Thioester-mediated chemoselective aminoacylation of RNA in water.

Jyoti Singh,[1] Benjamin Thoma,[1] Daniel Whitaker,[1] Yuan Yao[1] and Matthew W. Powner[1]

1. Department of Chemistry, University College London, 20 Gordon Street, London, WC1H 0AJ, UK

Abstract: Selective aminoacylation of RNA in water is a key challenge that must be resolved to elucidate the origin of peptide biosynthesis. To date there are no chemical methods to effectively and selectively aminoacylate RNA-2',3'-diols, with the breadth of proteinogenic amino acids, in water. Here we demonstrate that prebiotic nitriles, aminoacyl-adenylates and *N*-carboxyanhydrides all react with thiols (including enzyme-cofactors) to selectively yield aminoacyl-thiols, which in turn aminoacylate RNA in high yield and selectivity for 2',3'-diols. Broad sidechain scope for both thioester synthesis and aminoacylation are demonstrated, including Ala, Asp, Arg, Glu, Gln, Gly, His, Leu, Lys, Phe, Pro, Ser and Val. Asp reacts with high (kinetic) α-aminoacylation selectivity, even upon dual activation of α- and β-carboxylates, and Glu γ-aminoacylation is blocked by pyroglutamate formation. Arg-aminoacylation is catalysed and enhanced by unprecedented sidechain nucleophilic catalysis. Thioester formation suppresses amine coupling and native-duplex is shown to direct chemoselective 2',3'-aminoacylation of RNA. Our results suggest an important role for thiol cofactors in RNA-aminoacylation prior to the evolution of proteinaceous synthetase enzymes.

Introduction

At life's functional core there is a complex and inseparable interplay between nucleic acids and proteins, but the origin of this relationship remains a mystery. Whilst nucleic acids store, replicate and transmit sequence information through their inherent structural capacity for molecular (self)recognition,[1,2,3,4] proteins are the molecular and structural workhorses of life. Unlike nucleic acids, peptides do not innately replicate in a sequence specific-manner.[5,6] Therefore, life must control and transmit peptide sequences, which are essential to its survival, via nucleic acid-encoding, but this flow of information is unidirectional– sequence information does not flow back from peptides into nucleic acids. Thus, the translation and expression of nucleic acid information as peptides is fundamental to all life on Earth.[7] Understanding how peptide biosynthesis could have first emerged is a major gap in our understanding of life, but this is a formidable challenge due to the immense complexity and antiquity of peptide biosynthesis.[8] Peptides are built by life through two main routes – ribosomal and non-ribosomal synthesis (Fig. 1a, 1b). Although the macromolecules, encoding mechanisms, and amino acid activation strategies (i.e., ribonucleosidyl esters or pantetheinyl thioesters, respectively) which orchestrate these two processes are very different, there are striking chemical similarities. Both proceed in the *N*-to-*C* terminal direction by insertion of a nucleophilic 'amino acid' monomer into the electrophilically-activated growing peptide chain. To control this chemistry both the nucleophile and the electrophile are covalently linked to macromolecules, and the monomer (as well as the polymer) is electrophilically activated, so that monomer insertion yields an activated peptide and resets the polymer growth cycle. The (thio)ester covalent linkages are critical to align and arrange the activated amino acids for sequence specific peptide synthesis.(Thio)esters are energetically uphill with respect to amides but must ideally not spontaneously form random peptides in the absence of catalysis and macromolecular alignment for coded peptide synthesis. The covalent nature of this linkage is important, it must enable catalysed peptide synthesis to be (largely) agnostic to peptide side chain structure to facilitate the synthesis of diverse and evolving peptides. Accordingly, the crucial aspect of the emergence of macromolecule-controlled peptide synthesis, and therefore ultimately coded peptide synthesis, must be the formation of aminoacyl-thiols or aminoacyl-ribonucleotides. However, neither amino acid thioesters nor amino acid esters are known to form selectively in water, posing the question of their prebiotic provenance. Consequently, how these monomers, which are both electrophilically activated and nucleophilic, can be

effectively formed in water is the key challenge that must be resolved to begin to unpick the origin of peptide biosynthesis.

Esters and thioesters are classically formed by dehydration of their respective carboxylates under anhydrous conditions in the presence of an acid or a condensing agent.[9,10,11] Typically, dehydration reactions are considered to be disfavoured in water.[12] This problem is exacerbated for (thio)esters of amino acids, where the presence of a nucleophilic amine makes condensation poorly selective due to both hydrolysis and polymerisation.[13] Synthetically, the amine of an amino acid (**1**) must be blocked by a protecting group prior to condensation, often in dry conditions or anhydrous solvent, preventing significant hydrolysis and self-reaction.[14] However, life performs chemistry in water and does not protect amino acids prior to functionalisation. Life's amino acid (thio)esters are inherently hydrolytically labile and exist transiently under physiological conditions, and their (non-enzymatic) accumulation must balance their rate of formation and consumption. Importantly, the formation of activated amino acid (thio)esters must occur without uncontrolled polymerisation:[15] this is a prerequisite of controlled peptide synthesis. The formation of coded peptides must outcompete random peptide synthesis to rise above the noise-threshold for coding.[16] The universal and fundamental role that amino acid esters play in biochemistry led us to suspect aminoacyl-thiols (or even aminoacyl-RNAs) would form spontaneously and selectively in water without the need for complex macromolecular catalysts, and importantly without spontaneous formation of peptides.

