – $R_{L4}$  ligation (Supplementary Figs. 55–59). Replacing the chimeric template with an RNA template, under otherwise identical conditions, gave  $R_{P2}$  (65%),  $D_{P1}$  (12%) and 62% of  $RD_{P1}$  and  $RD_{P2}$  ( $R_{L3}$ – $D_{L2}$ ), which indicates that the RNA template also gave rise to significant cross-ligation products (Supplementary Figs. 60 and 61). Changing the ratios of the RNA ligands ( $R_{L3}+R_{L4}$ ) to the DNA ligands ( $D_{L1}+D_{L2}$ ) affected the product distribution

(Supplementary Fig. 59), which implies that the generation of chimeric oligomers (along with homogeneous backbone oligomeric products) has to be reckoned with; these chimeric oligomer products should, in turn, help in the formation of homogeneous RNA and DNA ligation products. Although this hypothesis is reinforced by the results in Fig. 3, it was demonstrated to be so by isolating  $RD_{P1}$  and using it as a template with RNA ligands to produce  $R_{P3}$  efficiently in a 108% yield (Supplementary Fig. 62). The above results show that from a mixed system with two different oligonucleotides (for example, RDNA) there is, indeed, the possibility of the simultaneous emergence of the two respective homogeneous nucleotide polymers (for example, RNA and DNA).

**RDNA chimeric templates are better in overcoming template-product inhibition.** The above observations suggest that chimeric templates could provide a solution to the problem of product inhibition (Fig. 4a), in which the continuous production of the product is curtailed due to the strong association of the initially formed template–product complex<sup>45–48</sup>. For instance, RNA ligands  $R_{L3}$  and  $R_{L4}$  in the presence of the  $R_{P2}$ – $R_{P3}$  RNA duplex under EDC-activation conditions showed no production of  $R_{P2}$ , even after 24 hours, indicative of a classic product inhibition behaviour; but the addition of the chimeric template  $C_{T2}$  led to the formation of more  $R_{P2}$  within a matter of a few hours (Supplementary Figs. 67 and 68). As outlined in Fig. 4b, if there was the adventitious presence and/or formation of a complementary RNA partner ( $R_{P3}$ , from its corresponding ligands

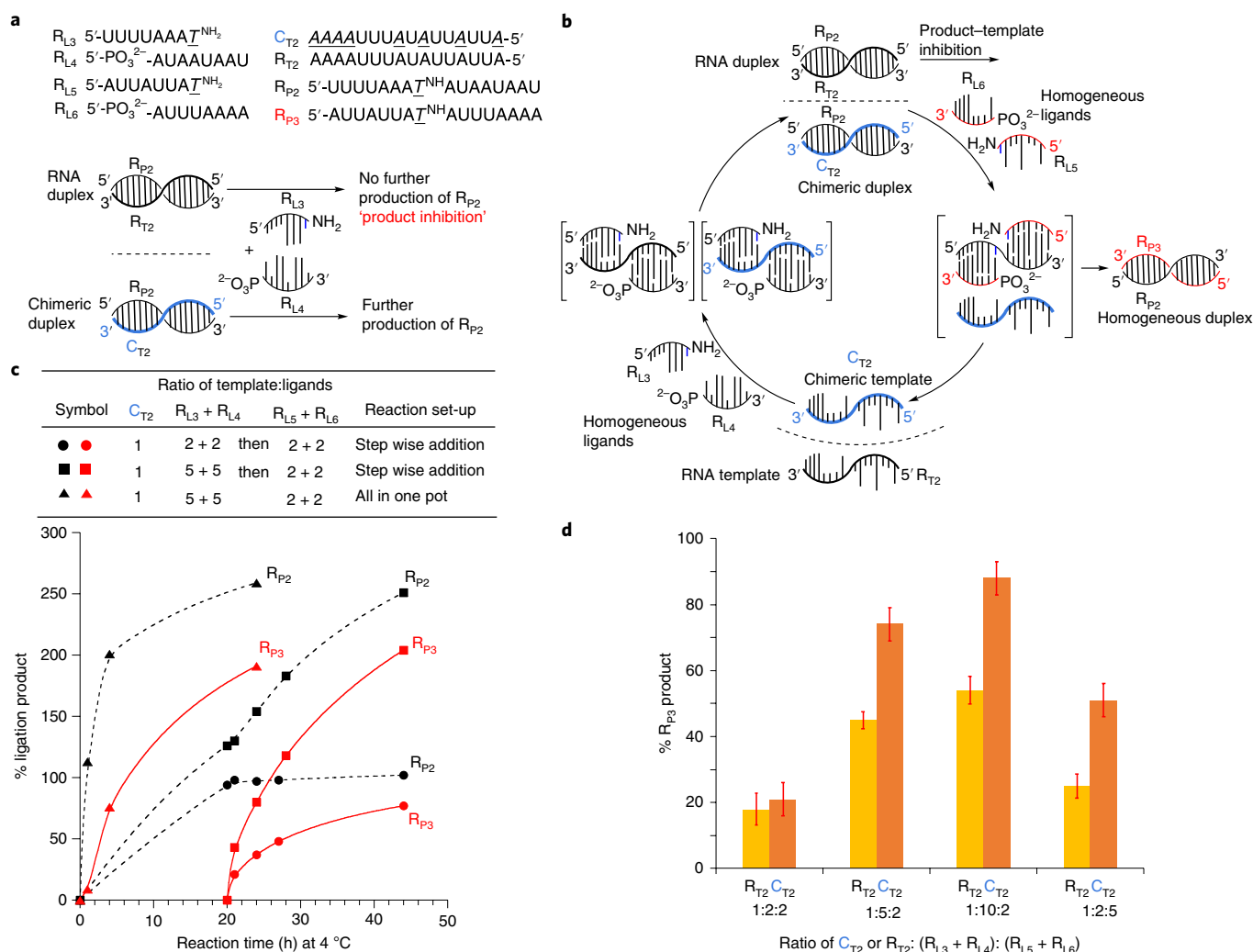


**Fig. 3 | Chimeric RDNA templates preferentially associate and ligate homogeneous RNA and DNA ligands over chimeric ligands.** **a**, A list of the homogeneous and chimeric sequences used in this study. **b**, Comparison of ligation efficiency by the hexadecameric (AU)-RDNA template  $C_{T2}$  with RNA ( $R_{L3}$ ,  $R_{L4}$ ), DNA ( $D_{L1}$ ,  $D_{L2}$ ) and chimeric RDNA ( $C_{L3}$ ,  $C_{L4}$ ) ligands shows the consistent preferential formation of homogeneous ligation products,  $R_{P2}$  and  $D_{P1}$ , over the chimeric ligation products  $C_{P2}$ . **c,d**, Comparison of ligation efficiency by the octameric (A, U/T, G, C)-RDNA template  $C_{T4}$  with RNA ( $R_{L7}$ ,  $R_{L8}$ ) and chimeric RDNA ( $C_{L5}$ ,  $C_{L6}$ ) ligands (**c**) shows the influence of temperature on the preferential formation of homogeneous ligation products,  $R_{P4}$ , over the chimeric ligation products,  $C_{P3}$  (**d**) (Supplementary Figs. 50–52 give the EDC-ligation reaction conditions). A, U, G, C = RNA;  $\overline{A}$ ,  $\overline{T}$ ,  $\overline{G}$ ,  $\overline{C}$  = DNA. Lines in **b** are drawn as a guide to indicate the trend and are not mathematical curve fittings. Percentage yields were calculated with respect to the template  $C_{T2}$  or  $R_{T2}$  or  $C_{T4}$ . Experiments were run in triplicate and the error range was less than  $\pm 5\%$ ; error bars represent s.d.

