Prebiotic synthesis and triphosphorylation of 3′-amino-TNA nucleosides

Authors Daniel Whitaker and Matthew W. Powner*

Affiliations Department of Chemistry, University College London, 20 Gordon Street, London WC1H 0AJ, UK. *e-mail: matthew.powner@ucl.ac.uk

Abstract Nucleosides are essential to the emergence of life, and so their synthesis is a key challenge for prebiotic chemistry. Whilst amino-nucleosides have enhanced reactivity in water compared with ribonucleosides, they are assumed to be prebiotically irrelevant due to perceived difficulties with their selective formation. Here we demonstrate 3′-amino-TNA nucleosides are formed diastereoselectively and regiospecifically from prebiotic feedstocks in four high-yielding steps. Phosphate provides an unexpected resolution, leading to spontaneous purification of the genetically relevant *threo*-isomer. Furthermore, 3′-amino-TNA nucleosides are shown to be phosphorylated directly in water, under mild conditions with cyclic trimetaphosphate, forming a nucleoside triphosphate (NTP) in a manner not feasible for canonical nucleosides. Our results suggest 3′-amino-TNA nucleosides may have been present on the early Earth, and the ease with which these NTPs form, alongside the inherent selectivity for the Watson-Crick base pairing *threo*-monomer, warrants further study of the role they could play during the emergence of life.

A significant obstacle to understanding life's emergence on Earth is how the first nucleotides, which are essential for Darwinian evolution,[1] could have been synthesised. The cellular machinery currently used to enable replication of DNA is too complex to have emerged without invoking evolution, so must be the product of a simpler ancestral system.[2] The "RNA world" hypothesis posits that this ancestral system was composed of self-replicating RNA which both encodes genetic information and can exhibit catalytic function. While its capacity to act as both genotype and phenotype in extant biology makes RNA an attractive candidate for the first genetic material, it is plausible that a different – perhaps constitutionally simpler – nucleotide may have preceded it,[3] or that the first genetic material was chimeric in nature, composed of several different classes of nucleotide.[4]

Non-canonical nucleotides have been considered as possible precursors to RNA (and DNA) for decades.[5,6] A variety of different nucleotide structures, with changes both to the sugar

backbone and the nucleobases, have been investigated.[7,8] The key property of a potential genetic polymer is information transfer, i.e. the capacity to pass on the information contained within its own sequence. Ultimately, for a genetic polymer to have preceded D/RNA it must be able to pass on genetic information to these polymers as well as to itself. Changes to the nucleobase are rarely hereditable,[9] while changes to the sugar often prevent stable duplex formation (and so information transfer) with D/RNA.[10] However, almost 20 years ago, as part of a seminal study evaluating the potential of non-canonical nucleotides to have been the first genetic polymers of life, Eschenmoser evaluated RNA, threose nucleic acid (TNA),[11] 2′-amino-TNA and 3′-amino-TNA.[3,12] Unusually for non-canonical nucleotides TNA and amino-TNA polymers are able to form stable duplexes with both RNA and DNA as well as with themselves, and can also form catalytic polymers akin to ribozymes,[13,14] setting the stage for a potential genetic takeover and for them to have played a role in the origins of life on Earth.[3,12]

Eschenmoser considered the differences in 'constitutional' verses 'generational' complexity of TNA and RNA nucleotides, comparing their structural complexity with how readily they could have formed. He concluded that TNA's four carbon sugar means that its monomers are "constitutionally" simpler than RNA's monomers, but because the most difficult steps for the emergence of nucleotides seemed to be formation of the glycosidic bond and phosphorylation of the nucleosides, they were not necessarily "generationally" simpler. On the other hand, he considered amino-TNA's monomers to be both generationally and constitutionally more complex than RNA's monomers due to the additional difficulty in regioselective incorporation of an amine into a sugar. Indeed, when Eschenmoser considered the effect that nitrogenous compounds would have on the aldol chemistry that, at the time, seemed the most likely source of sugars on the early Earth,[15] he concluded that "huge chemical complications" would arise due to Amadori and Mannich chemistry occurring alongside the already unselective aldol reactions. He nevertheless recognised that these complications may "harbour structural and transformational opportunities that may be relevant" to the origins of life. We therefore decided to re-examine the suggestion that amino-TNA's monomers are more generationally complex than RNA's monomers, in light of chemistry discovered since, with the aim of discovering a facile and selective amino-nucleoside synthesis. Our re-evaluation indicates that one such opportunity for monomer synthesis is the regioselective incorporation of an amine moiety at the 3' position of a sugar during aminooxazoline synthesis (Figure 1).

How the first nucleotides formed on Earth is still an open question,[16] but there is a strong growing consensus that the earliest life would have emerged from a system containing peptides and lipids as well as nucleotides.[17] Despite this, most proposed mechanisms by which nucleotides could have formed on the early Earth[18] are still based on two assumptions – first, that the sugar and base moieties formed separately and then combined at a late stage,[19] and second, that nucleotide synthesis and amino acid/peptide synthesis occurred separately. We have previously challenged the first assumption, discovering highly diastereoselective routes to ribonucleotides[20,21] in which the sugar and bases are constructed together in the same environment on the same scaffold. This sequence represents the only diastereoselective prebiotically plausible formation of ribonucleotides known so far, and proceeds via the coupling of glycolaldehyde **1** with cyanamide **2** to form 2-aminooxazole **3**, which then reacts with glyceraldehyde **4** forming the key intermediate aminooxazoline **5** (Figure 1b).

The advantages delivered through unification of nitrile and sugar chemistry in our synthesis have led us to consider the relationship between peptides and nucleotides, and how their syntheses would intersect.[22] Aminonitriles **6** have long been considered to be precursors of amino acids[23,24] since their formation via the Strecker reaction is both facile and predisposed to form biogenic α-aminonitriles. However, the intrinsic reactivity of these molecules has been overlooked, with a major focus on how to improve their relatively sluggish hydrolysis to amino amides and then acids. Unlike amino acids, aminonitriles are predisposed towards reactivity in water due to the kinetically inert but high energy nitrile moiety and the suppressed pK_a of the amine rendering them nucleophilic at neutral pH. We recently demonstrated that this dual reactivity could be exploited to form peptide bonds in water, with high selectivity for biologically relevant (proteinogenic) α-peptides over unnatural isomers.[25,26] But how then does this chemistry interact with nucleotide synthesis? And can it be used to overcome the perceived complexity of aldol, Mannich and Amadori reactions (Figure 2a) expected in nitrogenous aldol chemistry?

Results and Discussions

Amino-sugar synthesis

We have previously reported that at pH 5–6 in the presence of 5-aminoimidazole-4-carboxamide (AICA) the reaction between sugars and 2-aminooxazole **3** is out-competed by a three component reaction. Imine formation between AICA and the sugar (e.g. glycolaldehyde

1) preceded reaction with **3**, forming an aminooxazoline product (**11**) regiospecifically functionalised by an amine at C3' (Figure 2b).[27] In a related reaction between amino acids, glyceraldehyde **4** and 2-aminooxazole **3**, Blackmond observed a mixture of both 2- and 3-component coupling products (Figure 2d).[28] The 3-component reaction was diastereoselective with enantiopure proline which, due to the low chemoselectivity, led to enantioenrichment of glyceraldehyde **4** and therefore consequently enantioenrichment of the 2-component aminooxazoline products **5**. We recognised that if a variant of this reaction could be rendered selective for 3-component coupling with an $-NH_2$ source, a simple route to 2,3'-diaminooxazolines **7** would be uncovered. We envisaged that aminonitriles **6**, due to their low amine pK_a, would be efficiently incorporated and that, due to the reversibility of the Strecker reaction, the free amine could then be liberated after coupling, and therefore deliver a selective route to 3'-amino-nucleosides.

