nature chemistry

ARTICLES

https://doi.org/10.1038/s41557-020-00564-3

Harnessing chemical energy for the activation and joining of prebiotic building blocks

Ziwei Liu[1,2], Long-Fei Wu[1,2], Jianfeng Xu[1], Claudia Bonfio[1], David A. Russell[1] and John D. Sutherland[1]✉

Life is an out-of-equilibrium system sustained by a continuous supply of energy. In extant biology, the generation of the primary energy currency, adenosine 5′-triphosphate and its use in the synthesis of biomolecules require enzymes. Before their emergence, alternative energy sources, perhaps assisted by simple catalysts, must have mediated the activation of carboxylates and phosphates for condensation reactions. Here, we show that the chemical energy inherent to isonitriles can be harnessed to activate nucleoside phosphates and carboxylic acids through catalysis by acid and 4,5-dicyanoimidazole under mild aqueous conditions. Simultaneous activation of carboxylates and phosphates provides multiple pathways for the generation of reactive intermediates, including mixed carboxylic acid–phosphoric acid anhydrides, for the synthesis of peptidyl–RNAs, peptides, RNA oligomers and primordial phospholipids. Our results indicate that unified prebiotic activation chemistry could have enabled the joining of building blocks in aqueous solution from a common pool and enabled the progression of a system towards higher complexity, foreshadowing today's encapsulated peptide–nucleic acid system.

Short RNA oligomers or peptides could have formed on early Earth in dry state[1,2] or solution phase reactions[3,4] following the activation of nucleotide phosphates or amino acid/peptide carboxylates, respectively. While these reactions have been performed separately, the simultaneous synthesis of RNA oligomers and peptides under dry state conditions starting from mixtures of appropriate building blocks (monomers and short oligomers) has not yet been reported. In aqueous solution, high-energy molecules, such as cyanamide[5,6] and carbonyl sulfide[7,8], have been shown to drive condensation reactions of nucleotide phosphates and amino acids, respectively. However, although it can be released by outgassing due to geothermal activity, it is difficult to envisage how a reactive molecule like carbonyl sulfide could have accumulated on the surface of early Earth. Conversely, the kinetic stability of cyanamide that potentially allows its accumulation also makes it a sluggish activating agent. Although some form of catalysis might offer a solution to this problem[9,10], no combination of prebiotically plausible high-energy molecule and catalyst has been found to activate simultaneously phosphates and carboxylates in the solution phase. Whilst dry state and solution phase production of oligopeptides and oligonucleotides may have sufficed initially, solution phase RNA ligation/replication, peptide formation and peptidyl–RNA chemistry must have taken place to enable the eventual transition to biology according to the principle of continuity[11].

Methyl isonitrile has been detected in interstellar space[12], while Eschenmoser and co-workers[13] reported that it can also be produced by thermolysis of iminodiacetonitrile, which derives in short measure from hydrogen cyanide[14]. We have previously described an early Earth geochemical model that leads to methyl isonitrile–iron(II) complexes from ferrocyanide {$[Fe(CN)_6]^{4-}$}, nitric oxide and methylamine[15]. Methyl isonitrile could have been stored in the form of persistent Fe(II) complexes, such as $[Fe(CN)_5(CNCH_3)]^{3-}$ and released by ultraviolet irradiation[15]. Isonitriles can activate phosphate groups upon reaction with aldehydes or imines[15,16]. However, the activation of carboxylates using similar chemistry leads, via intramolecular acyl transfer, to esters and amides, respectively[17,18]. Accordingly, we wondered if the chemical energy of methyl isonitrile could be harnessed, under possible aldehyde- and imine-free conditions, to allow the simultaneous activation of prebiotic phosphate and carboxylate building blocks.

Results

Simultaneous activation of adenosine 3′-monophosphate and carboxylates, catalysis by 4,5-dicyanoimidazole and stereoselective acyl transfer. It has previously been shown that the hydrolysis of methyl isonitrile (**1**, $pK_{aH} \approx 1$) to *N*-methylformamide **2** is accelerated by carboxylic acids[19,20]. Therefore, we wondered if acid catalysis could induce the addition of carboxylate and phosphate nucleophiles to **1**. A model system was set up by mixing methyl isonitrile **1**, adenosine 3′-monophosphate (A3′P, **3**) and individual members selected from a library of acids, including carboxylic and amino acids, under mildly acidic aqueous conditions (Fig. 1a). Acetic acid (AcOH, $pK_a = 4.8$) promoted the formation of adenosine 2′,3′-cyclic monophosphate (A>P, **4**, yield of **4** was 56% versus 12% for blank after 24 h, Table 1, Entries 1 and 2, Supplementary Figs. 1 and 2), indicative of phosphate activation[5,9,10,16]. Free amino acids, such as glycine (Gly, $pK_a = 2.3$) exerted no obvious rate enhancement (11% yield of **4**, Table 1, Entry 3, Supplementary Fig. 3). The N-protected amino acid, *N*-acetyl-L-alanine (**Ac-L-Ala**, $pK_a = 3.8$), on the other hand, not only accelerated the formation of A>P **4** (in 18% yield), but also formed adenosine-2′-(*N*-acetyl-L-alanyl)-3′-monophosphate **5-L-a**, albeit in a yield of only 5% (Table 1, Entry 4, Supplementary Figs. 4 and 5). The formation of **5-L-a** proceeds via a mixed carboxylic acid–phosphoric acid anhydride intermediate **6** (refs. [21,22]), which can be observed using NMR spectroscopy by running the reaction at pH 4 (Table 1, Entry 5, Supplementary Figs. 6 and 7) and along with the concurrent formation of **4** is indicative of simultaneous activation of carboxylate and phosphate.

Encouraged by these findings, we investigated whether nucleophilic catalysis might also contribute to carboxylate activation and

[1]MRC Laboratory of Molecular Biology, Cambridge Biomedical Campus, Cambridge, UK. [2]These authors contributed equally: Ziwei Liu, Long-Fei Wu. ✉e-mail: johns@mrc-lmb.cam.ac.uk