$R_{L5}$  and  $R_{L6}$  in the mixture that contained the chimeric duplex ( $R_{P2}$ - $C_{T2}$ ), it would induce the formation of the stronger RNA ( $R_{P2}$ - $R_{P3}$ ) duplex. This should release the original chimeric RDNA template for another round of ligation of  $R_{L3}$  and  $R_{L4}$  to form more  $R_{P2}$  and result in a continuous accumulation of the duplex  $R_{P2}$ - $R_{P3}$ , with the chimeric template  $C_{T2}$  taking the role of a catalyst that produces more  $R_{P2}$  from its respective ligands. To test this scenario, we first conducted a stepwise addition of RNA ligands  $R_{L3}$  and  $R_{L4}$  to the RDNA chimeric template  $C_{T2}$ , which led to the formation of the product  $R_{P2}$  (97% in 20 hours (Fig. 4c)). Then, ligands  $R_{L5}$  and  $R_{L6}$  were added to this mixture. The formation of the second ligation product  $R_{P3}$  (21% in 1 h increasing to 77% in 24 h (Fig. 4c)), indicated that the in situ generated first ligation product  $R_{P2}$  was, indeed, acting as a template (Supplementary Figs. 64–66). More encouragingly, with higher ligand ratios of  $R_{L3}$  and  $R_{L4}$ , an increased amount of the first ligation product  $R_{P2}$  (251% with respect to  $C_{T2}$ ) and of the second ligation product  $R_{P3}$  (204%) was observed after 24 hours (Fig. 4c, Supplementary Fig. 65). This indicates that the chimeric template  $C_{T2}$  was, indeed, being released to take part in a turnover, which in turn led to the formation of more  $R_{P2}$ . Pertinent control experiments confirmed the need for all the components to be present for this system to operate; importantly,  $C_{T2}$  itself did not serve as a template to ligate  $R_{L5}$  and  $R_{L6}$  and did not produce  $R_{P3}$  (Supplementary Figs. 45 and 66). Encouraged by these results, we set up a one-pot experiment in which all the components,  $C_{T2}$ + $R_{L3}$ + $R_{L4}$ + $R_{L5}$ + $R_{L6}$ , were mixed from the beginning and

observed the concomitant production of the two RNA ligation products  $R_{P2}$  and  $R_{P3}$  (as efficiently as the stepwise addition experiment) (Fig. 4c). The presence of the chimeric template  $C_{T2}$  in a mixed one-pot system not only initiated the ligation process, but also acted as a turnover intermediary downstream, which potentially enables the continuous production of  $R_{P2}$  and  $R_{P3}$  by mitigating the inhibition by the template-product complex. This process was mainly driven by the preference of a thermodynamically stable homogeneous backbone duplex  $R_{P2}$ - $R_{P3}$ . Control reactions for  $R_{L3}$ + $R_{L4}$  or  $R_{L5}$ + $R_{L6}$  ligands without the  $C_{T2}$  template showed no observable background ligation reactions. However, when all four ligands (absent  $C_{T2}$ ) were mixed together, 33%  $R_{P2}$ , 20%  $R_{P3}$  and 13% cross-ligation products (probably from  $R_{L3}$ + $R_{L6}$  and/or  $R_{L5}$ + $R_{L4}$ ) were formed, but more slowly at 24 h (Supplementary Fig. 70), as opposed to 259% of  $R_{P2}$  and 191% of  $R_{P3}$  with no cross-ligation products in the presence of the chimeric template  $C_{T2}$  (Supplementary Fig. 69). The background ligation reactions were eliminated when ligand concentrations were lowered from 200  $\mu$ M to 20  $\mu$ M each; and only in the presence of 10  $\mu$ M chimeric template  $C_{T2}$  was the formation of  $R_{P2}$  (83%) and  $R_{P3}$  (18%) in 24 hours was observed (Supplementary Figs. 73–76). Furthermore, we tested whether the presence of the complementary ligands ( $C_{L3}$ + $C_{L4}$  (Fig. 3a)), which led to the  $C_{T2}$ - $C_{P2}$  duplex would prevent further copying of the first two RNA ligands  $R_{L4}$ + $R_{L3}$  and also impact the next round of copying when all four RNA ligands  $R_{L4}$ + $R_{L3}$ + $R_{L5}$ + $R_{L6}$  are present. In both cases, in 24 hours at 4 °C, corresponding RNA products formed in good yields: 92% of





**Fig. 4 | The beneficial role of the chimeric RDNA template in overcoming the template-product inhibition based on the thermodynamic stability of the duplexes. a**, The expected difference between the chimeric RDNA-RNA duplex and the homogeneous RNA-RNA duplex in being able to overcome the template-product inhibition. **b**, Schematic representation of the proposal that the hexadecameric (AU)-RDNA template  $C_{T2}$  with RNA ligands  $R_{L3}$  +  $R_{L4}$  produces  $R_{P2}$ , which, in the presence of  $R_{L5}$  and  $R_{L6}$ , is expected to lead to  $R_{P3}$ , based on the greater thermodynamic stability of the  $R_{P2}$ - $R_{P3}$  duplex over the  $R_{P2}$ - $C_{T2}$  duplex, and release the  $C_{T2}$  for another round of ligation reaction. **c**, Time course of the EDC-mediated ligation experiments that documents the effect of the change in ratio of the ligands, and the sequential addition of ligands  $R_{L5}$  and  $R_{L6}$  (0 h) followed by  $R_{L5}$  and  $R_{L6}$  (at 20 h) versus the all-in-one-pot reaction on the production of  $R_{P2}$  and  $R_{P3}$ . **d**, Comparison of the amount of  $R_{P3}$  formed by the homogeneous RNA template  $R_{T2}$  versus the chimeric RDNA template  $C_{T2}$  (at 48 h) demonstrates the higher efficiency of  $C_{T2}$  in mediating the formation of  $R_{P3}$  by overcoming the template-product inhibition (Supplementary Figs. 63–78 give the EDC-ligation conditions). A, U = RNA; A, T = DNA. Lines in **c** are drawn as a guide to indicate the trend and are not mathematical curve fittings. Percentage yields were calculated with respect to the template  $C_{T2}$  or  $R_{T2}$ . Experiments were run in triplicate and the error range was less than  $\pm 5\%$ ; error bars represent s.d.