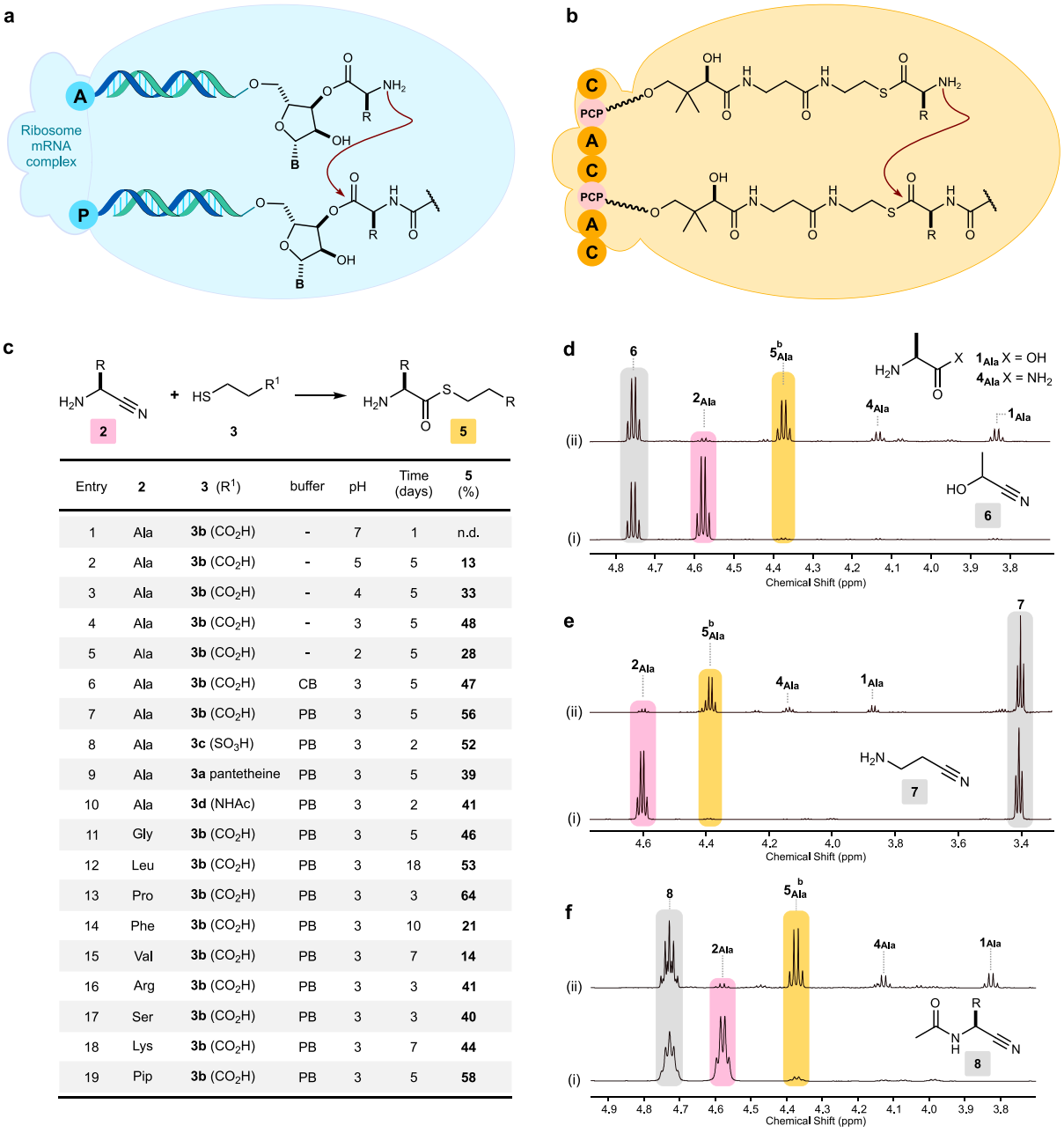
Aminonitriles (**2**) are ambiphilic molecules, capable of reacting either as a nucleophile or an electrophile, however, importantly, they exhibit kinetic stability whilst being thermodynamically higher in energy than peptides. We have previously discovered efficient catalytic mechanisms by which nitriles can be converted into peptides with excellent selectivity, tolerating all proteinogenic side chains.[17,18,19] Thiol (**3**) nucleophiles were the key catalysts that unlocked the energy in these nitriles. In principle, thiol **3** is also a handle by which the nitrile can be covalently linked to a macromolecule, raising the question of whether prebiotic nitrile chemistry can shift directly into the mode of peptide synthesis used by extant life. Furthermore, we have recently discovered a selective route by which pantetheine (**3a**), the universal biological thioester-bearing moiety that is central to non-ribosomal peptide synthesis (Fig. 1b), forms in water from nitriles.[20] This generational link between nitriles, biological peptides and biological thiols, led us to consider the direct link between nitriles and activated amino acid (thio)esters more closely.

Considering how nitrile chemistry could smoothly transition into biochemistry, we realised that thiols might play a crucial role. Thiols react with nitriles to form thioimidates that react spontaneously with amino acids (**1**) or amino amides (**4**) to form proteinogenic α-peptides.[18] However, under acidic conditions general acid-catalysed hydrolysis to thioesters is observed,[21,22] and we suspected that the lower reactivity of thioesters would inhibit spontaneous peptide synthesis. This led us to suppose that thioesters would be accessible from nitriles under mild aqueous conditions. Specifically, we thought that aminoacyl-thiols (**5**), the active intermediates in non-ribosomal peptide synthesis (Fig. 1b), would be accessible from the reaction of an aminonitrile (**2**) and a thiol (**3**). We also suspected that this reactivity would unlock a prebiotic (non-enzymatic) synthesis of aminoacyl-RNAs (Fig. 1a), a long-standing and hereto unresolved challenge,[23,24,25] and the first step towards elucidating the origins of coded peptide biosynthesis.

Aminoacyl thioester synthesis

We began our investigation by studying the reaction of alanine nitrile ($\mathbf{2_{Ala}}$) with 3-mercaptopropionic acid (**3b**), which we had previously observed to be the most effective catalyst for catalytic peptide ligation.[18] No thioester was observed at pH 7 – 9 (Fig. 1c, Entry 1), but under mildly acidic conditions (pH 5–3) alanine thioester ($\mathbf{5^{b}_{Ala}}$) was formed in good yield (13–48%) (Fig. 1c, Entry 2–4). We reasoned that the selectivity for aminoacyl-thiol **5** (over amino amide **4**) formation would be improved by general-acid catalysis, and phosphate was found to modestly increase the yield of thioester $\mathbf{5^{b}_{Ala}}$ (56%, Fig. 1c, Entry 7). A range of

simple and biological thiols, including pantetheine (**3a**) and coenzyme M (**3c**), reacted with aminonitriles (**2**) to yield aminoacyl-thiols (**5**) under these conditions. The non-enzymatic synthesis of pantetheine aminoacyl-thiol ($\mathbf{5}^{\mathbf{a}}_{\mathbf{Aaa}}$), coupled with our recent selective synthesis of pantetheine (**3a**),[20] completes a direct synthetic chain from simple nitriles to the aminoacyl-thiols used by life to facilitate non-ribosomal peptide synthesis (Fig. 1b).