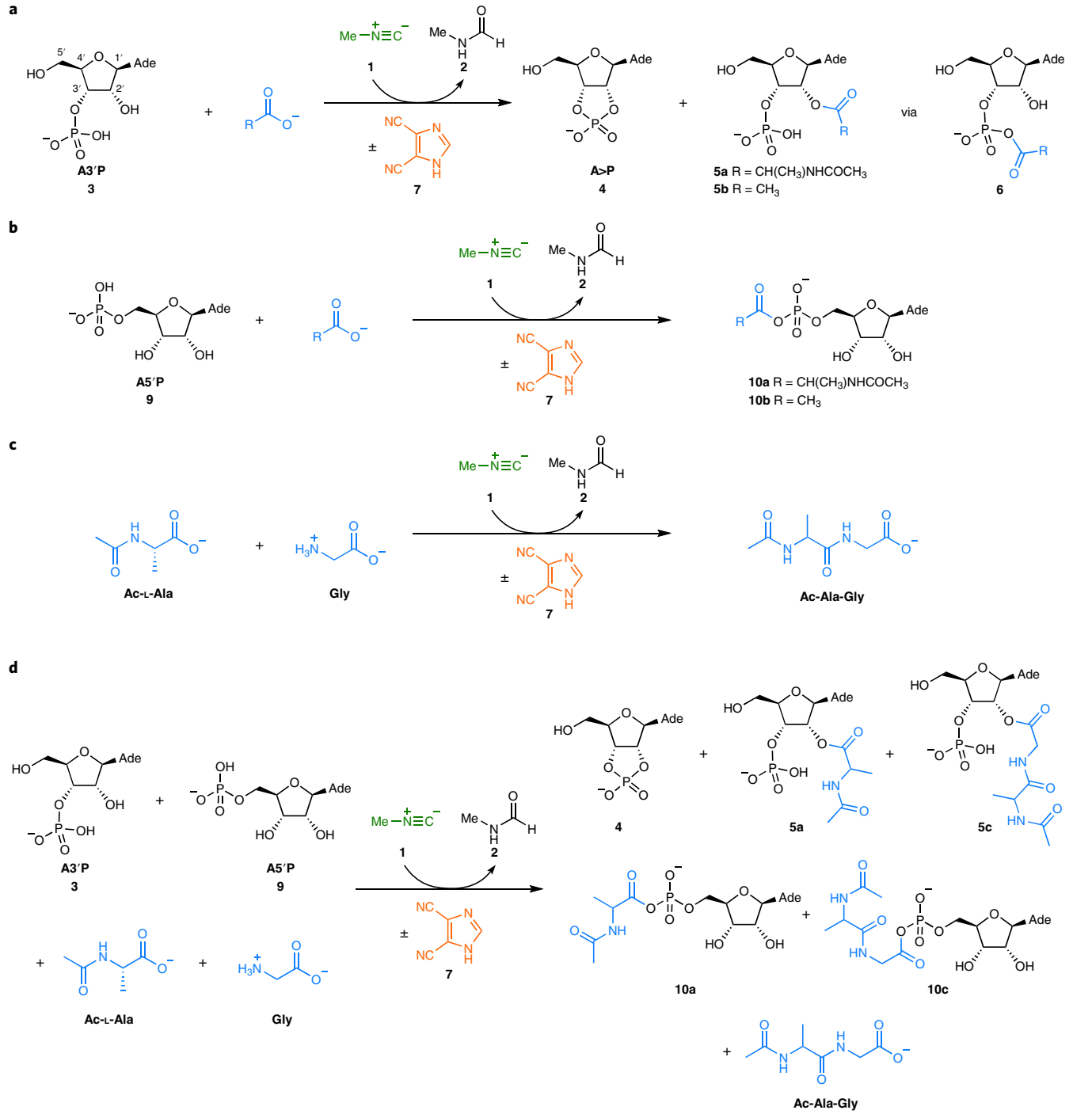

Fig. 1 | Joining of prebiotic building blocks driven by common activation chemistry in aqueous solution. a, Model reaction of A3′P **3** and carboxylates upon activation. The mixed anhydride intermediate **6** can give either **4** or **5**. **b**, Model reaction of A5′P 9 and carboxylates upon activation. The mixed anhydride **10** is reminiscent of structurally analogous intermediates common to peptide and phospholipid biosynthesis. **c**, Formation of peptides upon activation. **d**, Simultaneous activation of carboxylates and phosphates under unified activation conditions. Carboxylates (aliphatic, peptidic and amino acids) are highlighted in blue, the catalyst (DCI **7**) is shown in orange and the activating reagent, methyl isonitrile **1**, is shown in green. Ade: adenine.

transfer on a wider range of substrates, including free and protected amino acids, peptides and aliphatic acids of different chain lengths. Based on the well-established nucleophilic catalysis of acyl transfer reactions by imidazole[23,24], we tested several prebiotically plausible azoles, including imidazole[25] ($pK_{aH}=7$), 5-aminoimidazole-4-carboxamide[26] (AICA, $pK_{aH}=5.2$), 2-aminoimidazole[27] (2-AI, $pK_{aH}=8.5$) and 4,5-dicyanoimidazole (DCI, **7**, $pK_a=5.2$). The latter azole is generated in a yield of 21% alongside adenine as a product of cyanide polymerization in formamide[28] (Fig. 2 and Supplementary Fig. 8). In the reaction with **Ac-L-Ala**, DCI **7** increased the yield of **5-L-a** (6-fold after 12 h, Table 1, Entries 6–8) relative to the reaction in the absence of an azole (Supplementary Figs. 9–13). Remarkably, when a 1:1 mixture of **Ac-L-Ala** and **Ac-D-Ala** was reacted with **1**, A3′P **3** and **7**, acyl adducts **5-L-a** and **5-D-a** were

Table 1 | The yields of 5 from different carboxylic acids and A3′P 3, activated by methyl isonitrile 1 with different concentrations of DCI 7

Entry	Carboxylic acid	DCI 7 (mM)	pH	Time (h)	Yield[a] of 5 (%)	Yield[a] of 4 (%)
1	-	0	5.2	12	n. d.[b]	12
2	Acetic acid	0	5.2	12	n. d.	56
3	Glycine	0	5.2	12	n. d.	11
4	**Ac-L-Ala**	0	5.2	12	**5-L-a**: 5; **5-D-a**: <1	18
5	**Ac-L-Ala**	0	4.0	12	**5-L-a**: 2; **5-D-a**: <1	20
6	**Ac-L-Ala**	5	5.2	12	**5-L-a**: 10; **5-D-a**: <1	20
7	**Ac-L-Ala**	25	5.2	12	**5-L-a**: 15; **5-D-a**: <1	20
8	**Ac-L-Ala**	100	5.2	12	**5-L-a**: 29; **5-D-a**: 2	23
9	**Ac-L-Ala** and **Ac-D-Ala**[c]	100	5.2	12	**5-L-a**: 21; **5-D-a**: 5	25
10	**Ac-D-Ala**	100	5.2	12	**5-L-a**: 4; **5-D-a**: 8	27
11	Acetic acid	100	5.2	12	32	47
12	**Ac-L-Ala**	100	4.0	12	**5-L-a**: 20; **5-D-a**: <1	17
13	Gly-Gly[d]	100	4.0	12	27	55
14	Gly-Gly[d]	100	5.2	12	6	68
15	Gly-Gly[d]	0	4.0	7	1	63
16	Gly-Gly-Gly[d]	100	4.0	12	34	43
17	L-Ala-L-Ala[d]	100	4.0	8	51[e]	37
18	D-Ala-D-Ala[d]	100	4.0	8	12[e]	61
19	Formyl-Gly-Gly[f]	100	4.0	9	31	31
20	Formyl-Gly-Gly[f]	100	5.2	9	28	47

Data reported are representative of distinct samples, $n = 2$ replicates. [a]Conditions used for the formation of **5**: A3′P **3** (10 mM), methyl isonitrile **1** (100 mM), carboxylic acid (100 mM), DCI **7** as indicated, at 23 °C unless specified. Yields were calculated based on ^{1}H NMR spectroscopy. [b]Product not detected. [c]**Ac-L-Ala** (50 mM) and **Ac-D-Ala** (50 mM). [d]Methyl isonitrile (200 mM) [e]Total yield [f]*N*-Formyl-Gly-Gly (200 mM)

observed in a 4:1 ratio after 12 h (Table 1, Entry 9, Supplementary Fig. 14). When **Ac-L-Ala** and **Ac-D-Ala** were separately mixed with **1**, **3** and **7**, both **5-L-a** and **5-D-a** were again produced, indicating racemization during the activation chemistry (Table 1, Entries 6, 9 and 10, Supplementary Figs. 15 and 16). No 2′-acyl adducts could be observed when free amino acids were used as substrates, possibly due to the reduced nucleophilicity of their carboxylates ($pK_{aH} \approx 2.2$). In the reaction of A3′P **3** and acetic acid with **1** and **7**, 2′-acetyl adduct **5b** was observed in 32% yield after 12 h. Similar adducts have been shown to be able to play a key role in abiotic RNA proofreading[22,29] (Table 1, Entry 11, Supplementary Fig. 17). Furthermore, **Ac-L-Ala**, di- and tri-peptides activated by the same chemistry gave good yields of 2′-acyl adducts **5** at pH 4 in the presence of **7** (Table 1, Entries 12–20, Supplementary Figs. 18–25).