We envisaged the reaction of glycolaldehyde **1**, 2-aminooxazole **3** and an aminonitrile **6** would form tetrose 3'-amino-aminooxazolines **7**. The nucleobase could then be constructed on this compound by reaction with cyanoacetylene **8**, and anomerisation of the base would furnish the base-pairing α-*threo*-nucleoside *threo*-**10** (Figure 1c). We therefore began our investigation by looking at the three-component reaction between glycolaldehyde **1**, 2-aminooxazole **3** and glycine nitrile **6a** (R=H) in water at neutral pH. Although the reaction was not selective, **12a** was clearly visible in the ^{1}H NMR spectrum. Employing a slight excess of **6a** improved the chemoselectivity, with the yield increasing to 40% (Table 1, entry 1). Based on our previous work with AICA we expected that three-component coupling would be favoured at low pH, and indeed at pH 5 we observed very efficient three-component coupling, forming **12a** in 94% yield after 16 h (Table 1, entry 2). The yield reached a maximum at pH 5. At lower pH lower yields were observed – at pH 4 the reaction stalled after four hours and 60% conversion, but all three starting materials were accounted for (Table 1, entry 3). When initiated at pH 4, the solution was observed to drop to pH 2 as the reaction proceeded, and at pH 2 coupling is observed to proceed only very slowly (Supplementary Figure 1). **6a** (pK_{aH} = 5.2) can act as the ideal buffer, as well as a coupling partner, keeping the pH close to the optimum value of pH 5. This buffering effect is highly apparent in reactions initiated at pD 4.5, where with 1 equiv. of **6a** the reaction stops at 55% conversion, but is then restarted if the solution is adjusted back from pD 2.5 to pD 4.5 reaching 75% conversion. However, when the reaction was initiated at pD 4.5 with 3 equiv. of **6a** it did not stall, the pH was maintained and **12a** was furnished in

90% yield. At pH 5 3-component coupling was seen to dominate even at concentrations as low as 10 mM, albeit more slowly, with **12a** formed in 70% yield after 10 days (Table 1, entry 4 & Supplementary Figure 12).

If glycyl aminooxazoline **12a** is heated at 60 °C a slow retro-Strecker reaction occurs, liberating a free amino group to form **7** (Table 1, entry 5 & Supplementary Figure 26). However, hydrolysis of the aminooxazoline moiety of **12a** to oxazolidinone **13** occurred concomitantly with glycyl-retro-Strecker upon heating. Similar hydrolysis of pentose aminooxazolines (**5**) to their corresponding oxazolidinones has previously been reported to occur on heating in phosphate solution.[29] This C2-hydrolysis can be supressed with more sterically encumbered aminonitriles which undergo a more rapid retro-Strecker reaction. For example, α-alkylated aminonitriles **12b-e** (Table 1, entries 5–9) all undergo retro-Strecker reaction at room temperature, proceeding efficiently over several days (Extended Data Figure 1). The cyanohydrin formed in this process is not consumed and would be recycled into aminonitriles in the presence of a suitable source of ammonia.[30] Indeed, in the case of the hydrophobic R groups (e.g. ^{i}Pr and ^{i}Bu) cyanohydrin crystallisation was seen to occur spontaneously from the reaction mixture.

Although aminonitriles are known to be unstable compared to cyanohydrins and ammonium at low concentrations of free ammonia[31] they are reasonably kinetically stable at neutral or acidic pH. Commeyras, however, has shown that α,α-disubstituted aminonitriles undergo retro-Strecker reactions easily even at acidic pH at room temperature, and that N-alkylation further increases the rate of retro-Strecker.[32] On the other hand, Commeyras reported that alanine nitrile (**6b**) was roughly 20 times more stable than glycine nitrile (**6a**),[33] but our data suggests that α-alkylation promotes more rapid retro-Strecker (**12b** > **12a**) and that **6a** and its derivatives are kinetically more stable to retro-Strecker reactions than the more substituted aminonitriles.

Aminonitriles **6** are more reactive coupling partners than either amino acids or ammonia under the optimal conditions for 3-component coupling. For example, when 200 mM **1**, **3**, **6a** and glycine were mixed at pH 5, **12a** (80%) was seen to form as the predominant product after 24 h (Supplementary Figure 20). Furthermore, if **1** and **3** were mixed in the presence of 600 mM **6a** in 1 M NH_4Cl at pH 5, **12a** was seen to form in 95% yield after 24 h (Supplementary Figure 16).

When the reactions of α-substituted aminonitriles initiated at pH 5 were monitored for several weeks the concentration of **7** was seen to fall slightly over time, without concomitant increases in either **12** or oxazolidinone **13**. The concentration of *threo*-**7** seemed to fall less than *erythro*-**7** (Supplementary Figure 28). To explore this intriguing behaviour further we isolated **7** as a mixture of diastereomers.

threo-Selective crystallisation

When **7** (3:2 *threo*:*erythro;* Figure 3a) was mixed with 200 mM phosphate at near-neutral pH a colourless precipitate was seen to form rapidly. Proton NMR spectra indicated that the supernatant was significantly enriched in *erythro*-**7**, and *threo*-**7** had precipitated (Figure 3b). Over time *erythro*-**7** was then seen to transform into a new compound, assigned as the bridged bicyclic compound **14** on the basis of 2D NMR and mass spectral data (Figure 3c & Supplementary Figures 42–44).[3] **14** was found to be unreactive in the subsequent synthetic steps. The high pK_{aH} of the guanidine moiety (pK_{aH} of guanidinium – 13.7) means that **14** is fully protonated, and so its nucleophilicity is quenched at neutral pH. The same rearrangement would also be expected to occur in 3′-amino-*ribo*-aminooxazolines, since *ribo* and *erythro* aminooxazolines have the same stereochemical orientation at their 2′ and 3′ carbons, which would preclude the formation of 3′-amino-RNA.

Analysis of the precipitate demonstrated that the phosphate salt of *threo*-**7** had indeed undergone a highly diastereoselective crystallisation (Figure 3d). This provides a natural route for separation and concentration of the genetically useful diastereomer, *threo*-**7**, while removing the potentially inhibitory *erythro*-isomer. This *threo*-selective precipitation is observed at concentrations of *threo*-**7** as low as 10 mM in 100 mM phosphate. In contrast, precipitation of *ribo*-aminooxazoline **5**, thought to provide a selection mechanism for RNA,[22,29,34] occurs efficiently only above 100 mM and does not co-accumulate **5** with phosphate (Supplementary Figures 37 – 40).

The diastereoselectivity imparted in **7** by phosphate is therefore twofold: phosphate drives the co-precipitation, purification and concentration of the *threo*-isomer, whilst catalysing the conversion of the *erythro*-isomer into an unreactive by-product. That the chemistry is wholly predisposed to favour the genetically useful isomer is an interesting and unprecedented result.

Nucleobase elaboration

Phosphate-buffered addition of cyanoacetylene **8** to pentose aminooxazolines **5** in near-neutral water quantitatively converts them to 2,2′-anhydrocytidines,[11,20] building the nucleobase on the sugar scaffold. When *threo*-**7** was mixed with **8** at pH 7 competitive addition to both nucleophilic sites on the molecule was observed, forming *threo*-**9** and *threo*-**15** as the predominant products in 40% yield after 24 h (Supplementary Figure 53). A portion of *threo*-**7** remains undissolved as its solid phosphate salt. If the salt is dissolved at pH 8.5 before addition of cyanoacetylene **8**, and the pH is then lowered to 7.0, a higher 74% combined yield is obtained (Figure 4a & Supplementary Figure 56). The NMR spectrum in D_2O very clearly shows that only two species derived from *threo*-**7** are present after 24 h, *threo*-$\mathbf{9_d}$ and *threo*-$\mathbf{15_d}$ (Supplementary Figure 62). However, addition of **8** to the 3′-amine does not significantly inhibit addition to the aminooxazoline,[35] and once excess cyanoacetylene **8** has been consumed *threo*-**15** slowly converts to *threo*-**9** and cyanoacetaldehyde **16**. If no adjustment to the reaction conditions are made, hydrolysis of *threo*-**15** to *threo*-**9** (33%) is complete after 28 days, but this sluggish hydrolysis is then accompanied by hydrolysis to *threo*-**17**, which forms in 22% yield after 28 days (Supplementary Figure 57). Some precipitation of nucleoside-derived products is seen to occur over this time alongside significant precipitation of cyanoacetaldehyde oligomer products which leads to partial enrichment of nucleosides in the supernatant. However, the sluggish hydrolysis of *threo*-**15** to *threo*-**9** can be greatly accelerated, for example by acid catalysis, and at pH 2 the conversion is rapid (< 10 min) and quantitative.