Entry	2	3 (R^1)	buffer	pH	Time (days)	5 (%)
1	Ala	**3b** (CO_2H)	-	7	1	n.d.
2	Ala	**3b** (CO_2H)	-	5	5	**13**
3	Ala	**3b** (CO_2H)	-	4	5	**33**
4	Ala	**3b** (CO_2H)	-	3	5	**48**
5	Ala	**3b** (CO_2H)	-	2	5	**28**
6	Ala	**3b** (CO_2H)	CB	3	5	**47**
7	Ala	**3b** (CO_2H)	PB	3	5	**56**
8	Ala	**3c** (SO_3H)	PB	3	2	**52**
9	Ala	**3a** pantetheine	PB	3	5	**39**
10	Ala	**3d** (NHAc)	PB	3	2	**41**
11	Gly	**3b** (CO_2H)	PB	3	5	**46**
12	Leu	**3b** (CO_2H)	PB	3	18	**53**
13	Pro	**3b** (CO_2H)	PB	3	3	**64**
14	Phe	**3b** (CO_2H)	PB	3	10	**21**
15	Val	**3b** (CO_2H)	PB	3	7	**14**
16	Arg	**3b** (CO_2H)	PB	3	3	**41**
17	Ser	**3b** (CO_2H)	PB	3	3	**40**
18	Lys	**3b** (CO_2H)	PB	3	7	**44**
19	Pip	**3b** (CO_2H)	PB	3	5	**58**

Figure 1 | Synthesis of aminoacyl-thiols in water. a) Schematic representation of ribosomal peptide synthesis, which is mediated by aminoacyl-RNA esters. **b)** Schematic representation of non-ribosomal peptide synthesis, which is mediated by aminoacyl-pantetheine thioesters. **c)** The yield of aminoacyl-thiol **5** from the reaction of α-aminonitrile **2** (200 mM) and thiol **3** (9 equiv.) at room temperature. PB = phosphate; CB = citrate; n.d. = not detected; Pip = pipecolic acid. **d-f)** ^{1}H Nuclear magnetic resonance (NMR) spectra to show the selective synthesis of alanine thioester ($\mathbf{5}_{\mathbf{Ala}}$, 50–55%) alanine nitrile ($\mathbf{2}_{\mathbf{Ala}}$, 200 mM) and thiol **3b** (9 equiv.) in the presence of potential competitors: **d)** lactonitrile (**6**, 200 mM), alanine nitrile ($\mathbf{2}_{\mathbf{Ala}}$, 200 mM) and thiol **3b** (9 equiv.) at pH 4 after: (i) 1 hour and (ii) 5 days, which yielded thioester

$\mathbf{5^{b}_{Ala}}$ (55%) alongside residual **6** (97%); or **e)** β-alanine nitrile (**7**, 200 mM), alanine nitrile ($\mathbf{2_{Ala}}$, 400 mM) and thiol **3b** (9 equiv.) at pH 3 after: (i) 1 hour and (ii) 5 days, which yielded thioester $\mathbf{5^{b}_{Ala}}$ (50%) alongside residual **7** (96%); or **f)** *N*-acetyl alanine nitrile (**8**, 200 mM), alanine nitrile ($\mathbf{2_{Ala}}$, 200 mM) and thiol **3b** (9 equiv.) at pH 3 after: (i) 1 hour and (ii) 5 days, which yielded thioester $\mathbf{5^{b}_{Ala}}$ (50%) alongside residual **8** (96%).

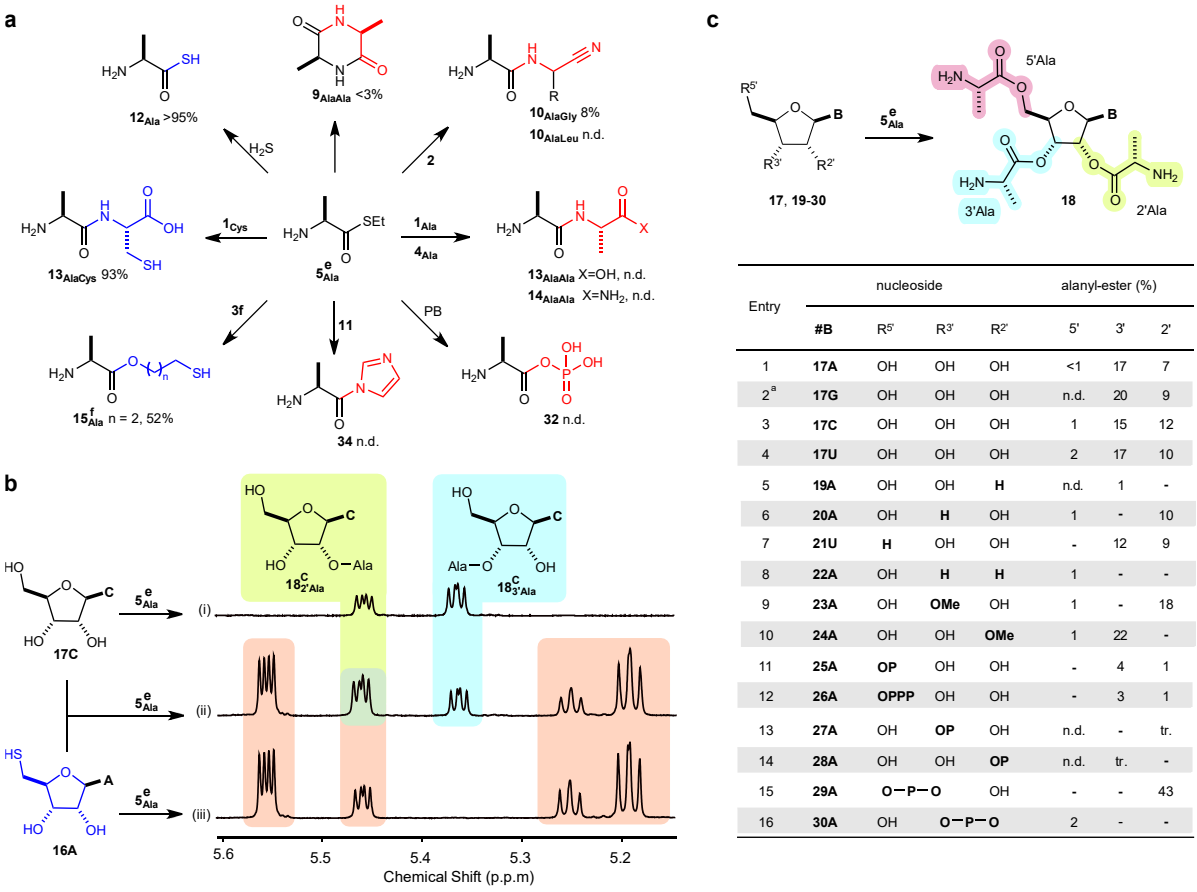
The formation of thioester **5** from aminonitrile **2** was most effective under acidic conditions, which initially seemed prebiotically unrealistic. However, phosphate, as well as being a catalyst for thioester formation, could drive (reversible) pH change from neutral to acidic conditions. Freezing sodium phosphate solutions leads to an ideal drop in pH, due to the lower solubility of dibasic phosphate, the eutectic phases formed on freezing neutral phosphate solutions stabilised at pH 3.5,[26] the sweet spot for aminoacyl-thiol **5** formation. Accordingly, we found that solutions of nitrile $\mathbf{2_{Ala}}$ and thiol **3b**, buffered by phosphate at pH 7, yielded alanine thioester $\mathbf{5^{b}_{Ala}}$ efficiently, albeit slowly, at -7 ºC. At high dilution (without freezing) minimal reaction between aminonitrile $\mathbf{2_{Ala}}$ and thiol **3b** was observed. However, with freezing the initial conditions could in principle be almost indefinitely dilute due to the concentrating effects of forming the eutectic phase, and incubating alanine nitrile $\mathbf{2_{Ala}}$ (1 mM) with thiol **3b** (10 mM) at -7 ºC led to effective synthesis of alanine thioester $\mathbf{5^{b}_{Ala}}$ (48%) after 30 days.