Simultaneous activation of adenosine 5′-monophosphate and carboxylates, formation of mixed anhydrides and peptides. The stereochemical results obtained with both enantiomers of Ac-Ala can be explained by the intermediacy of 5(4*H*)-oxazolone **8a** (Fig. 2), which is known to undergo racemization via an aromatic tautomer **8a′** (refs.[30–32]). This was further confirmed by an H/D exchange experiment (Supplementary Fig. 26). Interestingly, Tamura and Schimmel have previously reported a 4:1 L- over D-enantiomer preference in the transfer of (*N*-acyl) aminoacyl groups from the 5′-phosphate of a DNA strand to the 2′,3′-diol of an abutting RNA strand in a templated system, using a mixed anhydride pre-synthesized by conventional procedures[33,34] (Supplementary Fig. 27). The observation of mixed anhydrides of 3′-phosphates **6** suggested that the chemistry might enable the synthesis of mixed anhydrides of 5′-phosphates and this could provide a link between prebiotic activation chemistry and the Tamura and Schimmel chemistry and suggest an underlying reason for the relative stereochemistry of peptides and nucleic acids in biology. For simplicity and to allow analysis using NMR spectroscopy, we employed adenosine 5′-monophosphate (A5′P, **9**, Fig. 1b) as a proxy for an oligonucleotide-5′-phosphate. Thus, we investigated the reaction of carboxylic acids and methyl isonitrile **1** at pH 5.2 with A5′P **9**. In the absence of DCI **7**, a yield of 25% of the 5′-mixed anhydride **10a** was obtained after 24 h when **Ac-L-Ala** was used (Table 2, Entry 1, Supplementary Fig. 28). Unexpectedly, the yield of **10a** decreased as the concentration of DCI **7** was increased, although the reactions were also accelerated, based on the reduced time needed to reach maximum yield (Table 2, Entries 1–5, Supplementary Figs. 29–31, Supplementary Table 1). A pre-synthesized mixed anhydride **10** was incubated with different concentrations of DCI under similar conditions, which clearly showed that DCI catalysed the hydrolysis of mixed anhydride **10** (Supplementary Fig. 32). No stereoselectivity was observed in the reaction of A5′P **9** with **Ac-L-Ala** and **Ac-D-Ala** (Table 2, Entries 1, 5 and 6, Supplementary Figs. 33–35). The reaction of acetic acid with A5′P **9** in the absence of **7** gave only 2% of mixed anhydride **10b** after 24 h (Table 2, Entry 7, Supplementary Fig 36, Supplementary Table 2). When unprotected di- or tri-peptides were added to a solution containing A5′P **9** and methyl isonitrile **1**, very low yields of mixed anhydrides were detected using ^{31}P NMR spectroscopy (less than 3% yield, data not shown). We reasoned that aminolysis of activated peptides, or the intramolecular formation of diketopiperazines in the case of dipeptides[35], reduced the yields of mixed anhydride. In fact, oligopeptide formation was observed using ^{1}H NMR and confirmed by mass spectrometry (up to 12-mers, Supplementary Fig. 37). An analogous reaction of A5′P **9** with *N*-formylglycylglycine, a model N-terminal protected peptide, was performed and a 19% yield of mixed anhydride was observed after 36 h (Table 2, Entry 8, Supplementary Fig. 38), further supporting our hypothesis. To quantify the efficiency of peptide formation[4], **Ac-L-Ala** (100 mM), glycine (50 mM) and methyl isonitrile **1** (200 mM) were mixed together at pH 4 or 5.2 and incubated at 23 °C. The yield of the major peptide product **Ac-Ala-Gly** reached 64% after 24 h at pH 4 based on ^{1}H NMR spectroscopy and a yield of 59% was obtained after 10 days for the same reaction at pH 5.2 (Supplementary Figs. 39–48).

Formation of peptides and peptidyl–RNAs in a unified system. We wondered how the reactions described above might interfere and interact when operated together in a more complex system. Reaction mixtures with A3′P **3** (10 mM), A5′P **9** (10 mM), **Ac-L-Ala** (100 mM), glycine (50 mM), DCI **7** (20 mM) and methyl isonitrile **1** (200 mM) at pH 4 or 5.2 (Fig. 1c, Supplementary Figs. 49 and 50) were monitored using ^{1}H and ^{31}P NMR spectroscopy at 23 °C. At pH 4, the yields of 2′-acyl adduct **5a** and mixed anhydride **10a** were 10 and 8%, respectively, after 12 h and the yield of peptide **Ac-Ala-Gly** was 60% after 24 h, these yields being comparable to those obtained in separate reactions (Supplementary Figs. 51–56). However, notably, two new species were identified in yields of 4 and 2% after 16 h using NMR spectroscopy. These were shown to be the 2′-Ac-Ala-Gly adduct of A3′P **5c** and the **Ac-Ala-Gly** mixed anhydride of A5′P **10c**, respectively (Supplementary Figs. 57–60). At pH 5.2, the yields of **5a** and **10a** were 19 and 7%, respectively, after 12 h and the yield