These results demonstrate that the skeletal structure of 3'-amino-*threo*-nucleosides can be regioselectively established in only two steps from prebiotically plausible two- and three-carbon units. Formation of amino-nucleotides would then require anomerisation of the base and phosphorylation. We set out to consider how these could have occurred on the early Earth. Anhydronucleoside hydrolysis seems to be the disposed pathway for anhydronucleosides in water (Figure 4b). It is very sluggish, taking weeks to complete at neutral pH, but is accelerated at higher pH occurring within 6 hours at pH 8 or within 10 mins at pH 11 to furnish *threo*-**17** (quant.). However, *threo*-**17** has not been shown to be an effective component of a proto-genetic material, and by comparison with the weaker duplex formation in α-DNA,[36] would be expected to form less stable duplexes than polymers of *threo*-**10**. Therefore, if *threo*-**17** were to be a precursor to a genetic polymer on the early Earth anomerisation of the 1′ stereocenter would likely have been highly beneficial.

Photochemical anomerisation

Irradiation at 254 nm is the most promising mechanism uncovered so far for anomerisation of pyrimidine nucleosides. However, for canonical nucleobases this process is still inefficient, with their nucleotides undergoing condensation to form oxazolidinones as the major product upon irradiation, with anomerisation only occurring in 4-6% yield.[37] In contrast, 2-thiocytosine ribonucleosides undergo extremely clean anomerisation to preferentially form a *trans*-disposition between the 1' and 2' substituents.[38] We recently demonstrated that the same anomerisation can occur on a *threo* skeleton,[11] and therefore decided to investigate amino-nucleoside photoanomerisation. As expected, when *threo*-**17** was irradiated at 254 nm in neutral water it slowly formed oxazolidinone **13** along with base release to give cytosine **18** (Figure 4b & Supplementary Figure 78–79). We therefore considered how sulfur would have been incorporated into *threo*-**9** to give *threo*-**19.**

Thiolysis of α-anhydrocytidines occurs in the presence of sodium hydrosulfide in formamide.[38,11] Although these conditions are not trivial to rationalise as prebiotically plausible, all the chemical inputs are individually plausible, and the demonstrated efficiency and value of this chemical transformation outweighs speculative plausibility at this stage – it seems reasonable that there may be a different, and more plausible route to *threo*-**19** and it would be counterproductive to discount uniquely effective chemistry on unverifiable grounds of plausibility. We therefore tested the reaction of *threo*-**9** and hydrosulfide in dimethylformamide (DMF) and observed quantitative conversion to *threo*-**19** in the crude NMR (Supplementary Figure 74). Similar, although less clean, thiolysis was observed in formamide (Supplementary Figure 75).

When **19** was irradiated at 254 nm in the presence of H_2S in neutral water clean anomerisation was found to occur, providing *threo*-**20** (65%) along with 7% 2-thiocytosine (Figure 4c + 4d & Supplementary Figures 76–77). 2-Thiopyrimidines are intriguing molecules from the point of view of Watson-crick base pairing. For example, 2-thiouridine and 2-thiocytidine are found in tRNA[39,40] and the enhanced selectivity for U:A base pairing over the U:G wobble pair when uridine is replaced by 2-thiouridine[41] has been exploited to improve the fidelity of non-enzymatic primer extension[42] and ribozyme-catalysed copying of RNA.[43] Further to this, 5-methyl-2-thiocytidine has been shown to selectively bind with inosine,[44] which has been considered as a likely precursor to guanosine in early genetic materials due to its easier synthesis and greatly increased solubility.[45] 2-thioU:A and 2-thioC:I Watson-Crick pairing deserves further investigation from the point of view of duplex stability and copying efficiency,

particularly in light of these compounds potentially being synthetic precursors to the oxygenic variants.

How the cytosine-nucleoside *threo*-**10** could have formed from its 2-thiocytosine presents an interesting problem here due to the nucleophilicity of the 3'-amine. Activation of the sulfur, either by alkylation or oxidation, would render it a better leaving group and so greatly accelerate substitution at this position,[11] but in the case of *threo*-**20** such activation leads to rapid cyclisation to **24** (Figure 5 & Supplementary Figure 80), even at pH 5 where the amine would be highly protonated (we have not observed either cyclisation or hydrolysis of *threo*-**20** in the absence of an activating agent in rigorously deoxygenated water, even over several weeks, although we cannot rule out that substitution occurs extremely slowly).

Considering the issue of cyclisation, we next reflected on our target structure and realised it was the phosphorylated amino-TNA nucleo*tide*, not the unphosphorylated nucleo*side*. Accordingly, we recognised that this cyclisation problem may have arisen by assuming that the nucleoside should be an intermediate in the synthesis. The nucleophilicity of the amine would be suppressed by phosphorylation, and therefore the solution to monomer cyclisation may lie in incorporation of the intermediate into a nucleotide – where we expected cyclisation to be suppressed in favour of hydrolysis – before converting the pyrimidine to the canonical nucleobase.

Triphosphorylation

Nucleoside phosphorylation is extremely challenging in water and has remained a long standing problem, with reasonably efficient phosphorylation only being seen for vicinal diols in water.[46] Furthermore, mechanisms by which nucleoside triphosphates, nature's chosen monomers and universal energy currency, could have formed on the early Earth are even more poorly understood than the formation of nucleoside monophosphates – the most promising methods all employ nucleoside monophosphates as a starting point,[46,47] but such processes compete with formation of inorganic polyphosphates,[48] and leaves the question of how nucleoside monophosphates could form in the first place an open question. Amino-nucleosides provide a potential way around this problem due to the increased nucleophilicity of the amine moiety relative to the solvent water.[49] We reasoned that if an amino-nucleoside were exposed to a condensed phosphate species nucleoside polyphosphates could form directly and selectively in water.

We elected to study the interaction between *threo*-**20** and trimetaphosphate, $(PO_3Na)_3$.[50] In neutral water at room temperature no reaction was seen to occur over 5 days, so we investigated the effect of pH. If the phosphorylation is initiated in alkaline solution (pD 12) with 6 equiv. $(PO_3Na)_3$ 50% nucleoside triphosphate *threo*-**22** is observed to form over 25 days (Supplementary Figure 85). Over this period the solution also fell to pD 9, so we next initiated the phosphorylation at pD 9 in the presence of a divalent metal ion (Mg^{2+}) to accelerate the reaction and observed 60% *threo*-**22** form after only 5 days (Figure 5c & Supplementary Figure 82).

We next exposed *threo*-**22** to sulfur-activating agents (such as H_2O_2) and found that hydrolysis to *threo*-**23** and *threo*-**10** occurs, demonstrating that phosphorylation of the 3′-NH_2 enables selective hydrolysis to occur in place of the cyclisation that is observed without phosphorylation (Figure 5d & Supplementary Figure 86). Further hydrolysis of the cytosine base to uracil can also occur to give both canonical pyrimidine bases, a process which is accelerated by heating in phosphate buffer (Figure 5e & Supplementary Figure 88). *threo*-**10** undergoes hydrolysis (34% after 40 h) to furnish *threo*-**25** at 40 °C in 100 mM phosphate at neutral pH. Uridine *threo*-**25** is observed to also undergo reversible addition of its 3′-amine to the C6-carbon of the uracil base when heated to yield **26**.

Summary and outlook

We have found that regio- and stereo-selective incorporation of an amine into a nucleotide sugar moiety can be achieved with prebiotically plausible reactions. 3'-amino-TNA NTPs are formed readily from two and three carbon building blocks. The phosphate-mediated diastereoselectivity of these reactions, for *threo*-nucleotides, are unparalleled in prebiotically plausible nucleotide syntheses, and the 3'-amino-isomer that is exclusively formed has been shown to form more stable duplexes with both itself and with RNA than the 2'-amino-isomer.[12] The ease and selectivity for the formation of 3′-amino-TNA monomers suggests that they could have been present on the early Earth. Indeed, our results suggest that 3′-amino-TNA nucleosides are no more "generationally" complex than RNA nucleosides, in stark contrast to the predictions made by Eschenmoser.[3,15] Our initial expectation, at the outset of our work, was that selective 2′-alcohol phosphorylation would be required for 3′-amino-TNA synthesis. However, guided by the reactivity of the molecules, we have discovered that these amino-nucleosides react selectively with cTMP in water to yield NTPs.