We next turned our attention to other proteinogenic side chains, selecting those we considered informative (Fig. 1c, Entries 11–19). Simple (alkyl) aminonitriles $\mathbf{2_{Gly}}$, $\mathbf{2_{Pro}}$ and $\mathbf{2_{Leu}}$ underwent good conversion to their respective thioesters $\mathbf{5^{b}_{Aaa}}$ (46–64%). The conversion of $\mathbf{2_{Phe}}$ (21%) and $\mathbf{2_{Val}}$ (14%) were lower, likely due to thioimidate-tautomerisation. However, hydrophilic aminonitriles, $\mathbf{2_{Ser}}$ and $\mathbf{2_{Arg}}$, were readily converted to serine thioester $\mathbf{5^{b}_{Ser}}$ (40%) and arginine thioester $\mathbf{5^{b}_{Arg}}$ (41%) in good yield. Even lysine aminonitrile ($\mathbf{2_{Lys}}$), which contains a primary ε-amine in its sidechain, formed lysine thioester $\mathbf{5^{b}_{Lys}}$ (44%) in comparable yield, alongside (proline-homolog) pipecolate thioester $\mathbf{5^{b}_{Pip}}$ (26%).[27,28,29] These results demonstrate that a wide range of proteinogenic, and closely related, α-aminonitriles (**2**), are readily converted in comparable yield to their respective aminoacyl-thiols (**5**). On the other hand, strikingly, non-proteinogenic α-hydroxy nitrile **6**, β-aminonitrile **7**, and acylated α-aminonitrile **8**, all underwent <5% reaction with thiol **3b** under these conditions (Fig. 1d-f). The observed selectivity for the proteinogenic aminoacyl structure is chemically disposed by the highly electron withdrawing protonated free α-amine group, which is essential to biomimetic peptide synthesis.

Aminoacyl thioester reactivity

We next investigated the reactivity of aminoacyl-thiols **5**, and particularly whether their selective reaction (with other nucleophiles) might outcompete self-condensation. For ambiphilic molecules it is often difficult to avoid self-condensation.[30] However, surprisingly, at neutral pH in water we found that alanine thioester L-$\mathbf{5^{e}_{Ala}}$ (100 mM) reacted very inefficiently with itself, only affording a low (<3%) yield of diketopiperazine $\mathbf{9_{AlaAla}}$ after 5 days, alongside alanine ($\mathbf{1_{Ala}}$) (Fig. 2a). At higher pH more DKP (**9**) formation was observed, but the major pathway remained hydrolysis of thioester $\mathbf{5^{e}_{Ala}}$ to alanine ($\mathbf{1_{Ala}}$). Glycine nitrile ($\mathbf{2_{Gly}}$), the most nucleophilic α-aminonitrile, also reacted ineffectively with thioester $\mathbf{5^{e}_{Ala}}$, yielding H-Ala-Gly-CN $\mathbf{10_{AlaGly}}$ in remarkably low (8%) yield at pH 7. This starkly contrasts the near-quantitative $\mathbf{2_{Gly}}$ ligations that have been observed with activated peptide-thioacid electrophiles at neutral pH.[27, 31] Moreover, other aminonitriles (e.g., $\mathbf{2_{Leu}}$), as well as amino acid $\mathbf{1_{Ala}}$ and amide $\mathbf{4_{Ala}}$ were not observed to react with thioester $\mathbf{5^{e}_{Ala}}$ to furnish their respective dipeptides (Fig. 2a). The surprisingly low reactivity of thioester **5** with amines prompted us to investigate other nucleophiles.

Phosphate, glycerol and imidazole (**11**) did not react with thioester **5** (Fig. 2a). However, sulfur-containing nucleophiles react rapidly and efficiently with thioesters, so we were unsurprised that the reaction of aminoacyl-thiols (**5**) with hydrogen sulfide (H_2S) afforded thioacids (**12**) in excellent-to-quantitative yield (Fig. 2a). Cysteine ($\mathbf{1_{Cys}}$) also reacted with thioester **5** to yield cysteinyl-dipeptides (e.g., $\mathbf{13_{AlaCys}}$) in near-quantitative conversion, by sulfide exchange and intramolecular *S*-to-*N* acyl-shift. Similarly, mercapto-alcohols reacted with aminoacyl-thiols (**5**) to yield esters via intramolecular *S*-to-*O* exchange. Therefore γ-mercapto-propanol (**3f**) reacted sluggishly but cleanly to form ester **15** in good yield (Fig. 2a). However, whilst the shorter homolog, β-mercapto-ethanol (**3g**), was observed to form esters these underwent subsequent 3-*exo-tet* cyclisation to form thiirane and amino acid **1**.[32] The direct reaction of aminonitriles (**2**) with γ-mercapto-propanol (**3f**) also yielded esters **15** in good-to-excellent yield (72–87%). These facile intramolecular esterifications led us to consider how aminoacyl-thiols (**5**) could be intercepted by nucleic acids, which would ultimately enable nucleic acid-mediated peptide ligation, and be a necessary step towards the development of coded peptide synthesis.