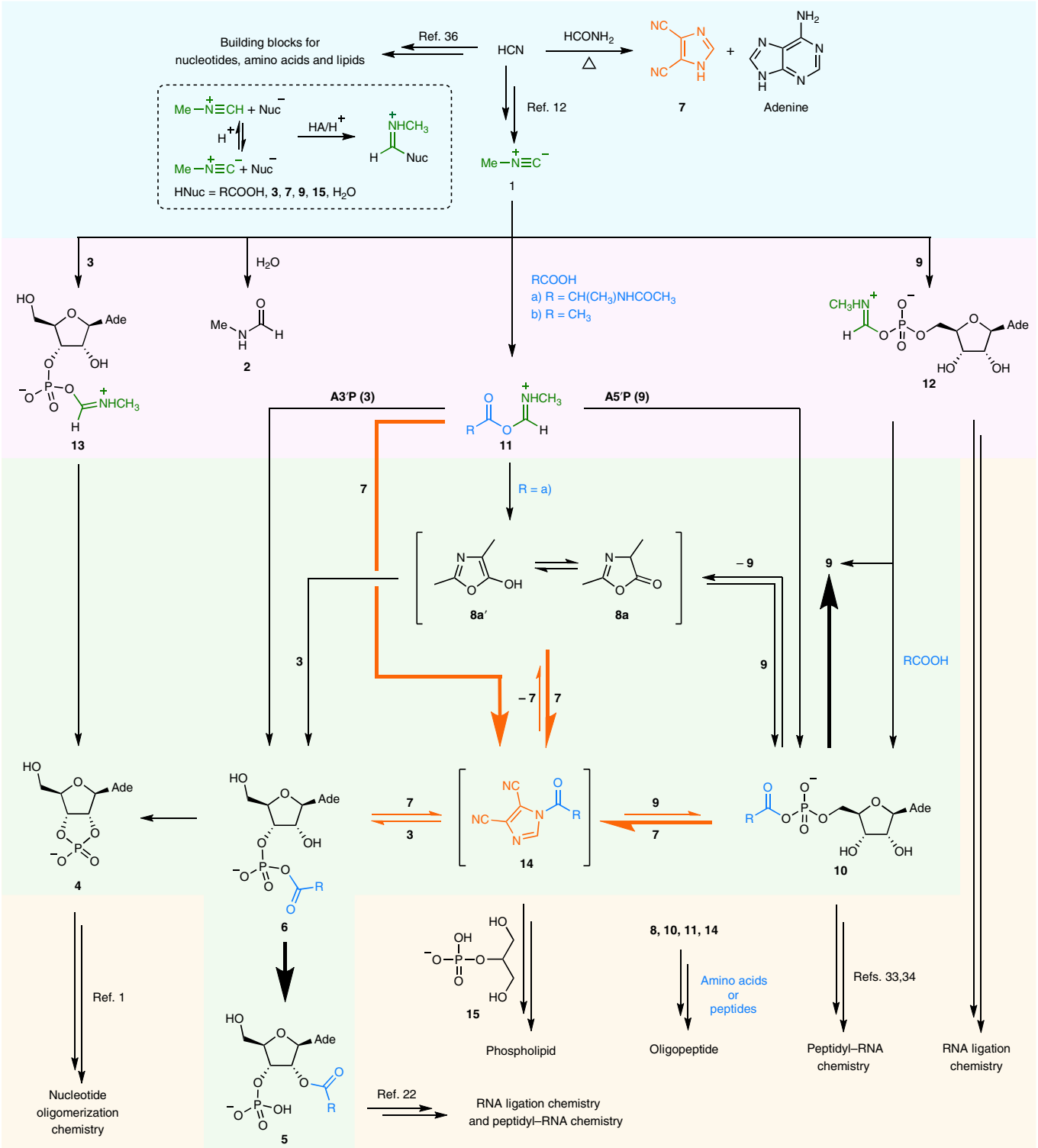

Fig. 2 | Proposed reaction scheme of the activation chemistry. From top to bottom, four distinct chemical stages are highlighted in different colours. The first stage (blue background) encompasses the prebiotic synthesis of building blocks, activating reagents, catalysts and other species. The second stage (pink background) is activating-reagent-dependent and the activating reagent is covalently attached to carboxylates or phosphates in the form of intermediates **11**, **12** and **13**. A dashed box includes two potential alternative pathways by which the activation chemistry mediated by methyl isonitrile **1** may be initiated. The third stage (green background) is activating-reagent-independent and is characterized by common intermediates such as **6**, **8**, **10** and **14**. The formation of these intermediates could in principle be driven by alternative activating reagents. The fourth stage (orange background) involves the condensation of various species to form oligopeptides, peptidyl–RNAs, RNAs and primordial phospholipids. Pathways involving catalysis by DCI **7** are highlighted in orange and bold arrows indicate steps preferentially taken. Carboxyl moieties including aliphatic carboxylic acids, *N*-acyl amino acids and peptides are highlighted in blue. All the activated carboxylates and phosphates are prone to hydrolysis but these reactions are omitted for the sake of clarity.

Table 2 | The yields of 10 from different carboxylic acids and A5′P 9, activated by methyl isonitrile 1 with different concentrations of DCI 7

Entry	Carboxylic acid	DCI 7 (mM)	pH	Time (h)	Yield[a] of 10 (%)
1	**Ac-L-Ala**	0	5.2	24	**10-L-a**: 18; **10-D-a**: 7
2	**Ac-L-Ala**	5	5.2	24	**10-L-a**: 14; **10-D-a**: 4
3	**Ac-L-Ala**	25	5.2	12	**10-L-a**: 9; **10-D-a**: 2
4	**Ac-L-Ala**	100	5.2	6	**10-L-a**: 5; **10-D-a**: 1
5	**Ac-D-Ala**	0	5.2	24	**10-L-a**: 7; **10-D-a**: 17
6	**Ac-L-Ala** and **Ac-D-Ala**[b]	0	5.2	24	**10-L-a**: 13; **10-D-a**: 15
7	Acetic acid	0	5.2	24	2
8	Formyl-Gly-Gly	0	5.2	36	19

Data reported are representative of distinct samples, $n=2$ replicates. [a]Conditions used for the formation of **10**: A5′P **9** (10 mM), methyl isonitrile **1** (100 mM), carboxylic acid (200 mM), at 23 °C unless specified. Yields were calculated based on ^{31}P NMR spectroscopy. [b]**Ac-L-Ala** (100 mM) and **Ac-D-Ala** (100 mM).

of **Ac-Ala-Gly** was 50% after 10 days (Supplementary Figs. 61 and 62). These results suggest that common activation chemistry could not only have brought about the activation and joining of carboxylate and phosphate building blocks in solution, but it could also have driven the system towards higher-order complexity.

Network of activation chemistry pathways driven by methyl isonitrile. Based on the results above, a reaction scheme for the activation chemistry was proposed (Fig. 2). In the first step, imidoyl carboxylates **11** and imidoyl phosphates **12** and **13** are formed following specific/general acid catalysed attack of carboxylates or phosphates on the isonitrile (Dashed box in Fig. 2). In later steps, mixed anhydrides **6** and **10** are generated via the putative oxazolone **8a**/**8a′** and *N*-acyl dicyanoimidazole intermediate **14**, followed by thermodynamically and/or kinetically favoured pathways to **4**, **5** and peptides, respectively. Additional experiments were run to support our reaction scheme and the formation of the key intermediate **14** connecting **6**, **8** and **10**, by DCI catalysis (Fig. 2). The reaction of A3′P **3** (10 mM) with premade 2-methyl-4(4-methoxybenzyl)-5(4*H*)-oxazolone[31] (50 mM) gave **5** and **6** more rapidly in the presence of DCI **7** (Pathways **8**→**6**→**5** and **8**→**14**→**6**→**5** in Fig. 2, Supplementary Fig. 63) than in its absence. Imidazoles are known to catalyse the hydrolysis of mixed anhydrides via nucleophilic catalysis[23] and this is consistent with our observation that DCI **7** catalyses both the formation and hydrolysis of **10a**, presumably via equilibrium with *N*-acyl dicyanoimidazolide **14** (Pathways **10**→**14** and **14**→**10** in Fig. 2). Interestingly, the role of DCI **7** in enhancing the formation of 2′-acyl adducts **5** can also result from the ability of DCI, as a good nucleophile, to take any activated carboxylate in the system to form *N*-acyl dicyanoimidazole **14**. Compound **14** could also be in equilibrium with 3′-mixed anhydride **6**, but the favourable intramolecular acyl transfer reaction of **6** drives the equilibrium towards the 2′-acyl adduct **5**.