Amino-nucleotides have been employed experimentally to overcome the issues plaguing non-enzymatic replication of RNA; despite decades of excellent work,[51,1] RNA-templated primer extensions using all four nucleobases and chemically activated ribonucleotides, has reached a copying limit of about 7 nucleotides.[52] The increased nucleophilicity of amine moieties compared with hydroxyls enables superior templated ligations,[53,54,55,56,57,58,59] but as non-biological molecules amino-nucleotides have so far been discussed explicitly as model substrates with no prebiotic relevance. Our results challenge this view. The 3′-amino-TNA NTPs we have discovered do not possess a free amine group, and so would not necessarily be expected to be more nucleophilic than ribonucleotides, but unlike ribonucleosides they are readily and selectively phosphorylated in water.

How ribonucleoside triphosphates (or other activated ribonucleoside-5′-phosphates) could have formed under aqueous conditions is an unsolved question. Phosphorylation of ribonucleotides with $(PO_3Na)_3$ in water not only requires extremely alkaline conditions ($pH > 12$), but also results in a mixture of deactivated 2′- and 3′-monophosphates.[60] In comparison 3′-amino-TNA nucleosides can be selectively triphosphorylated in water. It is worth noting that while chemical (phosphorus-anhydride) activation towards ligation[61,62] is retained, polymerisation of these amido-triphosphates has not yet been studied, and that although triphosphates are universally used as activated nucleotide monomers in biology, as well as being substrates for (ribozyme) polymerases, nucleoside triphosphates have been less well studied as substrates in non-enzymatic replication than the experimentally more convenient phosphorimidazolides.[52,59] This is due to their relatively slow (albeit more selective) ligation,[63] as well as the apparent difficulties of their synthesis on the early Earth. We hope that our discovery of a selective nucleoside triphosphate synthesis will promote further exploration of this space.

Overall, our results suggest 3′-amino-TNA triphosphates appear to be generationally *simpler* than the triphosphates of RNA or DNA. When combined with the known performance of 3′-amino-TNA in self and interspecies Watson-Crick base pairing with D/RNA, this may suggest these molecules played a role in the origins of life. Significant gaps remain in our understanding of how best to drive and control prebiotic (non-enzymatic) nucleotide ligation and replication for all nucleic acids. The implications of 3'-amino-*threo*-cytidine and uridine synthesis, and the scenario for continuous (in situ) monomer phosphorylation and activation, opened by their

enhanced reactivity in water, will be the object of further studies and future comparisons of the generational simplicity of RNA with 3'-amino-TNA.

Acknowledgements The Leverhulme Trust (RPG-2019-214 MWP), the Simons Foundation (318881FY19 MWP) and the Engineering and Physical Sciences Research Council (EP/P020410/1 MWP) provided financial support. We thank K. Karu and M. Puchnarewicz (Mass Spectrometry), M. Corpinot and K. Bucar (Crystallography) and A. E. Aliev (NMR spectroscopy).

Author contributions MWP conceived the research. MWP and DW designed and analysed the experiments and wrote the paper. DW conducted the experiments.

Competing Interests Statement The authors declare no competing financial interests.

Table 1 | Three-component coupling of aldehyde 1, oxazole 3 and aminonitrile 6 to yield amino-nucleotide precursor 7 in water. Optimisation of 3-component coupling of glycolaldehyde (**1**; 0.2 M) and 2-aminooxazole (**3**; 0.2 M) with aminonitrile (**6**; 0.6 M) at room temperature. [a] 10 mM **1**, 10 mM **3** and 30 mM **6**. [b] room temperature for 1 day, then 60 °C for 14 days. R' = H or CHRCN

Figure 1 | Intersections of prebiotic peptide, sugar and nucleobase syntheses. The three different classes of molecule presented (i.e. peptides, nucleotides and amino-nucleotides) are for clarity demarcated by coloured boxes, whilst glycolaldehyde **1**, which is a precursor and synthetic node in all intersecting syntheses shown, is positioned centrally within a white box. Glycolaldehyde **1** is a prebiotic precursor to **a**, peptides via serine nitrile **6** (R=CH_2OH), **b**, nucleotides via 2-aminooxazole **3** and its masked-aldol reaction with glyceraldehyde **4**, and **c**, amino-nucleotides **10** via 3-component Mannich reactivity.

Figure 2 | Formation of amino-sugar derivatives from C2 and C3 sugars. **a**, Uncontrolled aldol and Amadori reactions between ammonia and sugars proposed by Eschenmoser to yield a complex mixture of amino-sugars.[15] **b**, Selective 3-component coupling between 2-aminooxazole **3**, glycolaldehyde **1** and purine precursor, AICA at pH 5 to form tricyclic product **11**.[27] **c**, Selective 2-component coupling between 2-aminooxazole **3** and glycolaldehyde **1** in the presence of AICA to form aminooxazoline **5** at

pH 7.[27] **d**, Unselective incorporation of amino acids into a mixture of aminooxazolines, including **5**, via coupling of glyceraldehyde **4** with 2-aminooxazole **3** and proline.[28]

Figure 3 | Stereochemical resolution and crystallisation of *threo*-7. The precipitation of *threo*-**7** and rearrangement of *erythro*-**7** to **14** in phosphate buffer are observed to result in stereoselective synthesis of *threo*-**7**. ^{1}H NMR spectra to show the progression of stereochemical resolution upon addition of phosphate to *threo*-**7** and *erthro*-**7**: **a**, [400 MHz, D_2O, zg30, 7.0–3.3 ppm] a mixture of *threo*-**7** (blue) and *erythro*-**7** (orange) (200 mM, 3:2) in D_2O at pD 6.5 without addition of phosphate; **b**, [400 MHz, H_2O/D_2O 9:1, noesygppr1d, 7.0–3.3 ppm] mixture of *threo*-**7** (blue) and *erythro*-**7** (orange) (200 mM, 3:2) in 200 mM phosphate at pH 6.5, soluble fraction after 1 h; **c**, [400 MHz, H_2O/D_2O 9:1, noesygppr1d, 7.0–3.3 ppm] soluble fraction after 20 h showing formation of **14** (green). **d**, Crystal structure of *threo*-**7**·H_3PO_4; grey = carbon, white = hydrogen, red = oxygen, blue = nitrogen and orange = phosphorus. CCDC 2087673. N.B. All structures are racemates, L-enantiomers shown for clarity.

Figure 4 | Formation of anhydrocytidine *threo*-9 and its photochemical products. a, Reaction of *threo*-**7** with cyanoacetylene **8** yields anhydronucleoside *threo*-**9** and aldehyde **16**. **b**, Hydrolysis of *threo*-**9** to *threo*-**17** and subsequent photochemical formation of *threo*-**13** and cytosine **18**; **c**, Thiolysis of *threo*-**9** to *threo*-**19** and subsequent photochemical anomerisation of *threo*-**19** to *threo*-**20**, along with thiocytosine **21**. R=Me or H; **d**, ^{1}H NMR [700 MHz, H_2O/D_2O 9:1, noesygppr1d, 5.5–8.5 ppm] spectra showing the photoanomerisation of *threo*-**19** to yield *threo*-**20** and **21** on irradiation at 245 nm in water. N.B. All structures are racemates, L-enantiomers shown for clarity, and intermediate *threo*-**15** is observed to be a mixture of *E*- and *Z*-cyanovinyl isomers.