Entry		nucleoside			alanyl-ester (%)		
	#B	R5'	R3'	R2'	5'	3'	2'
1	**17A**	OH	OH	OH	<1	17	7
2[a]	**17G**	OH	OH	OH	n.d.	20	9
3	**17C**	OH	OH	OH	1	15	12
4	**17U**	OH	OH	OH	2	17	10
5	**19A**	OH	OH	**H**	n.d.	1	-
6	**20A**	OH	**H**	OH	1	-	10
7	**21U**	**H**	OH	OH	-	12	9
8	**22A**	OH	**H**	**H**	1	-	-
9	**23A**	OH	**OMe**	OH	1	-	18
10	**24A**	OH	OH	**OMe**	1	22	-
11	**25A**	**OP**	OH	OH	-	4	1
12	**26A**	**OPPP**	OH	OH	-	3	1
13	**27A**	OH	**OP**	OH	n.d.	-	tr.
14	**28A**	OH	OH	**OP**	n.d.	tr.	-
15	**29A**	**O–P–O**		OH	-	-	43
16	**30A**	OH	**O–P–O**		2	-	-

Figure 2 | Reactivity of amine, hydroxyl and sulfide nucleophiles with alanine thioester. a) Alanine thioester L-$\mathbf{5^e_{Ala}}$ (100 mM) reacted ineffectually with amines, imidazole (**11**) and phosphate nucleophiles (red) at neutral pH, but in excellent-to-quantitative yield with sulfur nucleophiles (blue). Yields are maximum %-conversion observed with specified nucleophile (150 mM)* at pH 7 and room temperature. PB = phosphate buffer. n.d. = not detected. * = 140 mM $\mathbf{2_{Gly}}$; 220 mM $\mathbf{2_{Leu}}$; 200 mM PB; 200 mM **11**. Structures are drawn neutral for clarity. **b)** ^{1}H NMR spectra to show the reaction of: (i) L-$\mathbf{5^e_{Ala}}$ (200 mM) and cytidine **17C** (20 mM) at pH 6.5 which yields aminoacyl-cytidine $\mathbf{18^C_{Ala}}$ (25%) after 24 hours, (ii) L-$\mathbf{5^e_{Ala}}$ (100 mM), **17C** (20 mM) and 5'-mercapto-adenosine **16A** (15 mM) at pH 6.5 after 24 hours, and (iii) L-$\mathbf{5^e_{Ala}}$ (100 mM) and **16A** (20 mM) at pH 6.5, where 2',3'-diol-aminoacylation (54%) is observed after 24 hours. **c)** Alcohol-aminoacylation yield for specified nucleosides (20 mM) with alanine thioester L-$\mathbf{5^e_{Ala}}$ (200 mM)

at pH 6.5 and room temperature after 18 hours. OP = monophosphate; OPPP = triphosphate; O–P–O = phosphodiester; tr. = trace (<0.5%); n.d. = not detected. a = 2 mM **17G**.

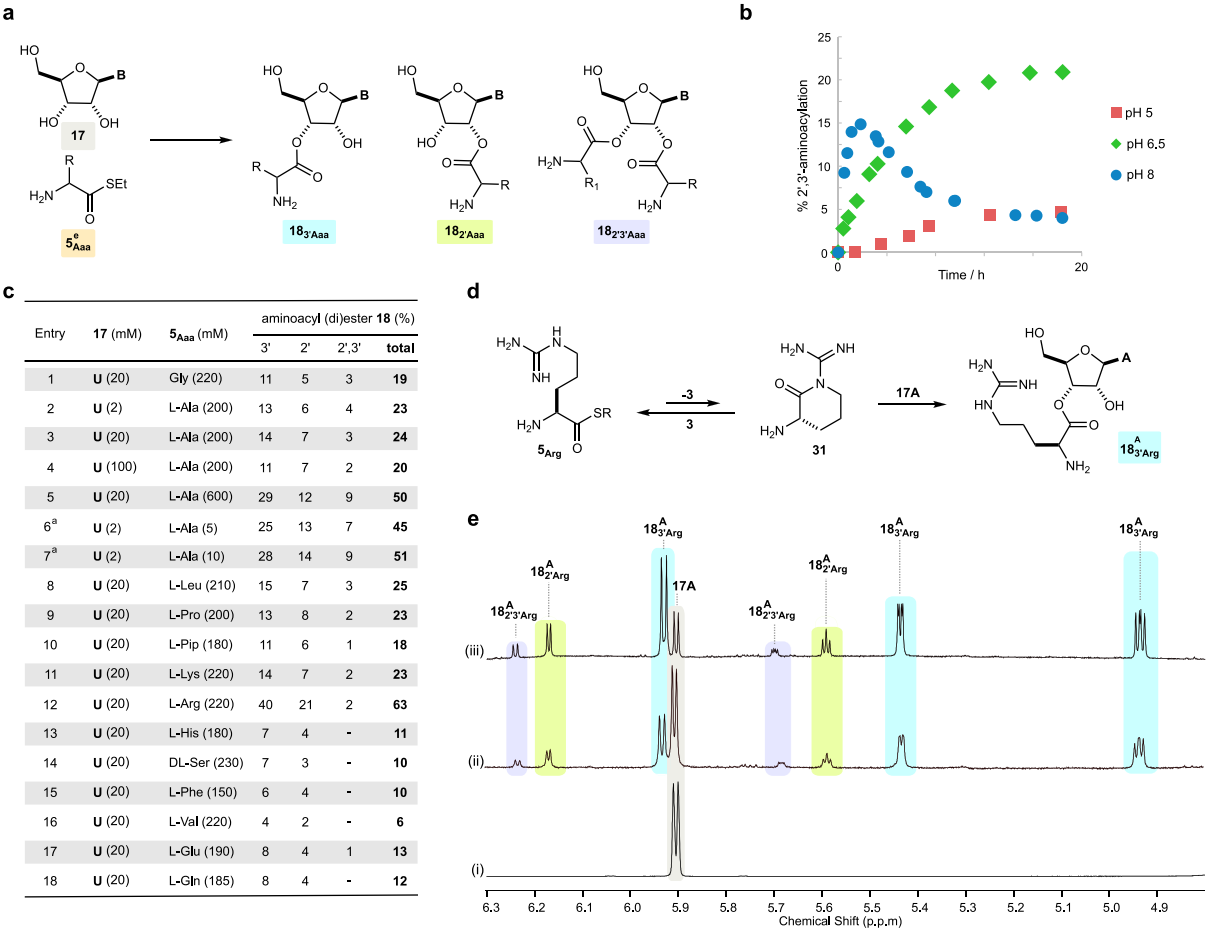
We therefore next incubated 5'-mercapto-adenosine **16A** with alanine thioester L-$\mathbf{5}^{\mathbf{e}}_{\mathbf{Ala}}$ and observed 2',3'-diol-aminoacylation (54%) (Fig. 2b). Surprisingly, addition of canonical nucleoside cytidine (**17C**), as a competitor nucleophile, led to intermolecular 2',3'-aminoacylation of cytidine (**17C**) (Fig. 2b). Indeed, incubating cytidine (**17C**, 20 mM) alone, with alanine thioester (L-$\mathbf{5}^{\mathbf{e}}_{\mathbf{Ala}}$, 200 mM) led to the formation of alanyl-cytidine ($\mathbf{18}^{\mathbf{C}}_{\mathbf{Ala}}$) in 25% yield at pH 6.5 (Fig. 2b). Ribonucleoside-2',3'-aminoacylation with thioester **5** was observed despite the large (~3000-fold) excess of water and 10-fold excess of amine relative to the diol. This efficient and selective nucleoside-aminoacylation stands in stark contrast to previous reports in which evolved catalysts or intramolecular tethering were required to enable RNA-aminoacylation in water.[33,34]