$R_{P2}$  (with 30% of  $C_{P2}$ ) for the first experiment and in the second scenario, 83%  $R_{P2}$  and 16%  $R_{P3}$ , with no discernible peak for  $C_{P2}$  in the chromatogram trace (Supplementary Figs. 47 and 77).

To assess the efficiency of the chimeric template  $C_{T2}$  versus that of the corresponding homogeneous backbone RNA counterpart  $R_{T2}$ , the all-in-one-pot reaction was repeated, but with the RNA template  $R_{T2}$  in place of  $C_{T2}$ . In this case, as expected, the production of  $R_{P2}$  at 48 hours was comparable (99% for  $R_{T2}$  versus 109% for  $C_{T2}$ ); however,  $R_{P3}$  formation dropped by almost half to 18% (for  $R_{T2}$ ) when compared to 30% (for  $C_{T2}$ ), which indicates that the template-product inhibition by the stronger  $R_{T2}$ - $R_{P2}$  complex meant that less  $R_{P2}$  was available to ligate  $R_{L5}$ + $R_{L6}$  (Fig. 4d). The advantage of  $C_{T2}$  over  $R_{T2}$  was more apparent when the ratio of the ligands was changed to 5( $R_{L3}$ + $R_{L4}$ ):2( $R_{L5}$ + $R_{L6}$ ) with  $C_{T2}$  producing 178% of  $R_{P2}$  and 77% of  $R_{P3}$  when compared to 119% of  $R_{P2}$  and 43% of  $R_{P3}$  with

the RNA template  $R_{T2}$  (Fig. 4d). This strongly suggests that  $C_{T2}$  is better able to dissociate from the  $C_{T2}$ - $R_{P2}$  template-product complex, whereas the RNA template  $R_{T2}$  is limited by the classic  $R_{T2}$ - $R_{P2}$  template-product inhibition and is, therefore, unable to recycle to produce more  $R_{P2}$  and  $R_{P3}$ . In fact,  $C_{T2}$  consistently outperformed  $R_{T2}$  in the production of  $R_{P3}$  for all other combinations of ligand ratios (Fig. 4d and Supplementary Fig. 78), indicative of the beneficial role played by chimeric templates in moving towards the emergence of homogeneous backbone sequences. However, for this to be possible, this phenomenon must hold good for other strands in terms of length and sequence diversity. Given the limitations imposed by the EDC-ligation chemistries and analysis of the chimeric sequences involved, we set up a proof-of-principle experiment as in Fig. 4b, but with octameric AUGC that contained the chimeric template  $C_{T4}$  (Supplementary Fig. 79), as it also showed a preference for the

complementary homogeneous ligands over the chimeric counterparts, as seen in Fig. 3d. As expected, the chimeric template  $C_{T4}$  was efficient in producing the homogeneous products  $R_{P4}$  and  $R_{P5}$  (Supplementary Fig. 80), overcoming the template–product inhibition even in the presence of all four ligands ( $R_{L7}+R_{L8}+R_{L9}+R_{L10}$ ), which parallels the observations for the AU-based system. Thus, the ability of the chimeric template to give rise to homogeneous backbones (the heterogeneity-to-homogeneity paradigm) seems to be still operative in this RDNA chimeric system even when shortening the length of the template and expanding the sequence diversity.

We then examined the effect of stepwise dilution (as a selection pressure) on the efficiency of the templates in overcoming the template–product inhibition, asking the question—which of the templates, the chimeric RDNA or the homogeneous RNA would produce the ligation products more efficiently as the stepwise dilution was continued? Using the AU system outlined in Fig. 4a, we conducted a stepwise dilution experiment in parallel with templates  $C_{T2}$  and  $R_{T2}$  that contain the complementary RNA ligands ( $R_{L3}$ ,  $R_{L4}$ ,  $R_{L5}$  and  $R_{L6}$ ) in which a portion of the reaction mixture was removed and fresh ligands and EDC were added every 24 hours, such that the concentrations of the ligands remained constant, but the template concentration decreased with each dilution step (Supplementary Figs. 91–93). As seen from Fig. 5a, as the stepwise dilution was implemented at 24-hour intervals, the formation of  $R_{P2}$  and  $R_{P3}$  was observed in both cases; although there was a concomitant drop in the product concentration (by 2  $\mu$ M) at each dilution step, the amount of  $R_{P2}$  and  $R_{P3}$  increased to a level greater than the previous value with the progress of time. The amount of the first ligation product  $R_{P2}$  was almost the same between the chimeric ( $C_{T2}$ ) and homogeneous ( $R_{T2}$ ) template-containing vials over the first two steps (48 hours) of dilution, with  $C_{T2}$  performing slightly better than  $R_{T2}$  as the dilution steps were continued (72–96 hours (Fig. 5b)). However, there was a remarkable difference in the production of the second ligation product  $R_{P3}$  with increasing stepwise dilutions; the chimeric template  $C_{T2}$  outperformed the homogeneous template  $R_{T2}$  in producing  $R_{P3}$  by ~250% (Fig. 5c), even as the concentrations of the templates were decreasing with each step of dilution. A comparison of the chromatogram traces at 96 hours (Fig. 5d) shows the dramatic difference and highlights the ability of  $C_{T2}$  to be a superior template<sup>17</sup> for the production of the homogeneous product  $R_{P3}$ , which demonstrates the ability of the chimeric template  $C_{T2}$  to better by-pass the template–product inhibition and turnover even under dilute conditions when compared to  $R_{T2}$ . Appropriate controls without the template showed no product formation (Fig. 5d).