Figure 5 | Aqueous phosphorylation of amino-nucleoside *threo*-20 and hydrolysis to canonical nucleobases. a, Phosphorylation of *threo*-**20** with cyclic trimetaphosphate $(PO_3Na)_3$ in water catalysed by $MgCl_2$, and subsequent oxidative hydrolysis to cytosine-containing nucleotide *threo*-**23**. ^{1}H NMR [700 MHz, D_2O, 3.25–8.25 ppm] spectra showing the reaction of 50 mM *threo*-**20**, 6 equiv. $(PO_3Na)_3$, 6 equiv. $MgCl_2$, pD 9 after **b**, 1 h; **c**, 5 days and **d**, subsequent reaction of solution shown in spectrum **c** with 3 equiv. H_2O_2 resulting in the formation of *threo*-**23**, *threo*-**10** and **24**. Note that the apparent change in multiplicity at C6-H is due to partial deuteration at C5-H. **e**, Phosphate catalysed hydrolysis of cyditine *threo*-**10** to uridine *threo*-**25**.

Extended Data Figure 1 | Three-component coupling of aldehyde 1, oxazole 3 and aminonitrile 6e to yield amino-nucleotide precursor 7 in water. ^{1}H NMR spectra [400 MHz, H_2O/D_2O (9:1),

7.6–5.2 ppm] to show the: **a**, formation of oxazoline **12e** (R = ^{s}Bu) after 2 h at room temperature and pH 4.5 and **b**, retro-Strecker of **12e** at room temperature and pH 4.5 to form **7** after 5 days.

References

[1] Szostak, J. W. The Narrow Road to the Deep Past: In Search of the Chemistry of the Origin of Life. *Angew. Chem. Int. Ed.* **56**, 11037–11043 (2017).
[2] Robertson, M. P. & Joyce, G. F. The origins of the RNA World. *Cold Spring Harb. Perspect. Biol.* **4**, 1 (2012).
[3] Schöning, K.-U. *et al.* The -L-Threofuranosyl-(3'→2')-oligonucleotide System ('TNA'): Synthesis and Pairing Properties. *Helv. Chim. Acta* **85**, 4111–4153 (2002).
[4] Bhowmik, S. & Krishnamurthy, R. The role of sugar-backbone heterogeneity and chimeras in the simultaneous emergence of RNA and DNA. *Nat. Chem.* **11**, 1009–1018 (2019).
[5] Joyce, G. F., Schwartz, A. W., Miller, S. L. & Orgel, L. E. The case for an ancestral genetic system involving simple analogues of the nucleotides. *Proc. Natl. Acad. Sci. U. S. A.* **84**, 4398–4402 (1987).
[6] Fialho, D. M., Roche, T. P. & Hud, N. V. Prebiotic Syntheses of Noncanonical Nucleosides and Nucleotides. *Chem. Rev.* **120**, 4806–4830 (2020).
[7] Becker, S., Schneider, C., Crisp, A. & Carell, T. Non-canonical nucleosides and chemistry of the emergence of life. *Nat. Commun.* **9**, 5174 (2018).
[8] Hud, N. V. Searching for lost nucleotides of the pre-RNA World with a self-refining model of early Earth. *Nat. Commun.* **9**, 5171 (2018).
[9] Probst, A. V., Dunleavy, E. & Almouzni, G. Epigenetic inheritance during the cell cycle. *Nat. Rev. Mol. Cell Biol.* **10**, 192–206 (2009).
[10] Beier, M., Reck, F., Wagner, T., Krishnamurthy, R. & Eschenmoser, A. Chemical Etiology of Nucleic Acid Structure: Comparing Pentopyranosyl-(2'→4') Oligonucleotides with RNA. *Science* **283**, 699–703 (1999).
[11] Colville, B. W. F. & Powner, M. W. Selective prebiotic synthesis of α-threofuranosyl cytidine by photochemical anomerization. *Angew. Chem. Int. Ed.* **60**, 10526-10530 (2021).
[12] Wu, X., Guntha, S., Ferencic, M., Krishnamurthy, R. & Eschenmoser, A. Base-Pairing Systems Related to TNA: α-Threofuranosyl Oligonucleotides Containing Phosphoramidate Linkages. *Org. Lett.* **4**, 1279–1282 (2002
[13] Wang, Y. *et al.* A Threose Nucleic Acid Enzyme with RNA Ligase Activity. *J. Am. Chem. Soc.* **143**, 8154–8163 (2021).
[14] Yu, H., 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).
[15] Eschenmoser, A. The TNA-Family of Nucleic Acid Systems: Properties and Prospects. *Orig. Life Evol. Biosph.* **34**, 277–306 (2004).
[16] Whitaker, D. & Powner, M. W. Prebiotic nucleic acids need space to grow. *Nat. Commun.* **9**, 5172 (2018).
[17] Islam, S. & Powner, M. W. Prebiotic Systems Chemistry: Complexity Overcoming Clutter. *Chem* **2**, 470–501 (2017).
[18] Yadav, M., Kumar, R. & Krishnamurthy, R. Chemistry of Abiotic Nucleotide Synthesis. *Chem. Rev.* **120**, 4766–4805 (2020).
[19] Becker, S. *et al.* Unified prebiotically plausible synthesis of pyrimidine and purine RNA ribonucleotides. *Science* **366**, 76–82 (2019).
[20] Powner, M. W., Gerland, B. & Sutherland, J. D. Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. *Nature* **459**, 239–242 (2009).
[21] Stairs, S. *et al.* Divergent prebiotic synthesis of pyrimidine and 8-oxo-purine ribonucleotides. *Nat. Commun.* **8**, 15270 (2017).
[22] Islam, S., Bučar, D.-K. & Powner, M. W. Prebiotic selection and assembly of proteinogenic amino acids and natural nucleotides from complex mixtures. *Nat. Chem.* **9**, 584–589 (2017).
[23] Ponnamperuma, C. & Woeller, F. α-Aminonitriles formed by an electric discharge through a mixture of anhydrous methane and ammonia. *Currents in Moden Biology 1*, 156–158 (1967).
[24] Patel, B. H., Percivalle, C., Ritson, D. J., D., D. & Sutherland, J. D. Common origins of RNA, protein and lipid precursors in a cyanosulfidic protometabolism. *Nat Chem* **7**, 301–307 (2015).