Having observed the aminoacylation of cytidine (**17C**) in water, we next considered the selectivity of nucleoside aminoacylation. Aminoacylation was 2',3'-selective, due to the low pK_a of the 2',3'-diol (Fig. 2c). Therefore, 2'-deoxynucleosides **19** furnished very low yields of aminoacylation, while non-canonical 3'-deoxyadenosine **20A** and 5'-deoxyuridine **21U** underwent substantial 2'-aminoacylation and 2',3'-aminoacylation, respectively. Aminoacylation was, as expected, nearly completely suppressed for 2',3'-dideoxyadenosine **22A**. On the other hand, thioester **5** reacted effectively with both 2'-*O*-methoxy-adenosine **23A** and 3'-*O*-methoxy-adenosine **24A**. Both **23A** and **24A** furnished comparable aminoacylation yields, suggesting that the 2'- and 3'-hydroxyls are comparably reactive in the RNA-2',3'-diol.

Next, we investigated the impact of nucleoside phosphorylation on aminoacylation. Nucleoside-5'-monophosphates (NMP, **25**) and nucleotide-5'-triphosphates (NTP, **26**) were substantially less reactive than nucleosides (**17**), whilst overall levels of thioester (**5**) turnover (to hydrolysis) remained similar, indicating that aminoacylation was slower, rather than aminoacyl-ester hydrolysis being faster. Interestingly, negligible aminoacylation of adenosine-3'-monophosphate (**27A**) or adenosine-2'-monophosphate (**28A**) was observed; consistent with the high pK_a of their 2'- and 3'-alcohols relative to a 2',3'-diol. This pK_a suppression was specific to phosphate monoesters (and their polyphosphates) – phosphodiesters such as adenosine-3',5'-cyclic phosphate (**29A**) underwent highly effective 2'-aminoacylation whilst 5'-aminoacylation of adenosine-2',3'-cyclic phosphate **30A** remained ineffective. The suppressed aminoacylation of 2',3'-diols in NMPs and NTPs may play an important role in directing aminoacylation to oligonucleotide-2',3'-diols, under conditions where monomers are concurrently available (for RNA-synthesis) alongside RNA-aminoacylation.

Aminoacylation was pH and thioester concentration dependent. Nucleoside aminoacylation was most rapid at pH 8.0, but was more 2',3'-diol selective at pH 6.5. Additionally, the product (i.e., aminoacyl-nucleoside **18**) is most stable at low pH, and so ribonucleoside aminoacylation was slow but highly effective at pH 6.5 (Fig. 3b). Furthermore, although thioester **5** dimerization was ineffectual under all conditions studied, the formation of DKP (**9**) was pH dependent. Therefore, whilst 15% DKP $\mathbf{9}_{\mathbf{AlaAla}}$ was observed at pH 8, negligible (<3%) $\mathbf{9}_{\mathbf{AlaAla}}$ was observed at pH 6.5. Aminoacylation was substantially higher yielding at high thioester concentrations, therefore 600 mM thioester L-$\mathbf{5}^{\mathbf{e}}_{\mathbf{Ala}}$ yielded 50% aminoacyl-uridine ($\mathbf{18}^{\mathbf{U}}_{\mathbf{Ala}}$) after 24 hours at pH 6.5. Minimal ribonucleoside aminoacylation was observed at low (<20 mM) thioester **5** concentration, however upon freezing the initial conditions could be highly dilute (1-10 mM) with respect to both the thioester **5** and the nucleoside **17**, and still yield nucleoside-aminoacylation. For example, incubating 10 mM alanine thioester $\mathbf{5}^{\mathbf{e}}_{\mathbf{Ala}}$ with 2 mM uridine (**17U**) at -7 °C led to highly effective aminoacylation, and alanyl-uridine $\mathbf{18}^{\mathbf{U}}_{\mathbf{Ala}}$ formed in 58% yield (51% 2',3'-Ala **18U** + 7% 5'-Ala **18U**). Nucleoside-aminoacylation results from judiciously balanced reactivity of the nucleoside-2',3'-diol and aminoacyl-thiol **5**. For example, reaction at the diol was blocked by the acetylation of thioester **5**; no

nucleotide-acylation was observed upon incubating nucleoside **17U** with amido-thioester Ac-Ala-SEt at pH 6.5, demonstrating the importance of the amine moiety, as well as the thioester moiety, in promoting selective reaction at the 2',3'-diol. Crucially, thioester-mediated ribonucleoside aminoacylation was extremely efficient relative to the poor reactivity displayed with amines (Fig. 2a). This led us to consider that side chain compatibility, which has not previously been probed in other aminoacylation studies, would be high.