**RDNA chimeric templates harbour the potential for cross-catalytic self-replication.** The promise of the turnover of RNA ligation (Fig. 4b) when coupled with the observation that RDNA ( $C_{P2}$ ) chimeric products can also be formed on the RNA template (Supplementary Fig. 39) suggested that the catalytic chimeric template ( $C_{T2}$ ) could also be regenerated in the same reaction mixture if the corresponding chimeric ligands ( $C_{L7}+C_{L8}$ ) are present (Fig. 6). If this is possible, then the regeneration of the catalytic template  $C_{T2}$  could allow for a cross-catalytic cycle to be operative, which would be expected to lead to the amplification of the homogeneous RNA product  $R_{P2}$  (Fig. 6b). To test this possibility, we set up a one-pot EDC-ligation reaction with the RNA ligands  $R_{L3}+R_{L4}$  along with chimeric ligands  $C_{L7}+C_{L8}$  in the presence of the chimeric template  $C_{T2}$  (Fig. 6). We observed within 1–4 hours the formation of the expected product  $R_{P2}$  (90%), which could now act as the template for the chimeric ligands  $C_{L7}+C_{L8}$ . Indeed, by 24 hours, the formation of the dT-phosphoramidate-linked equivalent of  $C_{T2}$  ( $C_{T2}^{NH}$ , 16%) was clearly observed, and kept increasing with time to 36% in 48 hours and to 48% in 72 hours. In parallel, the amount of  $R_{P2}$  increased accordingly to 125% in 24 hours, to 148% in 48 hours and to 160% in 72 hours (Supplementary Fig. 94). This is well above the levels of

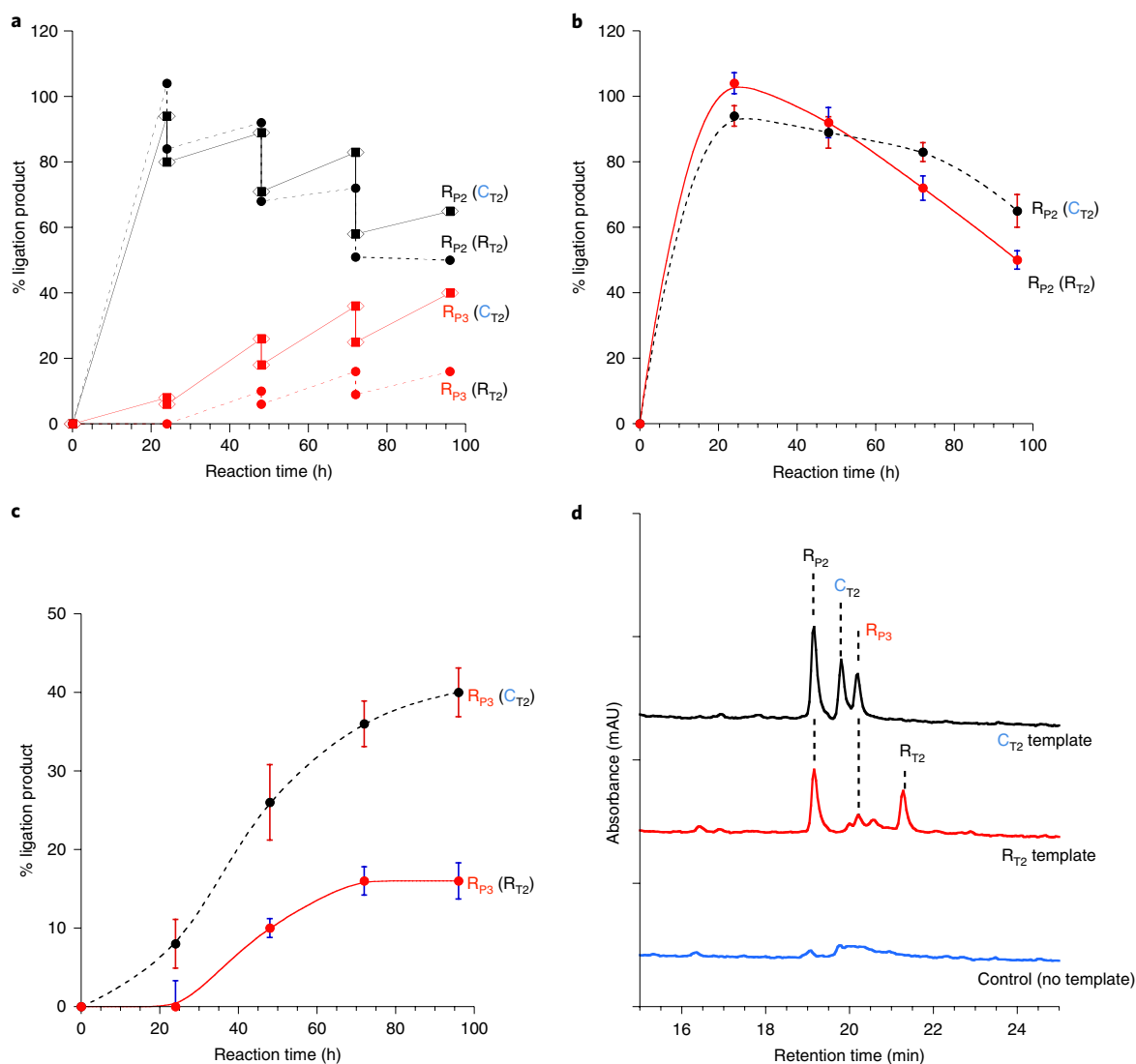
$R_{P2}$  produced in the ligation reaction mediated by  $C_{T2}$  in the presence of only  $R_{L4}+R_{L3}$  and lacking the chimeric ligands (Fig. 6c), in which the amount of  $R_{P2}$  levelled at around 108% by 72 hours. Thus, the chimeric template mediated ligation process shows potential for cross-catalytic self-replicating systems that can result in amplification of the downstream product. Further systematic investigations are ongoing to understand the scope and limitation of this system. In all the experiments described in this work no discernible degradation of the homogeneous or chimeric templates or products was observed (confirmed by comparison with an external standard of oligonucleotide dT<sub>24</sub> added to the samples just before analysis).

## Discussion

The results described in this work confirm experimentally the beneficial roles of chimeric sequences (backbone heterogeneity) in nucleic acid replication, which augments the evolution of functionality in mosaic nucleic acids<sup>17</sup>; they also suggest that the nucleobase sequence information encoded in heterogeneous backbones can, indeed, be heritable for chemical evolution (similar to homogeneous backbone systems). In these chimeric systems, there is the added advantage of (1) by-passing the template–product inhibition problem commonly encountered in the non-enzymatic replication of nucleic acids (unlike the homogeneous backbone systems) and (2) moving towards the (cross-catalytic) self-replication of the chimeric templates, which eventually are able to assist in the transition from heterogeneity to homogeneity in nucleic acid systems<sup>13,14</sup>. Whether the preference for homogeneous backbone ligands by chimeric templates (dictated by the thermodynamic stability of duplex formation) could be a general phenomenon for oligonucleotides composed of other different sugar backbones and/or nucleobases that are able to cross pair needs further examples (such as chimeras of 2',5'-RNA with 3',5'-RNA)<sup>19,49,50</sup> to validate its scope and limitations.