[25] Canavelli, P., Islam, S. & Powner, M. W. Peptide ligation by chemoselective aminonitrile coupling in water. *Nature* **571**, 546–549 (2019).
[26] Foden, C. S. *et al.* Prebiotic synthesis of cysteine peptides that catalyze peptide ligation in neutral water. *Science.* **370**, 865–869 (2020).
[27] Powner, M. W., Sutherland, J. D. & Szostak, J. W. Chemoselective Multicomponent One-Pot Assembly of Purine Precursors in Water. *J. Am. Chem. Soc.* **132**, 16677–16688 (2010).
[28] Hein, J. E., Tse, E. & Blackmond, D. G. A route to enantiopure RNA precursors from nearly racemic starting materials. *Nat. Chem.* **3**, 704–706 (2011).
[29] Powner, M. W. & Sutherland, J. D. Phosphate-Mediated Interconversion of Ribo- and Arabino-Configured Prebiotic Nucleotide Intermediates. *Angew. Chem. Int. Ed.* **49**, 4641–4643 (2010).
[30] Ashe, K. *et al.* Selective prebiotic synthesis of phosphoroaminonitriles and aminothioamides in neutral water. *Commun. Chem.* **2**, 23 (2019).
[31] Moutou, G. *et al.* Equilibrium of α-aminoacetonitrile formation from formaldehyde, hydrogen cyanide and ammonia in aqueous solution: Industrial and prebiotic significance. *J. Phys. Org. Chem.* **8**, 721–730 (1995).
[32] Taillades, J. & Commeyras, A. Systemes de strecker et apparentes—I Etude de la decomposition en solution aqueuse des α-alcoyl-aminonitriles tertiaires. Mécanisme d'élimination du groupement nitrile. *Tetrahedron* **30**, 127–132 (1974).
[33] Taillades, J. *et al.* N-carbamoyl-α-amino acids rather than free α-amino acids formation in the primitive hydrosphere: A novel proposal for the emergence of prebiotic peptides. *Orig. Life Evol. Biosph.* **28**, 61–77 (1998).
[34] Springsteen, G. & Joyce, G. F. Selective derivatization and sequestration of ribose from a prebiotic mix. *J. Am. Chem. Soc.* **126**, 9578–9583 (2004).
[35] Ferris, J. P., Sanchez, R. a & Orgel, L. E. Studies in prebiotic synthesis: III. Synthesis of pyrimidines from cyanoacetylene and cyanate. *J. Mol. Biol.* **33**, 693–704 (1968).
[36] Ni, G. *et al.* Review of α-nucleosides: From discovery, synthesis to properties and potential applications. *RSC Adv.* **9**, 14302–14320 (2019).
[37] Sanchez, R. A. & Orgel, L. E. Studies in prebiotic synthesis V. Synthesis and photoanomerization of pyrimidine nucleosides. *J. Mol. Biol.* **47**, 531–543 (1970).
[38] Xu, J. *et al.* A prebiotically plausible synthesis of pyrimidine β-ribonucleosides and their phosphate derivatives involving photoanomerization. *Nat. Chem.* **9**, 303–309 (2017).
[39] Grosjean, H., de Crécy-Lagard, V. & Marck, C. Deciphering synonymous codons in the three domains of life: Co-evolution with specific tRNA modification enzymes. *FEBS Lett.* **584**, 252–264 (2010).
[40] Ajitkumar, P. & Cherayil, J. D. Thionucleosides in transfer ribonucleic acid: diversity, structure, biosynthesis, and function. *Microbiol. Rev.* **52**, 103–113 (1988).
[41] Testa, S. M., Disney, M. D., Turner, D. H. & Kierzek, R. Thermodynamics of RNA-RNA duplexes with 2- or 4-thiouridines: Implications for antisense design and targeting a group I intron. *Biochemistry* **38**, 16655–16662 (1999).
[42] Heuberger, B. D., Pal, A., Del Frate, F., Topkar, V. V. & Szostak, J. W. Replacing uridine with 2-thiouridine enhances the rate and fidelity of nonenzymatic RNA primer extension. *J. Am. Chem. Soc.* **137**, 2769–2775 (2015).
[43] Prywes, N., Michaels, Y. S., Pal, A., Oh, S. S. & Szostak, J. W. Thiolated uridine substrates and templates improve the rate and fidelity of ribozyme-catalyzed RNA copying. *Chem. Commun.* **52**, 6529–6532 (2016).
[44] Ohkubo, A. *et al.* Formation of new base pairs between inosine and 5-methyl-2-thiocytidine derivatives. *Org. Biomol. Chem.* **10**, 2008–2010 (2012).
[45] Kim, S. C., O'Flaherty, D. K., Zhou, L., Lelyveld, V. S. & Szostak, J. W. Inosine, but none of the 8-oxo-purines, is a plausible component of a primordial version of RNA. *Proc. Natl. Acad. Sci. USA* **115**, 13318–13323 (2018).
[46] 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).
[47] Lohrmann, R. Formation of nucleoside 5′-polyphosphates from nucleotides and trimetaphosphate. *J. Mol. Evol.* **6**, 237–252 (1975).
[48] Yamagata, Y. Prebiotic formation of ADP and ATP from AMP, calcium phosphates and cyanate in aqueous solution. *Orig. Life Evol. Biosph.* **29**, 511–520 (1999).
[49] Feldmann, W. & Thilo, E. Zur Chemie der kondensierten Phosphate und Arsenate. XXXVIII. Amidotriphosphat. *Zeitschrift für Anorg. und Allg. Chemie* **328**, 113–126 (1964).

[50] Yamagata, Y., Watanabe, H., Saitoh, M. & Namba, T. Volcanic production of polyphosphates and its relevance to prebiotic evolution. *Nature* **352**, 516–519 (1991).
[51] Weimann, B. J., Lohrmann, R., Orgel, L. E., Schneider-Bernloehr, H. & Sulston, J. E. Template-Directed Synthesis with Adenosine-5′-phosphorimidazolide. *Science* **161**, 387–387 (1968).
[52] Li, L. *et al.* Enhanced nonenzymatic RNA copying with 2-aminoimidazole activated nucleotides. *J. Am. Chem. Soc.* **139**, 1810–1813 (2017).
[53] Zielinski, W. S. & Orgel, L. E. Oligomerization of activated derivatives of 3′-amino-3′-deoxyguanosine on poly(C) and poly(dC) templates. *Nucleic Acids Res.* **13**, 2469–2484 (1985).
[54] Röthlingshöfer, M. *et al.* Chemical primer extension in seconds. *Angew. Chem. Int. Ed.* **47**, 6065–6068 (2008).
[55] Chen, J. J., Cai, X. & Szostak, J. W. N2′→P3′ Phosphoramidate Glycerol Nucleic Acid as a Potential Alternative Genetic System. *J. Am. Chem. Soc.* **131**, 2119–2121 (2009).
[56] Blain, J. C., Ricardo, A. & Szostak, J. W. Synthesis and Nonenzymatic Template-Directed Polymerization of 2′-Amino-2′-deoxythreose Nucleotides. *J. Am. Chem. Soc.* **136**, 2033–2039 (2014).
[57] Zhou, L., O'Flaherty, D. K. & Szostak, J. W. Template-Directed Copying of RNA by Non-enzymatic Ligation. *Angew. Chem. Int. Ed.* **59**, 15682–15687 (2020).
[58] Hänle, E. & Richert, C. Enzyme-Free Replication with Two or Four Bases. *Angew. Chem. Int. Ed.* **57**, 8911–8915 (2018).
[59] O'Flaherty, D. K., Zhou, L. & Szostak, J. W. Nonenzymatic Template-Directed Synthesis of Mixed-Sequence 3′-NP-DNA up to 25 Nucleotides Long Inside Model Protocells. *J. Am. Chem. Soc.* **141**, 10481–10488 (2019).
[60] Saffhill, R. Selective Phosphorylation of the cis-2′,3′-Diol of Unprotected Ribonucleosides with Trimetaphosphate in Aqueous Solution. *J. Org. Chem.* **35**, 2881–2883 (1970).
[61] Mullen, L. B. & Sutherland, J. D. Formation of potentially prebiotic amphiphiles by reaction of β-hydroxy-n-alkylamines with cyclotriphosphate. *Angew. Chem. Int. Ed.* **46**, 4166–4168 (2007).
[62] Krishnamurthy, R., Guntha, S. & Eschenmoser, A. Regioselective α-Phosphorylation of Aldoses in Aqueous Solution. *Angew. Chem. Int. Ed.* **39**, 2281–2285 (2000).
[63] Rohatgi, R., Bartel, D. P. & Szostak, J. W. Nonenzymatic, Template-Directed Ligation of Oligoribonucleotides Is Highly Regioselective for the Formation of 3′−5′ Phosphodiester Bonds. *J. Am. Chem. Soc.* **118**, 3340–3344 (1996).

Methods

Caution statement

Cyanide and (hydro)sulfide are highly toxic poisons by inhalation, contact, and ingestion. They generate poisonous hydrogen cyanide (pK_a = 9.2) and hydrogen sulfide (pK_a = 7.1) gas at neutral or acidic pH. Solutions containing cyanide or (hydro)sulfide, or compounds which may generate these must be handled in a well-ventilated fumehood equipped with appropriate chemical quenches, such as sodium hypochlorite (bleach) or iron(II) sulfate solution. Read and follow the material safety data sheet (MSDS) instructions for personnel handling, exposure, and disposal information.

General procedure for 3-component coupling to form diaminooxazolines

Glycolaldehyde **1** (12 mg, 0.2 mmol), 2-aminooxazole **3** (17 mg, 0.2 mmol) and aminonitrile hydrochloride **6**.HCl (0.6 mmol) were dissolved in H_2O (800 μL), and the solution was adjusted to pH

5 with 4 M NaOH and 4 M HCl. Methylsulfonylmethane [internal standard] (100 µL of a 0.083 M solution in D_2O) was added, then the solution volume was increased to 1 mL with H_2O. ^1H NMR spectra were then obtained periodically.