To support the proposed DCI-catalysed equilibrium between **10** and **6**, proceeding via **14** (Pathway **10**→**14**→**6**→**5** as shown in Fig. 2), the reaction of premade mixed anhydride **10e** (≈20 mM), A3′P **3** (20 mM) and DCI **7** (0 mM, 5 mM or 20 mM) in MES buffer (250 mM, pH 5.2) was monitored using ^{31}P NMR spectroscopy (Supplementary Fig. 64). When 5 or 20 mM of DCI was used the yields of A>P **4** after 12 h were 4 and 6%, respectively, while the yields of ester **5i** were 5 and 10%, respectively. In the absence of DCI, only **4** was observed in 2% yield after 12 h (Supplementary Fig. 64). These results show that DCI catalyses the transfer of acyl groups from A5′P **9** to A3′P **3** in a concentration-dependent manner. This is best explained by nucleophilic catalysis proceeding via the acyl-dicyanoimidazolide intermediate **14** (refs. [23,24]). Reaction of A5′P **9** (10 mM), Ac-Gly (50 mM), DCI **7** (5 mM) and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC, 50 mM) in MES buffer (250 mM, pH 6) to give mixed anhydride **10e** was monitored using ^{31}P NMR spectroscopy. This showed that both the formation (Pathways **8**→**10** and **8**→**14**→**10** in Fig. 2, Supplementary Fig. 65) and the hydrolysis (Pathway **10**→**14** in Fig. 2) of mixed anhydride **10e** are catalysed by DCI **7**. These results suggest that **6**, **8**, **10** and **14** are common intermediates/products and that the pathways leading to these species do not require a specific activating reagent.

Formation of primordial phospholipids. With this mechanism in mind, we reasoned that this chemistry could potentially be used to generate phospholipids. Thus, glycerol-2-phosphate[36] **15** was mixed with methyl isonitrile **1** and DCI **7** in the presence of carboxylic acids of different chain lengths. Mono-acyl glycerol-2-phosphates and mono-acyl glycerol-2,3-cyclic phosphates were detected (19 and 4% yield, respectively, for decanoic acid) using ^{31}P NMR spectroscopy (Supplementary Fig. 66). Importantly, we and others have previously shown that the phospholipids derived from octanoic acid and decanoic acid are capable of self-assembling into vesicles[37,38].

Template-directed ligation of oligonucleotides. Finally, we wondered if this activation chemistry could drive template-directed RNA ligation in solution. Oligonucleotides collectively constituting a nicked duplex were mixed with methyl isonitrile **1**, DCI **7**, divalent metal ions (Mn^{2+}, Zn^{2+}, Mg^{2+} or Fe^{2+}) and *N*-methylimidazole at pH 6 (Supplementary Table 3) and the mixture was incubated at 23 °C. The best yield of the expected 15 nt ligation product was obtained with Mn^{2+} (45% after 5 days and 72% after 15 days). The yield of the ligation dropped to less than 1 and 10% after 5 days when the same reaction was performed at pH 4 and pH 5 respectively. A 16 nt side product was proven to be the 5′,5′-pyrophosphate of the 5′-phosphorylated ligator oligonucleotide (Supplementary Fig. 67).

Discussion

We have previously shown that RNA, peptide and lipid building blocks can be synthesized by a common cyanosulfidic chemistry from hydrogen cyanide[36,39]. Herein, we describe how a combination of methyl isonitrile **1** and DCI **7**, which are both derived from cyanide, allow the activation of multiple prebiotic building blocks in solution (Fig. 2). Recently, diamidophosphate has been reported to phosphorylate and enable the condensation of amino acids and nucleotides, respectively[37]. However, low water activity and paste-like conditions are required for the formation of the products and mixed species such as peptidyl–RNA adducts are not observed. In our system, simultaneous activation of peptides, RNA and amino acids gives higher-order peptides and peptidyl–RNA in aqueous solution. Invoking pH cycling within the system[40], unified prebiotic activation chemistry could have potentially supported the synthesis of longer peptides, peptidyl–RNA and RNA in aqueous solution at room temperature. Primitive peptidyl–RNAs might have possessed catalytic properties analogous to modern ribonucleoproteins[41]. Therefore, they represent species worthy of consideration in either a co-evolving peptide–RNA world or a more nuanced variant of the RNA world scenario. In addition, the observed preferential transfer of L- over D-*N*-acyl amino acids and peptides suggests that activation chemistry could have potentially played a role in the stereoselective RNA-templated synthesis of peptides on early Earth. If the absolute stereochemistry of nucleotides could have been fixed[42], peptide

synthesis based on RNA could have operated with some degree of stereocontrol. Furthermore, the activation of glycerol-2-phosphate in the presence of fatty acids gives primordial phospholipids, which could enable the transition from fatty acid to phospholipid vesicles[43]. Taken together these results suggest that unified activation chemistry could have potentially supported the synthesis of the functional polymers and membrane-forming components required for life, while simultaneously underpinning their intimate association in a compartmentalized RNA–peptide system[44].

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41557-020-00564-3.

Received: 3 December 2019; Accepted: 21 September 2020;
Published online: 22 October 2020