For the work described here, however, there are some issues still to be addressed: first, the use of EDC-mediated ligation combined with 3'-NH<sub>2</sub>-modified deoxynucleotide in this proof-of-principle study is not considered to be a plausibly prebiotic. To this end, we are exploring the use of other prebiotically plausible phosphorylation activation combined with oligomerization and ligation/recombination chemistries that may be compatible with the replication conditions<sup>51–54</sup>. We briefly explored the use of enzymes (T4 DNA ligase and T4 RNA ligase 2) with canonical RNA, DNA and RDNA chimeric sequences to check if ligases could be used to overcome the limitations of (1) the side reactions with chemical (EDC) activation<sup>55</sup> and (2) the need to synthesize sequences with the 3'-deoxy-NH<sub>2</sub> modification—so that we may be able to push towards many rounds of replication and sequence analysis within a shorter time span—but with limited success (Supplementary Figs. 95–103). We are exploring other ligases to expand the sequence space and length parameters to overcome the restrictions imposed by the EDC chemical ligation methods<sup>55,56</sup>.

Second, longer homogeneous products formed in the scenario described above are unlikely to work as continuous templates and may not provide the solution when moving towards a sustained replication of longer homogeneous strands that rely on thermodynamic-driven effects alone. One possible solution (alluded to in this work) is that chimeric templates can facilitate indirect replication by catalysing the accumulation of homogeneous strands. The product homogeneous strands can act as information storage, but cannot be directly replicated. Therefore, other mechanisms need to be invoked to allow the transfer of information stored in the homogeneous strands<sup>56</sup>. One straightforward pathway consistent with the above heterogeneity-to-homogeneity scenario is for the homogeneous RNA strands to give rise to functional ribozymes (ligase or polymerase) with the capability to take over the replication the homogeneous strands<sup>57</sup>. Other pathways could involve the beneficial effects

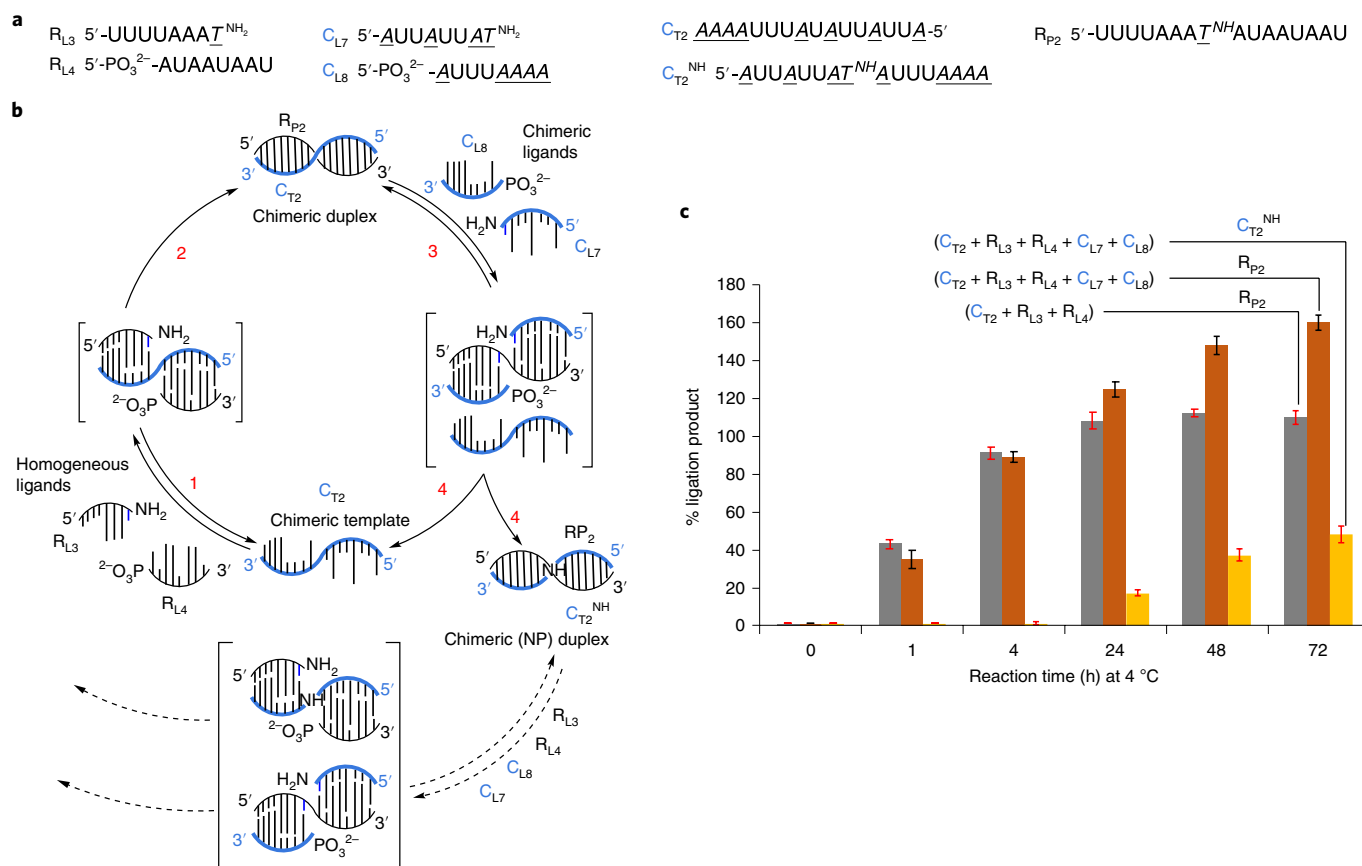


**Fig. 5 | Comparison of the efficiency between chimeric RDNA ( $C_{T2}$ ) and RNA ( $R_{T2}$ ) templates in producing the final ligation product  $R_{P3}$  under stepwise dilution conditions demonstrates the superior ability of  $C_{T2}$  to act as a template for ligation with turnover. **a**, Production of the ligation products  $R_{P2}$  and  $R_{P3}$  over period of 96 h in the stepwise dilution (in 24 h intervals) experiment with templates  $C_{T2}$  and  $R_{T2}$  and the four ligands  $R_{L3}$ ,  $R_{L4}$ ,  $R_{L5}$  and  $R_{L6}$ ; the drop in concentrations at 24, 48 and 72 h indicate the dilution step. **b**, Time course contrast between the templates  $C_{T2}$  and  $R_{T2}$  for the production of the first ligation product  $R_{P2}$  formed from  $R_{L3}$  and  $R_{L4}$ . **c**, Comparison of the efficiency of production of the second ligation product  $R_{P3}$  (from  $R_{L5}$  and  $R_{L6}$ ) between the templates  $C_{T2}$  and  $R_{T2}$ . **d**, Juxtaposed chromatogram traces at 96 h after three stepwise dilutions of the three parallel experiments in the presence of the  $C_{T2}$  template, the  $R_{T2}$  template and with no template. Supplementary Figs. 91–93 give the EDC-ligation conditions (at 4 °C).  $C_{T2}$ ,  $R_{T2}$ ,  $R_{P2}$ ,  $R_{P3}$ ,  $R_{L3}$ ,  $R_{L4}$ ,  $R_{L5}$  and  $R_{L6}$  are listed Fig. 4a. The lines in **a–c** are drawn as guides to indicate the trend and are not mathematical curve fittings. Percentage yields were calculated with respect to the template  $C_{T2}$ . Experiments were run in triplicate and the error range was less than  $\pm 5\%$ ; error bars represent s.d. AU, absorbance units.**

provided by different classes of molecules not considered in this study. For example, two other components, primordial (depsi)peptides<sup>58</sup> and protocells<sup>59</sup>, should be invoked, as they would have been an important part of any prebiotic scenario; they are as elementary as, if not more so than, the nucleotide building blocks<sup>53,60</sup>. Including them would be the next logical step to test the idea as to whether they could have not only aided in the transition from heterogeneity to homogeneity<sup>34</sup>, but also play a role in enabling the replication of information stored in the longer homogeneous RNA and DNA strands by overcoming the slower kinetics of strand exchange in the replication of homogeneous RNA and DNA strands as the strand lengths increase<sup>61,62</sup>.