General procedure for the reaction of *threo*-7 with cyanoacetylene 8

threo-**7**.H_3PO_4 (0.05 mmol) was suspended in H_2O (400 µL). The solution was adjusted to pH 8.5 to ensure full dissolution of the salt, then the solution was adjusted to the desired pH with 4 M NaOH or 4 M HCl. Cyanoacetylene **8** (0.5 mL of a 1 M solution in H_2O, 10 equiv.) was added, followed by methylsulfonyl methane [internal standard] (100 µL of a 0.083 M solution in D_2O, 0.0083 mmol). The volume of the resultant solution was increased to 1 mL with H_2O and NMR spectra were obtained periodically.

General procedure for the thiolysis of *threo*-9 in formamide

threo-**9**.2HCl (50 mg, 0.2 mmol) and NaSH.xH_2O (110 mg, 1 mmol) were sealed under argon in a flame-dried flask. Anhydrous formamide (2 mL) was added and the yellow suspension was stirred at room temperature for 3 days. The reaction was quenched with HCl (0.04 M, 10 mL), and argon was bubbled through the suspension into a bleach trap for 30 mins to remove excess H_2S. The suspension was concentrated under reduced pressure, and co-evaporated with D_2O (3×3 mL – to partially deuterate the formamide NH_2) then mixed with D_2O (3 mL) and analysed by NMR spectroscopy.

General procedure for the photochemical anomerisation of *threo*-19

threo-**19** (2.4 mg, 0.01 mmol) and NaSH.xH_2O (2.6 mg, 0.02 mmol) dissolved in H_2O (0.45 mL), and the solution was adjusted to pH 7 with 4 M HCl. Methyl sulfonylmethane [internal standard] (40 µL of a 0.083 M solution in D_2O, 0.003 mmol) was added then the solution was transferred to a quartz NMR tube and irradiated with UV light (principal emission at 254 nm). NMR spectra were obtained periodically.

General procedure for the aqueous triphosphorylation of *threo*-20

threo-**20** (0.02 mmol), cyclic trimetaphosphate $(PO_3Na)_3$ (36 mg, 0.12 mmol, 6 equiv.) and $MgCl_2$ (6 mg, 0.06 mmol) were suspended in D_2O (0.4 mL), the solution was adjusted to pD 9 with 4 M NaOD, and then NMR spectra were obtained periodically. After 5 days the solution had fallen to pD 7.8. The solution was adjusted to pD 9 with 4 M NaOD, and then NMR spectra were again acquired periodically.

Data Availability All data (experimental procedures and characterization data) supporting the findings of this study are available within the article and its Supplementary Information.

Crystallographic data for the structure reported in this Article has been deposited at the Cambridge Crystallographic Data Centre (CCDC), under deposition numbers: 2087673. Copies of the data can be obtained free of charge from CCDC via https://www.ccdc.cam.ac.uk/structures/

---

[1] Szostak, J. W. The Narrow Road to the Deep Past: In Search of the Chemistry of the Origin of Life. *Angew. Chem. Int. Ed.* **56**, 11037–11043 (2017).
[2] Robertson, M. P. & Joyce, G. F. The origins of the RNA World. *Cold Spring Harb. Perspect. Biol.* **4**, 1 (2012).
[3] Schöning, K.-U. *et al.* The -L-Threofuranosyl-(3'→2')-oligonucleotide System ('TNA'): Synthesis and Pairing Properties. *Helv. Chim. Acta* **85**, 4111–4153 (2002).
[4] Bhowmik, S. & Krishnamurthy, R. The role of sugar-backbone heterogeneity and chimeras in the simultaneous emergence of RNA and DNA. *Nat. Chem.* **11**, 1009–1018 (2019).
[5] Joyce, G. F., Schwartz, A. W., Miller, S. L. & Orgel, L. E. The case for an ancestral genetic system involving simple analogues of the nucleotides. *Proc. Natl. Acad. Sci. U. S. A.* **84**, 4398–4402 (1987).
[6] Fialho, D. M., Roche, T. P. & Hud, N. V. Prebiotic Syntheses of Noncanonical Nucleosides and Nucleotides. *Chem. Rev.* **120**, 4806–4830 (2020).
[7] Becker, S., Schneider, C., Crisp, A. & Carell, T. Non-canonical nucleosides and chemistry of the emergence of life. *Nat. Commun.* **9**, 5174 (2018).
[8] Hud, N. V. Searching for lost nucleotides of the pre-RNA World with a self-refining model of early Earth. *Nat. Commun.* **9**, 5171 (2018).
[9] Probst, A. V., Dunleavy, E. & Almouzni, G. Epigenetic inheritance during the cell cycle. *Nat. Rev. Mol. Cell Biol.* **10**, 192–206 (2009).
[10] Beier, M., Reck, F., Wagner, T., Krishnamurthy, R. & Eschenmoser, A. Chemical Etiology of Nucleic Acid Structure: Comparing Pentopyranosyl-(2'→4') Oligonucleotides with RNA. *Science* **283**, 699–703 (1999).
[11] Colville, B. W. F. & Powner, M. W. Selective prebiotic synthesis of α-threofuranosyl cytidine by photochemical anomerization. *Angew. Chem. Int. Ed.* **60**, 10526-10530 (2021).
[12] Wu, X., Guntha, S., Ferencic, M., Krishnamurthy, R. & Eschenmoser, A. Base-Pairing Systems Related to TNA: α-Threofuranosyl Oligonucleotides Containing Phosphoramidate Linkages. *Org. Lett.* **4**, 1279–1282 (2002
[13] Wang, Y. *et al.* A Threose Nucleic Acid Enzyme with RNA Ligase Activity. *J. Am. Chem. Soc.* **143**, 8154–8163 (2021).
[14] Yu, H., 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).
[15] Eschenmoser, A. The TNA-Family of Nucleic Acid Systems: Properties and Prospects. *Orig. Life Evol. Biosph.* **34**, 277–306 (2004).
[16] Whitaker, D. & Powner, M. W. Prebiotic nucleic acids need space to grow. *Nat. Commun.* **9**, 5172 (2018).
[17] Islam, S. & Powner, M. W. Prebiotic Systems Chemistry: Complexity Overcoming Clutter. *Chem* **2**, 470–501 (2017).
[18] Yadav, M., Kumar, R. & Krishnamurthy, R. Chemistry of Abiotic Nucleotide Synthesis. *Chem. Rev.* **120**, 4766–4805 (2020).
[19] Becker, S. *et al.* Unified prebiotically plausible synthesis of pyrimidine and purine RNA ribonucleotides. *Science* **366**, 76–82 (2019).