References

1. Verlander, M. S., Lohrmann, R. & Orgel, L. E. Catalysts for the self-polymerization of adenosine cyclic 2′,3′-phosphate. *J. Mol. Evol.* **2**, 303–316 (1973).
2. Lambert, J.-F. Adsorption and polymerization of amino acids on mineral surfaces: A review. *Orig. Life Evol. Biosph.* **38**, 211–242 (2008).
3. Kanavarioti, A., Monnard, P.-A. & Deamer, D. W. Eutectic phase in ice facilitate nonenzymatic nucleic acid synthesis. *Astrobiology* **1**, 271–281 (2001).
4. Canavelli, P., Islam, S. & Powner, M. Peptide ligation by chemoselective aminonitrile coupling in water. *Nature* **571**, 546–549 (2019).
5. Lohrmann, R. & Orgel, L. E. Prebiotic synthesis: phosphorylation in aqueous solution. *Science* **161**, 64–66 (1968).
6. Ibanez, J. D., Kimball, A. P. & Oró, J. Possible prebiotic condensation of mononucleotides by cyanamide. *Science* **173**, 444–446 (1971).
7. Leman, L., Orgel, L. E. & Ghadiri, M. R. Carbonyl sulfide-mediated prebiotic formation of peptides. *Science* **306**, 283–286 (2004).
8. Leman, L., Orgel, L. E. & Ghadiri, M. R. Amino acid dependent formation of phosphate anhydrides in water mediated by carbonyl sulfide. *J. Am. Chem. Soc.* **128**, 20–21 (2006).
9. Tsanakopoulou, M. & Sutherland, J. D. Cyanamide as a prebiotic phosphate activating agent – catalysis by simple 2-oxoacid salts. *Chem. Commun.* **53**, 11893–11896 (2017).
10. Liu, Z. et al. Tuning the reactivity of nitriles using Cu(II) catalysis – potentially prebiotic activation of nucleotides. *Chem. Sci.* **9**, 7053–7057 (2018).
11. Orgel, L. E. Evolution of the genetic apparatus. *J. Mol. Biol.* **38**, 381–393 (1968).
12. Remijan, A. J., Hollis, J. M., Lovas, F. J., Plusquellic, D. F. & Jewell, P. R. Interstellar isomers: the importance of bonding energy differences. *Astrophys. J.* **632**, 333–339 (2005).
13. Xiang, Y.-B., Drenkard, S., Baumann, K., Hickey, D. & Eschenmoser, A. Chemie von α-aminonitrilen 12. Mitteilung. Sondierungen über thermische umwandlungen von α-aminonitrilen. *Helv. Chim. Acta* **77**, 2209–2250 (1994).
14. Xu, J. et al. Photochemical reductive homologation of hydrogen cyanide using sulfite and ferrocyanide. *Chem. Commun.* **54**, 5566–5569 (2018).
15. Mariani, A., Russell, D. A., Javelle, T. & Sutherland, J. D. A light-releasable potentially prebiotic nucleotide activating agent. *J. Am. Chem. Soc.* **140**, 8657–8661 (2018).
16. Mullen, L. B. & Sutherland, J. D. Simultaneous nucleotide activation and synthesis of amino acid amides by a potentially prebiotic multi-component reaction. *Angew. Chem. Int. Ed.* **46**, 8063–8066 (2007).
17. Pirrung, M. C. & Sarma, K. D. Multicomponent reactions are accelerated in water. *J. Am. Chem. Soc.* **126**, 444–445 (2004).
18. Paprocki, D., Koszelewski, D., Walde, P. & Ostaszewski, R. Efficient Passerini reactions in an aqueous vesicle system. *RSC Adv.* **5**, 102828–102835 (2015).
19. Sung, K. & Chen, C.-C. Kinetics and mechanism of acid-catalysed hydrolysis of cyclohexyl isonitrile and pK_a determination of *N*-cyclohexylnitrilium ion. *Tetrahedron Lett.* **42**, 4845–4848 (2001).
20. Lim, Y.-Y. & Stein, A. R. Acid-catalysed solvolysis of isonitriles. I. *Can. J. Chem.* **49**, 2455–2459 (1971).
21. Biron, J., Parkes, A., Pascal, R. & Sutherland, J. D. Expeditious, potentially primordial, aminoacylation of nucleotides. *Angew. Chem. Int. Ed.* **117**, 6889–6892 (2005).
22. Bowler, F. R. et al. Prebiotically plausible oligoribonucleotide ligation facilitated by chemoselective acetylation. *Nat. Chem.* **5**, 383–389 (2013).
23. Jencks, W. P. & Carriuolo, J. Imidazole catalysis. II. Acyl transfer and the reactions of acetyl imidazole with water and oxygen anions. *J. Biol. Chem.* **234**, 1272–1279 (1959).
24. Lacey, J. C. Jr. & White, W. E. Jr. Aminoacyl transfer: chemical conversion of an aminoacyl adenylate to an imidazolide. *Biochem. Biophys. Res. Commun.* **47**, 565–573 (1972).
25. Ferris, J. P. & Kuder, J. E. Chemical evolution. III. The photochemical conversion of enaminonitriles to imidazoles. *J. Am. Chem. Soc.* **92**, 2527–2533 (1970).
26. Oró, J. & Kimball, A. P. Synthesis of purines under possible primitive earth conditions. I. Adenine from hydrogen cyanide. *Arch. Biochem. Biophys.* **94**, 217–227 (1961).
27. Fahrenbach, A. et al. Common and potentially prebiotic origin for precursors of nucleotide synthesis and activation. *J. Am. Chem. Soc.* **139**, 8780–8783 (2017).
28. Hudson, J. S. et al. A unified mechanism for abiotic adenine and purine synthesis in formamide. *Angew. Chem. Int. Ed.* **51**, 5134–5137 (2012).
29. Mariani, A. & Sutherland, J. D. Non-enzymatic RNA backbone proofreading through energy-dissipative recycling. *Angew. Chem. Int. Ed.* **56**, 6563–6566 (2017).
30. Danger, G. et al. 5(4*H*)-Oxazolones as intermediates in the carbodiimide- and cyanamide-promoted peptide activations in aqueous solution. *Angew. Chem. Int. Ed.* **52**, 611–614 (2013).
31. Liu, Z., Beaufils, D., Rossi, J. & Pascal, R. Evolutionary importance of the intramolecular pathways of hydrolysis of phosphate ester mixed anhydrides with amino acids and peptides. *Sci. Rep.* **4**, 7440 (2014).
32. Liu, Z., Rigger, L., Rossi, J., Sutherland, J. D. & Pascal, R. Mixed anhydride intermediates in the reaction of 5(4*H*)-oxazolones with phosphate esters and nucleotides. *Chem. Eur. J.* **22**, 14940–14949 (2016).
33. Tamura, K. & Schimmel, P. R. Chiral-selective aminoacylation of an RNA minihelix. *Science* **305**, 1253 (2004).
34. Tamura, K. & Schimmel, P. R. Chiral-selective aminoacylation of an RNA minihelix: mechanistic features and chiral suppression. *Proc. Natl Acad. Sci. USA* **103**, 13750–13752 (2006).
35. Beaufils, D., Jepaul, S., Liu, Z., Boiteau, L. & Pascal, R. The activation of free dipeptides promoted by strong activating agents in water does not yield diketopiperazines. *Orig. Life Evol. Biosph.* **46**, 19–30 (2016).
36. Patel, B. H., Percivalle, C., Ritson, D. J., Duffy, C. D. & Sutherland, J. D. Common origins of RNA, protein and lipid precursors in a cyanosulfidic protometabolism. *Nat. Chem.* **7**, 301–307 (2015).
37. 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).
38. Bonfio, C. et al. Length-selective synthesis of diacylglycerol-phosphates through energy-dissipative cycling. *J. Am. Chem. Soc.* **141**, 3934–3939 (2019).
39. Sutherland, J. D. The origin of life—out of the blue. *Angew. Chem. Int. Ed.* **55**, 104–121 (2016).
40. Lorenz, M. R. et al. Proton gradients and pH oscillations emerge from heat flow at the microscale. *Nat. Commun.* **8**, 1897 (2017).
41. Cech, T. R. Evolution of biological catalysis: ribozyme to RNP enzyme. *Cold Spring Harbor Symp. Quant. Biol.* **74**, 11–16 (2009).
42. 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).
43. Blain, J. C. & Szostak, J. W. Progress towards synthetic cells. *Annu. Rev. Biochem.* **83**, 615–640 (2014).
44. Sutherland, J. D. Opinion: Studies on the origin of life – the end of the beginning. *Nat. Rev. Chem.* **1**, 0012 (2017).