Finally, in a prebiotic context, the possibility of oligomerization on chimeric templates starting with monomeric building blocks has to be considered alongside the ligation chemistry demonstrated in

this study<sup>8</sup>. In our work, we were influenced by the duplex stabilities and reasoned that (1) the selectivity expressed at the ligand–template level may not translate to the level of weaker monomer–template associations and (2), based on earlier studies<sup>8,63</sup>, the oligomerization of monomers would be biased towards G- and C-containing sequences (due to their stronger association) over A and U residues. Also, as argued by others<sup>64,65</sup>, the presence of dimers and trimers along with monomers in a prebiotic clutter may lead to the selective incorporation of the higher-order oligomers (dimers and trimers) over the monomers and, therefore, the ligation process may have an advantage over the oligomerization process. It is necessary to test the limits of oligomerization with monomers in a chimeric scenario to observe what the preference is, both in terms of the effects of the sugar and base residue (based on the nearest-neighbouring nucleotide)<sup>49,50</sup>.



**Fig. 6 | Experiment to test the possibility of cross-catalytic amplification in oligonucleotide replication via regeneration of the chimeric RDNA**

**( $C_{T2}$ ) template.** **a**, The sequences of oligonucleotides used in this investigation;  $C_{T2}^{NH}$  is the same as the chimeric template  $C_{T2}$ , but with a single dT-phosphoramidate (NP) link at the ligation junction. **b**, Schematic representation of the hypothesis that the presence of chimeric ligands  $C_{L7}$  and  $C_{L8}$  (complementary to  $R_{P2}$ ) could induce the regeneration of the chimeric template  $C_{T2}^{NH}$  and lead to further production of  $R_{P2}$ . The concomitant release of  $C_{T2}$  also creates the potential for another round of the ligation reaction. **c**, Comparison of the amount of  $R_{P2}$  produced from the combination of  $C_{T2} + R_{L3} + R_{L4}$  (1:5:5) versus the combination of  $C_{T2} + R_{L3} + R_{L4} + C_{L7} + C_{L8}$  (1:5:5:2:2) demonstrates the regeneration of the chimeric template  $C_{T2}^{NH}$  along with the higher and increasing production of  $R_{P2}$  in the latter combination. Supplementary Fig. 94 gives the experimental conditions. Percentage yields were calculated with respect to the template  $C_{T2}$ . Experiments were run in duplicate and the error range was less than  $\pm 5\%$ ; error bars represent s.d.

The results reported in this study have twofold implications for the emergence of homogeneous backbone nucleic acids. First, starting from a mixture of binary chimeric systems, for example, RDNA (a possibility that is strengthened by the recent report<sup>44</sup> of Sutherland and co-workers on the plausibly prebiotic conversion of RNA nucleotides into DNA nucleos(t)ides), there is the potential for the simultaneous emergence of the two respective homogeneous polymeric and communicating informational systems (RNA and DNA). This is opposed to the often-suggested sequential model with RNA as the forerunner and DNA as the successor. The successive replication cycles<sup>42,66,67</sup> are expected to lead, simultaneously, to the two respective strands that contain the homogeneous sugar backbone (RNA and DNA), as indicated by the results in Figs. 3–6. Therefore, if RNA and DNA could have appeared together, there is no need for a genetic takeover by the new informational system (DNA) from an older system (RNA), a suggestion that has been made implicitly and explicitly by others<sup>14,17,23–25,28,68,69</sup>, as there is neither a predecessor nor a successor in this scenario. This is also true for the supposed pre-RNA to RNA transition<sup>33</sup>; for example, there is no need for RNA to be the descendent of TNA, when tRNA can simultaneously give rise to TNA and RNA. Second, the generality of this phenomenon—exemplified by RDNA and tRNA chimera systems—lends experimental credence to a point that is implied in Fig. 1, and one that has been discussed before<sup>13,21,28</sup>; namely, a clean

and directed prebiotic synthesis of a nucleotide building block of a particular oligonucleotide (for example, TNA, RNA or DNA) is not an absolute requisite for a homogeneous backbone nucleic acid like RNA to emerge. In other words, as is suggested in Fig. 1, the appearance of a system with homogeneous nucleotide backbone repeat units can be achieved at the emergent level of a replicating polymer<sup>34</sup>. Therefore, a mixture of diverse nucleotides can, via the formation of mixtures of oligonucleotides and the ensuing emergent property of template-mediated ligation, tend towards homogeneous nucleotide backbone systems<sup>13</sup>. This process can include alternative linker units and alternative nucleobases<sup>10,17,19,70</sup>, and chirality of the building blocks<sup>71</sup>. This means that the appearance of homogeneous backbone homochiral polymers with a set of uniform building blocks from a prebiotic mixture is a natural outcome of chemical evolution<sup>14</sup>, without the need to invoke the predecessor–successor models of extant biology<sup>34,68,72</sup>.

#### Data availability

Full experimental details and data are provided in the Supplementary Information. The raw data that support the findings of this study are available from the corresponding author upon reasonable request.