[20] Powner, M. W., Gerland, B. & Sutherland, J. D. Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. *Nature* **459**, 239–242 (2009).
[21] Stairs, S. *et al.* Divergent prebiotic synthesis of pyrimidine and 8-oxo-purine ribonucleotides. *Nat. Commun.* **8**, 15270 (2017).
[22] Islam, S., Bučar, D.-K. & Powner, M. W. Prebiotic selection and assembly of proteinogenic amino acids and natural nucleotides from complex mixtures. *Nat. Chem.* **9**, 584–589 (2017).
[23] Ponnamperuma, C. & Woeller, F. α-Aminonitriles formed by an electric discharge through a mixture of anhydrous methane and ammonia. *Currents in Moden Biology 1*, 156–158 (1967).
[24] Patel, B. H., Percivalle, C., Ritson, D. J., D., D. & Sutherland, J. D. Common origins of RNA, protein and lipid precursors in a cyanosulfidic protometabolism. *Nat Chem* **7**, 301–307 (2015).
[25] Canavelli, P., Islam, S. & Powner, M. W. Peptide ligation by chemoselective aminonitrile coupling in water. *Nature* **571**, 546–549 (2019).
[26] Foden, C. S. *et al.* Prebiotic synthesis of cysteine peptides that catalyze peptide ligation in neutral water. *Science.* **370**, 865–869 (2020).
[27] Powner, M. W., Sutherland, J. D. & Szostak, J. W. Chemoselective Multicomponent One-Pot Assembly of Purine Precursors in Water. *J. Am. Chem. Soc.* **132**, 16677–16688 (2010).
[28] Hein, J. E., Tse, E. & Blackmond, D. G. A route to enantiopure RNA precursors from nearly racemic starting materials. *Nat. Chem.* **3**, 704–706 (2011).
[29] Powner, M. W. & Sutherland, J. D. Phosphate-Mediated Interconversion of Ribo- and Arabino-Configured Prebiotic Nucleotide Intermediates. *Angew. Chem. Int. Ed.* **49**, 4641–4643 (2010).
[30] Ashe, K. *et al.* Selective prebiotic synthesis of phosphoroaminonitriles and aminothioamides in neutral water. *Commun. Chem.* **2**, 23 (2019).
[31] Moutou, G. *et al.* Equilibrium of α-aminoacetonitrile formation from formaldehyde, hydrogen cyanide and ammonia in aqueous solution: Industrial and prebiotic significance. *J. Phys. Org. Chem.* **8**, 721–730 (1995).
[32] Taillades, J. & Commeyras, A. Systemes de strecker et apparentes—I Etude de la decomposition en solution aqueuse des α-alcoyl-aminonitriles tertiaires. Mécanisme d'élimination du groupement nitrile. *Tetrahedron* **30**, 127–132 (1974).
[33] Taillades, J. *et al.* N-carbamoyl-α-amino acids rather than free α-amino acids formation in the primitive hydrosphere: A novel proposal for the emergence of prebiotic peptides. *Orig. Life Evol. Biosph.* **28**, 61–77 (1998).
[34] Springsteen, G. & Joyce, G. F. Selective derivatization and sequestration of ribose from a prebiotic mix. *J. Am. Chem. Soc.* **126**, 9578–9583 (2004).
[35] Ferris, J. P., Sanchez, R. a & Orgel, L. E. Studies in prebiotic synthesis: III. Synthesis of pyrimidines from cyanoacetylene and cyanate. *J. Mol. Biol.* **33**, 693–704 (1968).
[36] Ni, G. *et al.* Review of α-nucleosides: From discovery, synthesis to properties and potential applications. *RSC Adv.* **9**, 14302–14320 (2019).
[37] Sanchez, R. A. & Orgel, L. E. Studies in prebiotic synthesis V. Synthesis and photoanomerization of pyrimidine nucleosides. *J. Mol. Biol.* **47**, 531–543 (1970).
[38] Xu, J. *et al.* A prebiotically plausible synthesis of pyrimidine β-ribonucleosides and their phosphate derivatives involving photoanomerization. *Nat. Chem.* **9**, 303–309 (2017).
[39] Grosjean, H., de Crécy-Lagard, V. & Marck, C. Deciphering synonymous codons in the three domains of life: Co-evolution with specific tRNA modification enzymes. *FEBS Lett.* **584**, 252–264 (2010).
[40] Ajitkumar, P. & Cherayil, J. D. Thionucleosides in transfer ribonucleic acid: diversity, structure, biosynthesis, and function. *Microbiol. Rev.* **52**, 103–113 (1988).
[41] Testa, S. M., Disney, M. D., Turner, D. H. & Kierzek, R. Thermodynamics of RNA-RNA duplexes with 2- or 4-thiouridines: Implications for antisense design and targeting a group I intron. *Biochemistry* **38**, 16655–16662 (1999).
[42] Heuberger, B. D., Pal, A., Del Frate, F., Topkar, V. V. & Szostak, J. W. Replacing uridine with 2-thiouridine enhances the rate and fidelity of nonenzymatic RNA primer extension. *J. Am. Chem. Soc.* **137**, 2769–2775 (2015).

[43] Prywes, N., Michaels, Y. S., Pal, A., Oh, S. S. & Szostak, J. W. Thiolated uridine substrates and templates improve the rate and fidelity of ribozyme-catalyzed RNA copying. *Chem. Commun.* **52**, 6529–6532 (2016).
[44] Ohkubo, A. *et al.* Formation of new base pairs between inosine and 5-methyl-2-thiocytidine derivatives. *Org. Biomol. Chem.* **10**, 2008–2010 (2012).
[45] Kim, S. C., O'Flaherty, D. K., Zhou, L., Lelyveld, V. S. & Szostak, J. W. Inosine, but none of the 8-oxo-purines, is a plausible component of a primordial version of RNA. *Proc. Natl. Acad. Sci.* **115**, 13318–13323 (2018).
[46] 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).
[47] Lohrmann, R. Formation of nucleoside 5′-polyphosphates from nucleotides and trimetaphosphate. *J. Mol. Evol.* **6**, 237–252 (1975).
[48] Yamagata, Y. Prebiotic formation of ADP and ATP from AMP, calcium phosphates and cyanate in aqueous solution. *Orig. Life Evol. Biosph.* **29**, 511–520 (1999).
[49] Feldmann, W. & Thilo, E. Zur Chemie der kondensierten Phosphate und Arsenate. XXXVIII. Amidotriphosphat. *Zeitschrift für Anorg. und Allg. Chemie* **328**, 113–126 (1964).
[50] Yamagata, Y., Watanabe, H., Saitoh, M. & Namba, T. Volcanic production of polyphosphates and its relevance to prebiotic evolution. *Nature* **352**, 516–519 (1991).
[51] Weimann, B. J., Lohrmann, R., Orgel, L. E., Schneider-Bernloehr, H. & Sulston, J. E. Template-Directed Synthesis with Adenosine-5′-phosphorimidazolide. *Science* **161**, 387–387 (1968).
[52] Li, L. *et al.* Enhanced nonenzymatic RNA copying with 2-aminoimidazole activated nucleotides. *J. Am. Chem. Soc.* **139**, 1810–1813 (2017).
[53] Zielinski, W. S. & Orgel, L. E. Oligomerization of activated derivatives of 3′-amino-3′-deoxyguanosine on poly(C) and poly(dC) templates. *Nucleic Acids Res.* **13**, 2469–2484 (1985).
[54] Röthlingshöfer, M. *et al.* Chemical primer extension in seconds. *Angew. Chem. Int. Ed.* **47**, 6065–6068 (2008).
[55] Chen, J. J., Cai, X. & Szostak, J. W. N2′→P3′ Phosphoramidate Glycerol Nucleic Acid as a Potential Alternative Genetic System. *J. Am. Chem. Soc.* **131**, 2119–2121 (2009).
[56] Blain, J. C., Ricardo, A. & Szostak, J. W. Synthesis and Nonenzymatic Template-Directed Polymerization of 2′-Amino-2′-deoxythreose Nucleotides. *J. Am. Chem. Soc.* **136**, 2033–2039 (2014).
[57] Zhou, L., O'Flaherty, D. K. & Szostak, J. W. Template-Directed Copying of RNA by Non-enzymatic Ligation. *Angew. Chem. Int. Ed.* **59**, 15682–15687 (2020).
[58] Hänle, E. & Richert, C. Enzyme-Free Replication with Two or Four Bases. *Angew. Chem. Int. Ed.* **57**, 8911–8915 (2018).
[59] O'Flaherty, D. K., Zhou, L. & Szostak, J. W. Nonenzymatic Template-Directed Synthesis of Mixed-Sequence 3′-NP-DNA up to 25 Nucleotides Long Inside Model Protocells. *J. Am. Chem. Soc.* **141**, 10481–10488 (2019).
[60] Saffhill, R. Selective Phosphorylation of the cis-2′,3′-Diol of Unprotected Ribonucleosides with Trimetaphosphate in Aqueous Solution. *J. Org. Chem.* **35**, 2881–2883 (1970).
[61] Mullen, L. B. & Sutherland, J. D. Formation of potentially prebiotic amphiphiles by reaction of β-hydroxy-n-alkylamines with cyclotriphosphate. *Angew. Chem. Int. Ed.* **46**, 4166–4168 (2007).
[62] Krishnamurthy, R., Guntha, S. & Eschenmoser, A. Regioselective α-Phosphorylation of Aldoses in Aqueous Solution. *Angew. Chem. Int. Ed.* **39**, 2281–2285 (2000).
[63] Rohatgi, R., Bartel, D. P. & Szostak, J. W. Nonenzymatic, Template-Directed Ligation of Oligoribonucleotides Is Highly Regioselective for the Formation of 3′−5′ Phosphodiester Bonds. *J. Am. Chem. Soc.* **118**, 3340–3344 (1996).