Methods

General methods. Reagents and solvents were obtained from Acros Organics, Alfa Aesar, Santa Cruz Biotechnology, Sigma-Aldrich, SYNTHON Chemicals GmbH & Co., KG and VWR International and were used without further purification unless otherwise stated. The 7 nt RNA oligonucleotide (7 nt primer, 5′-GAGAACC-3′), 8 nt RNA oligonucleotide (8 nt ligator, 5′-/phos/UACUGGCA/3Cy3Sp/-3′) and 13 nt RNA oligonucleotide (13 nt template, 5′-CCAGUAGGUUCUC-3′) were purchased in HPLC-purified Na^+ form from Integrated DNA Technologies. The 8 nt RNA oligonucleotide (5′-/phos/UACUGGCA-3′) was synthesized using an ÄKTA oligopilot plus 10 instrument (GE Healthcare). All photochemical reactions were carried out in Norell Suprasil quartz NMR tubes purchased from Sigma-Aldrich using Hg lamps with principal emission at 360 nm in a Rayonet photochemical chamber reactor RPR-200, acquired from The Southern New England Ultraviolet Company. A Mettler Toledo SevenEasy pH Meter S20 combined with a Thermo Fisher Scientific Orion 8103BN Ross semi-micro pH electrode was used to measure and adjust the pH to the desired value. NMR spectra (^{1}H, ^{31}P and ^{13}C) were acquired using a Bruker Ultrashield 400 Plus instrument or a Bruker Ascend 400 instrument operating at 400.13, 161.97 and 100.62 MHz, respectively. Samples consisting of H_2O/D_2O mixtures were analysed using HOD suppression to collect ^{1}H NMR spectroscopy data. The notations s, d, p and m represent the multiplicities singlet, doublet, quintet and multiple, respectively. Chemical shifts (δ) are shown in ppm. Mass spectra were acquired on an Agilent 1200 LC-MS system equipped with an electrospray ionization (ESI) source and a 6130 quadrupole spectrometer (LC solvents: A, 0.2% formic acid in H_2O and B, 0.2% formic acid in acetonitrile). Gel electrophoresis experiments using 20% polyacrylamide and 8 M urea gels (0.75 mm thick, 20 cm long) were typically run at 15 W in Tris/Borate/EDTA buffer. Fluorescence imaging was performed using an Amersham Typhoon imager (GE Healthcare) and quantified using Image Quant TL software (version 7.0). Oligonucleotide concentrations were determined by ultraviolet absorbance at 260 nm using a NanoDrop ND-1000 spectrophotometer.

Prebiotic synthesis of DCI 7. KCN (0.46 mmol, 30 mg, ^{13}C-labelled or natural abundance) was mixed with ammonium formate (0.06 mmol, 4 mg) in formamide (1 ml) and the mixture was heated at 165 °C for 2 h (ref. [28]). The solvent was evaporated under reduced pressure and the residue was extracted with hot water following evaporation. Then, the product mixture was analysed using ^{1}H and ^{13}C NMR spectroscopy (10% D_2O in H_2O). The formation of DCI **7** was confirmed by spiking the crude mixture with authentic material (purchased from Sigma-Aldrich) and recording the ^{1}H NMR spectrum and also by comparing the ^{13}C NMR spectrum of the crude mixture with that of the authentic material.

Chemical synthesis of *N*-formyl-glycylglycine (formyl-Gly-Gly)[45]. Glycyl-glycine (15.1 mmol, 2 g) was stirred in formic acid (35 ml) and heated at 55 °C while adding acetic anhydride (13.5 ml) dropwise. The solution was stirred at room temperature for 1 h then ice–water (12 ml) was added and the solution was concentrated under reduced pressure. The residual solution was left in a fridge at 4 °C overnight whereupon crystals formed at the bottom of the flask. The crystals were collected by filtration, washed with diethyl ether and dried overnight using a desiccator to give a: white solid; yield 810 mg (31%); ^{1}H NMR (400 Hz, D_2O): δ = 8.2 (s, 1H), 4.1 (s, 2H), 4.0 (s, 2H) ppm; ^{13}C NMR (101 Hz, D_2O): δ = 173.3, 171.6, 164.9, 41.1, 41.0 ppm.

Chemical synthesis of *N*-acetyl-alanylglycine (Ac-Ala-Gly)[46]. Alanyl-glycine (1 mmol, 146 mg) was dissolved in a minimum amount of saturated $NaHCO_3$ solution. Then, 2 equiv. of acetic anhydride were added. After 30 min at room temperature, Na^+ was removed using an H^+-formed Dowex-50 column. The solution was lyophilized to give a: white powder; yield 170 mg (83%); ^{1}H NMR (400 Hz, D_2O): δ = 8.22 (s, 1H), 8.04 (s, 1H), 4.27 (p, J = 7.1 Hz, 1H), 3.84–3.67 (m, 2H), 1.97 (s, 3H), 1.32 (d, J = 7.2 Hz, 3H) ppm; ^{13}C NMR (101 Hz, D_2O): δ = 175.59, 175.09, 174.17, 49.74, 42.84, 21.83, 16.80 ppm.

Standard procedure for carboxylate and phosphate activation in reactions containing A3′P 3. An aqueous solution (0.5 ml, H_2O/D_2O, 9:1) of A3′P **3** (10 mM), carboxylic acid (100 mM, if required) and heterocyclic compound (100 mM, if required) was adjusted to the corresponding pH value with HCl (1 M) and NaOH (1 M) solutions. Methyl isonitrile **1** (3 µl, final concentration to 100 mM) was then added and the reaction was analysed using ^{1}H and ^{31}P NMR spectroscopy at 23 °C. The pH had changed only slightly (<0.2) by the time the isonitrile was totally consumed.

Standard procedure for carboxylate and phosphate activation in reactions containing A5′P 9. An aqueous solution (0.5 ml, H_2O/D_2O, 9:1) of A5′P **9** (10 mM), carboxylic acids (200 mM) and heterocyclic compounds (100 mM, if required) was adjusted to the desired pH value with HCl (1 M) and NaOH (1 M) solutions. Methyl isonitrile **1** (3 µl, final concentration to 100 mM) was then added and the reaction was analysed using ^{1}H and ^{31}P NMR spectroscopy at 23 °C. The pH had changed only slightly (<0.2) by the time the isonitrile was totally consumed.

Standard procedure for carboxylate and phosphate activation in reactions containing A3′P 3 and A5′P 9. An aqueous solution (0.5 ml, H_2O/D_2O, 9:1) of A3′P **3** (10 mM), A5′P **9** (10 mM), **Ac-L-Ala** (100 mM), glycine (50 mM) and DCI **7** (20 mM) was adjusted to pH 4 or 5.2 with HCl (1 M) and NaOH (1 M) solutions. Methyl isonitrile **1** (6 µl, final concentration to 200 mM) was then added and the reaction was analysed using ^{1}H and ^{31}P NMR spectroscopy at 23 °C. The pH had changed only slightly (<0.2) by the time the isonitrile was totally consumed.

Preparation of adenosine-2′-*N*-acetyl-L (or D)-alanine-3′-monophosphate (L-standard or D-standard). 1,1′-Carbonyldiimidazole (8.9 mg, 55 µmol) was added to a suspension of **Ac-L-Ala** or **Ac-D-Ala** (6.5 mg, 50 µmol) in acetonitrile (50 µl) and the mixture was vortexed for 5 min at room temperature. The resulting solution, containing the acyl imidazolide derivative of *N*-acetyl-L (or D)-alanine, was then added to a solution of A3′P **3** (9.8 mg, 25 µmol) in H_2O/D_2O (9:1, 450 µl). The mixture was quickly analysed using ^{1}H and ^{31}P NMR spectroscopy to confirm that the desired 2′-acylated product had formed. The mixture was then used immediately to spike other experiments.