Received: 10 January 2019; Accepted: 31 July 2019;  
Published online: 16 September 2019



## References

- Gilbert, W. Origin of life—the RNA world. *Nature* **319**, 618–618 (1986).
- Joyce, G. F. & Orgel, L. E. in *The RNA World* 3rd edn, Vol. 43 (eds Gesteland, R. F., Cech, T. R. & Atkins, J. F.) 23–56 (Cold Spring Harbor Laboratory Press, 2006).
- Gesteland, R. F., Cech, T. R. & Atkins, J. F. *The RNA World* 2nd edn (Cold Spring Harbor Laboratory Press, 1999).
- Schwartz, A. in *The Molecular Origins of Life: Assembling Pieces of the Puzzle* (ed. Brack, A.) 237–254 (Cambridge Univ. Press, 1998).
- Orgel, L. E. Prebiotic chemistry and the origin of the RNA world. *Crit. Rev. Biochem. Mol. Biol.* **39**, 99–123 (2004).
- Anastasi, C. et al. RNA: prebiotic product, or biotic invention? *Chem. Biodivers.* **4**, 721–739 (2007).
- Robertson, M. P. & Joyce, G. F. The origins of the RNA world. *Cold Spring Harb. Perspect. Biol.* **4**, a003608 (2012).
- Szostak, J. The eightfold path to non-enzymatic RNA replication. *J. Sys. Chem.* **3**, 2 (2012).
- Higgs, P. G. & Lehman, N. The RNA world: molecular cooperation at the origin of life. *Nat. Rev. Genet.* **16**, 7–17 (2015).
- Hud, N. V., Cafferty, B. J., Krishnamurthy, R. & Williams, L. D. The origin of RNA and ‘my grandfather’s axe’. *Chem. Biol.* **20**, 466–474 (2013).
- Joyce, G. F. The antiquity of RNA-based evolution. *Nature* **418**, 214–221 (2002).
- Islam, S. & Powner, M. W. Prebiotic systems chemistry: complexity overcoming clutter. *Chem. 2*, 470–501 (2017).
- Krishnamurthy, R. On the emergence of RNA. *Isr. J. Chem.* **55**, 837–850 (2015).
- Gavette, J. V., Stoop, M., Hud, N. V. & Krishnamurthy, R. RNA–DNA chimeras in the context of an RNA world transition to an RNA/DNA world. *Angew. Chem. Int. Ed.* **55**, 13204–13209 (2016).
- Joyce, G. F., Schwartz, A. W., Miller, S. L. & Orgel, L. E. The case for an ancestral genetic system involving simple analogs of the nucleotides. *Proc. Natl Acad. Sci. USA* **84**, 4398–4402 (1987).
- Lazcano, A. & Miller, S. L. The origin and early evolution of life: prebiotic chemistry, the pre-RNA world, and time. *Cell* **85**, 793–798 (1996).
- Trevino, S. G., Zhang, N., Elenko, M. P., Luptak, A. & Szostak, J. W. Evolution of functional nucleic acids in the presence of nonheritable backbone heterogeneity. *Proc. Natl Acad. Sci. USA* **108**, 13492–13497 (2011).
- Sheng, J. et al. Structural insights into the effects of 2′–5′ linkages on the RNA duplex. *Proc. Natl Acad. Sci. USA* **111**, 3050–3055 (2014).
- Engelhart, A. E., Powner, M. W. & Szostak, J. W. Functional RNAs exhibit tolerance for non-heritable 2′–5′ versus 3′–5′ backbone heterogeneity. *Nat. Chem.* **5**, 390–394 (2013).
- Eschenmoser, A. The TNA-family of nucleic acid systems: properties and prospects. *Orig. Life Evol. Biosph.* **34**, 277–306 (2004).
- Powner, M. W., Zheng, S.-L. & Szostak, J. W. Multicomponent assembly of proposed DNA precursors in water. *J. Am. Chem. Soc.* **134**, 13889–13895 (2012).
- Szostak, J. W. An optimal degree of physical and chemical heterogeneity for the origin of life? *Phil. Trans. R. Soc. B* **366**, 2894–2901 (2011).
- Orgel, L. E. & Lohrmann, R. Prebiotic chemistry and nucleic acid replication. *Acc. Chem. Res.* **7**, 368–377 (1974).
- Woese, C. R. *The Genetic Code: The Molecular Basis for Genetic Expression*. (Harper and Row, 1967).
- Oró, J. & Stephen-Sherwood, E. in *Cosmochemical Evolution and the Origins of Life* (eds Oró, J., Miller, S. L., Ponnampertuma, C. & Young, R. S.) 159–172 (Springer, Dordrecht, 1974).
- Becker, S. et al. Wet–dry cycles enable the parallel origin of canonical and non-canonical nucleosides by continuous synthesis. *Nat. Commun.* **9**, 163 (2018).
- Islam, S. et al. Detection of potential TNA and RNA nucleoside precursors in a prebiotic mixture by pure shift diffusion-ordered NMR spectroscopy. *Chem. Eur. J.* **19**, 4586–4595 (2013).
- Sutherland, J. D. & Whitfield, J. N. Prebiotic chemistry: a bioorganic perspective. *Tetrahedron* **53**, 11493–11527 (1997).
- Usher, D. Early chemical evolution of nucleic acids: a theoretical model. *Science* **196**, 311–313 (1977).
- Mehta, A. P. et al. Bacterial genome containing chimeric DNA–RNA sequences. *J. Am. Chem. Soc.* **140**, 11464–11473 (2018).
- Song, X.-P., Maiti, M. & Herdewijn, P. Enzymatic synthesis of DNA employing pyrophosphate-linked dinucleotide substrates. *J. Sys. Chem.* **2**, 3 (2011).
- Schoning, K.-U. et al. The  $\alpha$ -L-thiofuranosyl-(3′→2′)-oligonucleotide system (‘TNA’): synthesis and pairing properties. *Helv. Chim. Acta* **85**, 4111–4153 (2002).
- Orgel, L. E. Origin of life: a simpler nucleic acid. *Science* **290**, 1306–1307 (2000).
- Krishnamurthy, R. Giving rise to life: transition from prebiotic chemistry to protobiology. *Acc. Chem. Res.* **50**, 455–459 (2017).
- Pallan, P. S. et al. Why does TNA cross-pair more strongly with RNA than with DNA? An answer from X-ray analysis. *Angew. Chem. Int. Ed.* **42**, 5893–5895 (2003).
- Butlerov, A. Formation synthetique d’une substance sucree. *Acad. Sci.* **53**, 145–147 (1861).
- Kim, H.-J. et al. Synthesis of carbohydrates in mineral-guided prebiotic cycles. *J. Am. Chem. Soc.* **133**, 9457–9468 (2011).
- Kim, H.-J. & Benner, S. A. Prebiotic stereoselective synthesis of purine and noncanonical pyrimidine nucleotide from nucleobases and phosphorylated carbohydrates. *Proc. Natl Acad. Sci. USA* **114**, 11315–11320 (2017).
- Weber, A. L. & Pizzarello, S. The peptide-catalyzed stereospecific synthesis of tetroses: a possible model for prebiotic molecular evolution. *Proc. Natl Acad. Sci. USA* **103**, 12713–12717 (2006).
- von Kiedrowski, G. A self-replicating hexadeoxynucleotide. *Angew. Chem. Int. Ed.* **25**, 932–935 (1986).
- Wu, X., Delgado, G., Krishnamurthy, R. & Eschenmoser, A. 2,6-Diaminopurine in TNA: effect on duplex stabilities and on the efficiency of template-controlled ligations. *Org. Lett.* **4**, 1283–1286 (2002).
- Derr, J. et al. Prebiotically plausible mechanisms increase compositional diversity of nucleic acid sequences. *Nucleic Acids Res.* **40**, 4711–4722 (2012).
- Yu, H. Y., Zhang, S. & Chaput, J. C. Darwinian evolution of an alternative genetic system provides support for TNA as an RNA progenitor. *Nat. Chem.* **4**, 183–187 (2012).
- Xu, J., Green, N., Gibard, C., Krishnamurthy, R. & Sutherland, J. Prebiotic phosphorylation of 2-thiouridine provides either nucleotides or DNA building blocks via photoreduction. *Nat. Chem.* **11**, 457–462 (2019).
- Fernando, C., Von Kiedrowski, G. & Szathmáry, E. A stochastic model of nonenzymatic nucleic acid replication: ‘elongators’ sequester replicators. *J. Mol. Evol.* **64**, 572–585 (2007).
- Grossmann, T. N., Strohbach, A. & Seitz, O. Achieving turnover in DNA-templated reactions. *ChemBioChem* **9**, 2185–2192 (2008).
- He, C., Gállego, I., Laughlin, B., Grover, M. A. & Hud, N. V. A viscous solvent enables information transfer from gene-length nucleic acids in a model prebiotic replication cycle. *Nat. Chem.* **9**, 318–324 (2017).
- Duim, H. & Otto, S. Towards open-ended evolution in self-replicating molecular systems. *Beil. J. Org. Chem.* **13**, 1189–1203 (2017).
- Ertem, G. & Ferris, J. P. Synthesis of RNA oligomers on heterogeneous templates. *Nature* **379**, 238–240 (1996).
- Prakash, T. P., Roberts, C. & Switzer, C. Activity of 2′,5′-linked RNA in the template-directed oligomerization of mononucleotides. *Angew. Chem. Int. Ed.* **36**, 1522–1523 (1997).
- Mutschler, H. et al. Random-sequence genetic oligomer pools display an innate potential for ligation and recombination. *eLife* **7**, e43022 (2018).
- Lutay, A. V., Chernolovskaya, E. L., Zenkova, M. A. & Vlassov, V. V. The nonenzymatic template-directed ligation of oligonucleotides. *Biogeosciences* **3**, 243–249 (2006).
- Gibard, C., Bhowmik, S., Karki, M., Kim, E.-K. & Krishnamurthy, R. Phosphorylation, oligomerization and self-assembly in water under potential prebiotic conditions. *Nat. Chem.* **10**, 212–217 (2018).
- Vaidya, N. et al. Spontaneous network formation among cooperative RNA replicators. *Nature* **491**, 72–77 (2012).
- Taran, O., Thoennessen, O., Achilles, K. & von Kiedrowski, G. Synthesis of information-carrying polymers of mixed sequences from double stranded short deoxynucleotides. *J. Sys. Chem.* **1**, 9 (2010).
- Edeleva, E. et al. Continuous nonenzymatic cross-replication of DNA strands with in situ activated DNA oligonucleotides. *Chem. Sci.* **10**, 5807–5814 (2019).
- Mutschler, H., Wochner, A. & Holliger, P. Freeze–thaw cycles as drivers of complex ribozyme assembly. *Nat. Chem.* **7**, 502–508 (2015).
- Forsythe, J. G. et al. Ester-mediated amide bond formation driven by wet–dry cycles: a possible path to polypeptides on the prebiotic Earth. *Angew. Chem. Int. Ed.* **127**, 10009–10013 (2015).
- Joyce, G. F. & Szostak, J. W. protocells and RNA self-replication. *Cold Spring Harb. Perspect. Biol.* **10**, a034801 (2018).
- Ruiz-Mirazo, K., Briones, C. & de la Escosura, A. Prebiotic systems chemistry: new perspectives for the origins of life. *Chem. Rev.* **114**, 285–366 (2014).
- Kamat, N. P., Tobe, S., Hill, I. T. & Szostak, J. W. Electrostatic localization of RNA to protocell membranes by cationic hydrophobic peptides. *Angew. Chem. Int. Ed.* **54**, 11735–11739 (2015).
- Chen, I. A. & Walde, P. From self-assembled vesicles to protocells. *Cold Spring Harb. Perspect. Biol.* **2**, a002170 (2010).
- Joyce, G. F., Inoue, T. & Orgel, L. E. Non-enzymic template-directed synthesis on RNA random copolymers. poly(C, U) templates. *J. Mol. Biol.* **176**, 279–306 (1984).
- Prywes, N., Blain, J. C., Del Frate, F. & Szostak, J. W. Nonenzymatic copying of RNA templates containing all four letters is catalyzed by activated oligonucleotides. *eLife* **5**, e17756 (2016).
- Pfeffer, D., Sosson, M. & Richert, C. Enzyme-free ligation of dimers and trimers to RNA primers. *Nucl. Acids Res.* **47**, 3836–3845 (2019).
- Leu, K., Obermayer, B., Rajamani, S., Gerland, U. & Chen, I. A. The prebiotic evolutionary advantage of transferring genetic information from RNA to DNA. *Nucl. Acids Res.* **39**, 8135–8147 (2011).
- Tupper, A., Shi, K. & Higgs, P. The role of templating in the emergence of RNA from the prebiotic chemical mixture. *Life* **7**, 41 (2017).