Entry	17 (mM)	5_{Aaa} (mM)	aminoacyl (di)ester 18 (%)			
			3'	2'	2',3'	total
1	U (20)	Gly (220)	11	5	3	19
2	U (2)	L-Ala (200)	13	6	4	23
3	U (20)	L-Ala (200)	14	7	3	24
4	U (100)	L-Ala (200)	11	7	2	20
5	U (20)	L-Ala (600)	29	12	9	50
6[a]	U (2)	L-Ala (5)	25	13	7	45
7[a]	U (2)	L-Ala (10)	28	14	9	51
8	U (20)	L-Leu (210)	15	7	3	25
9	U (20)	L-Pro (200)	13	8	2	23
10	U (20)	L-Pip (180)	11	6	1	18
11	U (20)	L-Lys (220)	14	7	2	23
12	U (20)	L-Arg (220)	40	21	2	63
13	U (20)	L-His (180)	7	4	-	11
14	U (20)	DL-Ser (230)	7	3	-	10
15	U (20)	L-Phe (150)	6	4	-	10
16	U (20)	L-Val (220)	4	2	-	6
17	U (20)	L-Glu (190)	8	4	1	13
18	U (20)	L-Gln (185)	8	4	-	12

Figure 3 | Chemoselective aminoacylation of ribonucleoside-2',3'-diol in water. a) The reaction of aminoacyl-thiol **5** with ribonucleoside **17** yields a mixture of 2'-aminoacyl-nucleoside ($\mathbf{18_{2'Aaa}}$), 3'-aminoacyl-nucleoside ($\mathbf{18_{3'Aaa}}$) and 2',3'-di(aminoacyl)-nucleoside ($\mathbf{18_{2'3'Aaa}}$). **b)** 2',3'-Aminoacyl-nucleoside **18** yield observed upon incubating alanine thioester L-$\mathbf{5^{e}_{Ala}}$ (200 mM) with adenosine **17A** (20 mM) at room temperature in 2-(*N*-morpholino)ethanesulfonic acid (MES; 1 M, pH 5–6.5) or 3-(*N*-morpholino)propanesulfonic acid (MOPS; 1 M, pH 8) buffer. **c)** Yield of aminoacyl-uridine ($\mathbf{18^{U}_{Aaa}}$) observed upon incubating uridine (**17U**) with ethyl aminoacyl-thiol $\mathbf{5^{e}_{Aaa}}$ at pH 6.5 and room temperature after 24 hours. a = eutectic reaction: ~30 minutes at -80 °C and then 5 days at -7 °C. **d)** The reaction of arginine thioester L-$\mathbf{5^{e}_{Arg}}$ was accelerated by intramolecular nucleophilic side chain catalysis, and in situ formation of guanidine **31**. **e)** ^{1}H NMR spectra to show: (i) adenosine **17A** (2 mM); and (ii–iii) the reaction of arginine thioester L-$\mathbf{5^{e}_{Arg}}$ (120 mM) with adenosine **17A** (2 mM) at pH 6.5 after: (ii) 1 hour, and (iii) 6 hours, which formed arginyl-adenosine $\mathbf{18^{A}_{Arg}}$ (81%) in excellent yield.

Aminoacyl-thiols with simple or lipophilic side chains (e.g., $\mathbf{5^{e}_{Gly}}$, $\mathbf{5^{e}_{Leu}}$, $\mathbf{5^{e}_{Pro}}$) reacted similarly to alanine thioester $\mathbf{5^{e}_{Ala}}$, although as expected β-branched thioesters (e.g., $\mathbf{5^{e}_{Val}}$) reacted sluggishly (Fig. 3c). Thioesters

with weakly nucleophilic side chains (e.g., $\mathbf{5^{e}_{His}}$, $\mathbf{5^{e}_{Glu}}$, $\mathbf{5^{e}_{Gln}}$) formed aminoacyl-nucleosides **18** in good, albeit slightly lower steady state, yield. However, surprisingly, arginine thioester L-$\mathbf{5^{e}_{Arg}}$ was extremely effective at ribonucleoside aminoacylation, furnishing arginyl-adenosine $\mathbf{18^{A}_{Arg}}$ in up to 81% yield (Fig. 3d, e). Cyclic arginine **31** formed in situ and aminoacylated nucleosides even more effectively than thioester $\mathbf{5^{e}_{Arg}}$. For example, 40 mM **31** reacted with 20 mM adenosine **17A** to furnish arginyl-adenosine $\mathbf{18^{A}_{Arg}}$ in 40% yield after only 30 min, indicating pronounced intramolecular nucleophilic catalysis. To our knowledge this mode of side chain nucleophilic catalysis has not previously been reported in enzyme catalysis, but it seems highly likely enzymes would have exploited this catalytic strategy. Given this surprising cyclisation (with Arg, a side chain normally considered non-nucleophilic), one might expect lysine thioester $\mathbf{5^{e}_{Lys}}$ to cyclise rather than aminoacylate. However, ribonucleoside aminoacylation with lysine thioester $\mathbf{5^{e}_{Lys}}$ proceeded as effectively as aminoacylation with alanine thioester $\mathbf{5^{e}_{Ala}}$. Serine thioester $\mathbf{5^{e}_{Ser}}$ contains an alcohol moiety and so could in principle react as an intermolecular alcohol-nucleophile like ribonucleosides. However, serine thioester DL-$\mathbf{5^{e}_{Ser}}$ was observed to aminoacylate uridine (**17U**) in moderate (10%) yield; moreover, no serine acylation was observed despite a 10-fold excess of the α,β-amino-alcohol over nucleoside **17**, demonstrating remarkable selectivity for the 2',3'-diol. Overall excellent proteinogenic sidechain tolerance was observed during nucleoside aminoacylation, suggesting aminoacyl-thiol (**5**) is a privileged substrate for the aminoacylation of RNA in water.