Chemical synthesis of 10. EDC (76.4 mg, 400 µmol) was added to 2 ml of aqueous solution containing A5′P **9** (8 mg, 200 µmol), Ac-Gly (47 mg, 400 µmol) and DCI **7** (2.4 mg, 20 µmol) at pH 6.0 and at room temperature. The pH of the mixture was kept around 6 by adding 5 M HCl solution. After 2 h, the reaction was titrated to pH 4 and quenched by adding 40 ml of cold $NaClO_4$ solution (50 mM $NaClO_4$ in acetone). The precipitate was collected by centrifugation, washed with cold acetone and dried under a stream of nitrogen. The yield of the product **10e** (40 to 70%) was checked using ^{31}P NMR spectroscopy at pH 5 in H_2O/D_2O.

DCI catalysed hydrolysis of 10e. Synthesized **10e** was dissolved in water (1 ml) at pH 5. Then, 150 µl of this solution, 200 µl of MES buffer (500 mM, pH 5.2), 50 µl D_2O and a corresponding volume of DCI **7** solution (500 mM, pH 5.2) were mixed together to give final concentrations of DCI of 0 mM, 10 mM and 50 mM. Water was added to a final volume of 500 µl if needed. The pH was checked to be 5.2, otherwise it was adjusted with HCl (1 M) and NaOH (1 M) solutions. The reactions were then monitored using ^{31}P NMR spectroscopy at 23 °C.

Formation of oligopeptides from a dipeptide under activation chemistry. An aqueous solution (0.5 ml, H_2O/D_2O, 9:1) of A5′P **9** (10 mM), L-Ala-L-Ala (100 mM) and DCI **7** (100 mM) was adjusted to pH 5.2 with HCl (1 M) and NaOH (1 M) solutions. Methyl isonitrile **1** (3 µl, final concentration to 100 mM) was then added and the mixture was incubated at room temperature (white precipitate was formed during the incubation). After 3 days, the precipitate was removed by centrifugation. The supernatant was diluted 100-fold with H_2O and analysed using mass spectrometry (ESI, positive ion mode).

Formation of dipeptide Ac-Ala-Gly in reactions of Ac-Ala and glycine under activation chemistry. An aqueous solution (0.5 ml, H_2O/D_2O, 9:1) of A5′P **9** (0 or 10 mM), **Ac-L-Ala** (100 mM), glycine (50 mM) and DCI **7** (0, 20 or 50 mM) was adjusted to pH 5.2 or 4 with HCl (1 M) and NaOH (1 M) solutions. Methyl isonitrile **1** (6 µl, final concentration to 200 mM) was then added and the mixture was monitored using NMR spectroscopy at 23 °C.

General procedure for desalting of RNA by ethanol precipitation. Oligonucleotides were desalted by addition of 2 M imidazole nitrate solution (pH 6.2, 1/10 the volume of the aliquot taken from the reaction), followed by a 3 M sodium acetate solution (pH 5.2, 1/10 the volume of the aliquot taken from the reaction) and absolute ethanol (to a final concentration of 75% (v/v)). The resulting mixture was kept at −20 °C for 3 h and then centrifuged for 30 min at 16,000*g*. The supernatant was removed and the pellets were washed with 75% (v/v) aqueous ethanol before additional centrifugation (10 min at 16,000*g*). The resulting pellets were air dried before being re-dissolved in water.

General procedure for oligonucleotide ligation reactions. To an aqueous solution of 7 nt primer (20 µM, 0.5 µl of 400 µM stock solution), 8 nt ligator RNA (10 µM, 0.5 µl of 200 µM stock solution), 13 nt template RNA (10 µM, 0.5 µl of 200 µM stock solution), DCI **7** (0 or 100 mM, 0 or 4 µl of 250 mM stock solution at pH 6), 1-methylimidazole (0 or 100 mM, 0 or 1 µl of 1 M stock solution at pH 6), divalent metal ion (Mn^{2+}, Mg^{2+}, Fe^{2+} or Zn^{2+}, 0 or 10 mM, 0 or 1 µl of 100 mM stock solution at pH 6) was added nuclease-free water to 8 µl, then methyl isonitrile **1** (400 mM, 2 µl of a 2 M aqueous stock solution) was added and the reaction was kept at room temperature. Aliquots of 3.0 µl were taken at the indicated time points and desalted by ethanol precipitation. For each aliquot, the resulting pellet was re-dissolved in 3.0 µl of nuclease-free water and 1.0 µl of the resulting solution was mixed with 4.0 µl of loading dye (90% (v/v) formamide, 5% (v/v) glycerol, 24 mM EDTA, Orange G). The resulting mixture was analysed using gel electrophoresis.

Pyrophosphate formation under activation chemistry. An aqueous solution (0.5 ml, H_2O/D_2O, 9:1) of pUACUGGCA (1 mM), DCI **7** (100 mM), $MgCl_2$ (10 mM) and *N*-methylimidazole (100 mM) was adjusted to pH 6 with HCl (1 M) and NaOH (1 M) solutions. Methyl isonitrile **1** (12 µl, final concentration to 400 mM) was then added and the mixture was monitored using NMR spectroscopy at 23 °C.

Standard procedure for reactions of glycerol-2-phosphate 15 or A3′P 3 with different aliphatic acids activated by methyl isonitrile 1. A solution (0.5 ml, formamide/D_2O, 9:1) of glycerol-2-phosphate **15** (10 mM) or A3′P **3** (10 mM), aliphatic acids (100 mM of decanoic acid, octanoic acid, butyric acid or acetic acid) and DCI **7** (100 mM) was adjusted to pH 5.2. Methyl isonitrile **1** (3 µl, final concentration to 100 mM) was then added and the reactions were analysed periodically using ^{1}H and ^{31}P NMR spectroscopy at 23 °C.

Data availability

All data generated or analysed during this study are included in the manuscript and the Supplementary Information.

References

45. Sheehan, J. C. & Yang, D.-D. H. The use of *N*-formyl amino acids in peptide synthesis. *J. Am. Chem. Soc.* **80**, 1154–1158 (1958).
46. Kim, K.-H., Martin, Y., Otis, E. & Mao, J. Inhibition of iodine-125 labeled ristocetin binding to *Micrococcus luteus* cells by the peptides related to bacterial cell wall mucopeptide precursors: quantitative structure–activity relationships. *J. Med. Chem.* **32**, 84–93 (1989).

Acknowledgements

This research was supported by the Medical Research Council (no. MC_UP_A024_1009 to J.D.S.) and the Simons Foundation (no. 290362 to J.D.S.). We thank all J.D.S. group members for fruitful discussions. We thank R. Pascal for helpful suggestions.

Author contributions

Z.L., L.-F.W., J.X., C.B. and D.A.R. carried out the experiments under the supervision of J.D.S. All authors wrote the manuscript. All 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/s41557-020-00564-3.

Correspondence and requests for materials should be addressed to J.D.S.

Reprints and permissions information is available at www.nature.com/reprints.