68. Dworkin, J. P., Lazcano, A. & Miller, S. L. The roads to and from the RNA world. *J. Theor. Biol.* **222**, 127–134 (2003).
69. Brewin, N. Catalytic role for RNA in DNA replication. *Nat. New Biol.* **236**, 101 (1972).
70. Krishnamurthy, R. RNA as an emergent entity: an understanding gained through studying its nonfunctional alternatives. *Synlett* **25**, 1511–1517 (2014).
71. Ribó, J. M., Hochberg, D., Crusats, J., El-Hachemi, Z. & Moyano, A. Spontaneous mirror symmetry breaking and origin of biological homochirality. *J. R. Soc. Interface* **14**, 20170699 (2017).
72. Krishnamurthy, R. Life's biological chemistry: a destiny or destination starting from prebiotic chemistry? *Chem. Eur. J.* **24**, 16708–16715 (2018).

## Acknowledgements

The work was supported by a grant from NASA (NNX14AP59G) and the Simons Foundation to R.K. (327124). S.B. thanks the NASA Astrobiology Postdoctoral Program for a fellowship. We thank the S. F. Dowdy laboratory for the use of their instrument for MALDI–TOF analysis. We thank J. Szostak, I. Chen, D. Braun, U. Muller, L. Leman, A. Lazcano and our lab members for helpful discussions.

## Author contributions

R.K. conceived the project. R.K. and S.B. designed the experiments. S.B. performed all the experiments. R.K. wrote the paper with inputs from S.B. Both authors discussed the results and commented on the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41557-019-0322-x>.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Correspondence and requests for materials** should be addressed to R.K.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2019