Thiol-catalysed aminoacylation

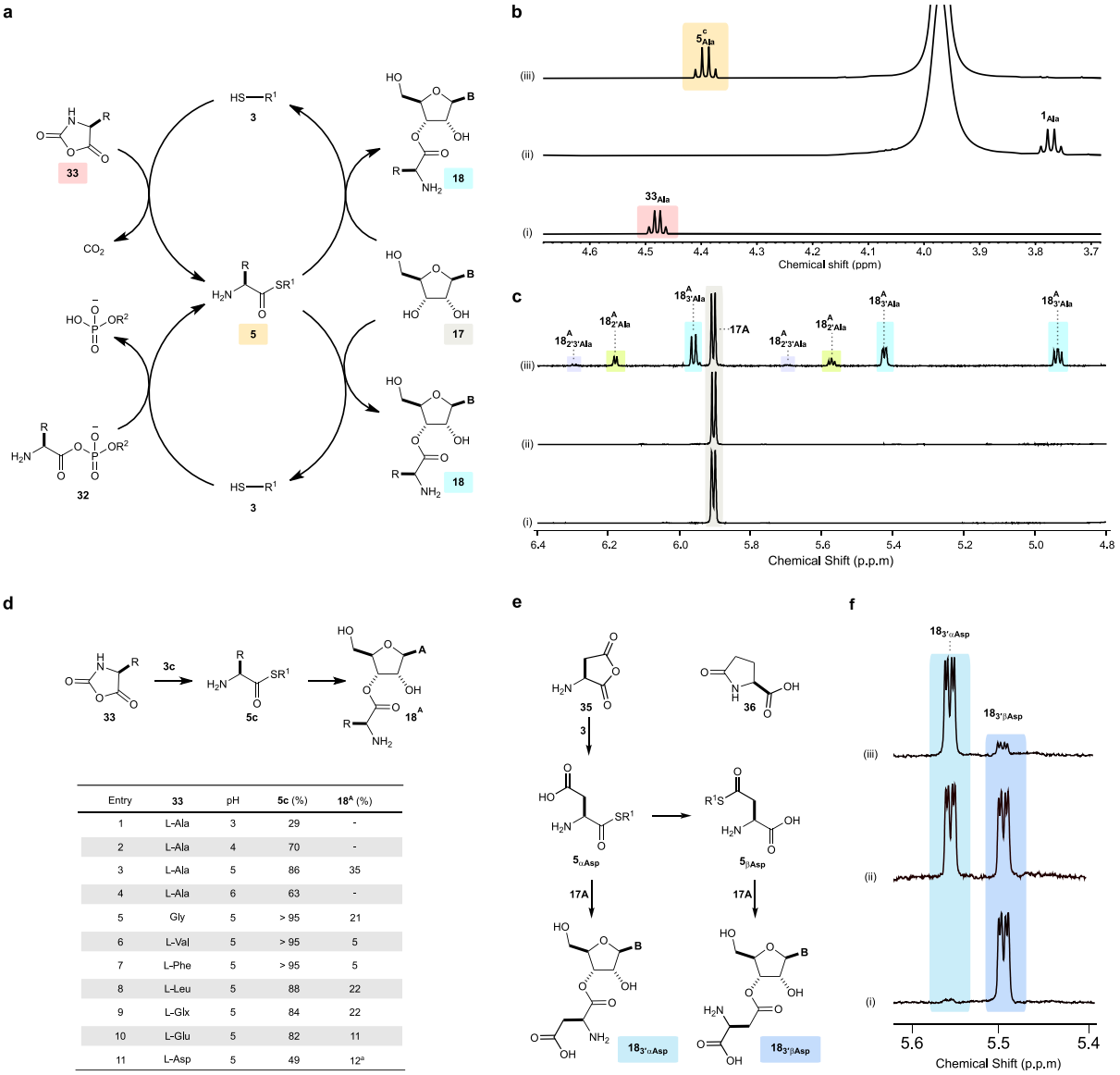
Nucleoside-acylation has been achieved with various acyl donors, including acyl-imidazoles,[35,36] acyl-thioacids[37] and acyl-phosphates.[38,39] However, previously aminoacylation (with an unprotected primary or secondary amine) has proven much more challenging due to the presence of the nucleophilic amine, and the poor stability of the resultant aminoacyl-ester relative to other esters.[40] Previous studies of nucleoside-aminoacylation have largely focused on aminoacyl-phosphates **32** (Fig. 4a),[38,39,41] but these phosphates are highly unstable in water and difficult to prepare without substantial hydrolysis. Problematically highly activated amino acid substrates, such as **32**, are not selective electrophiles for 2',3'-diols and so peptide formation is a major reaction pathway.[41] Additionally, how aminoacyl-phosphates could form in water remains an open question. Possible candidate syntheses have been suggested to proceed via *N*-carboxy anhydrides (NCAs, **33**)[42,43] or phosphorimidazolides **34**.[24] For example, NCAs **33**, albeit only with bulky lipophilic side chains (i.e., $\mathbf{33_{Phe}}$, $\mathbf{33_{Val}}$), react with nucleotide phosphates to form mixed anhydrides in low yields, and have been shown to aminoacylate (model) ribonucleoside-3'-monophosphates (**26A**), *via* intramolecular transfer from the 3'-phosphate to the 2'-hydroxyl.[42] However, NCA-aminoacylation of RNA-2',3'-diols (via aminoacyl-phosphates) has not been observed either intra- or intermolecularly due to the concentration dependent self-inhibition of aminoacylation (i.e., NCA **33** self-condensation). We suspected that the poor reactivity of aminoacyl-phosphates **32** and **33** would be ameliorated by a thiol catalyst, where the dominant pathway for reaction would be via aminoacyl-thiol **5**. Moreover, we expected that thiol-catalysed aminoacylation would allow unprecedented side chain scope for (in situ) aminoacylation of RNAs.

To test this hypothesis, we next incubated a range of NCAs **33** with thiols **3**. We observed rapid and highly efficient synthesis of aminoacyl-thiol **5** in good-to-quantitative (49-95%) yield from NCAs **33** (Fig. 4). Thioester synthesis was highly effective and selective from **33**. Simple and hydrophobic NCAs (e.g., $\mathbf{33_{Gly}}$, L-$\mathbf{33_{Phe}}$, L-$\mathbf{33_{Val}}$) furnished their respective aminoacyl-thiol (**5**) in near-quantitative yield. Aminoacyl-phosphate $\mathbf{32_{Ala}}$ (20 mM, R^2=adenosyl) also reacted with thiol **3d** to furnish alanine thioester $\mathbf{5^{d}_{Ala}}$ in reasonable unoptimized yield (42%). Notably, spontaneous self-condensation of these activated amino acids (**32**–**33**) was dramatically suppressed by thiol **3**; without thiols NCAs **33** underwent self-condensation

and hydrolysis, which blocked nucleoside-aminoacylation. Thiol-suppressed NCA-self-condensation intriguingly opened an opportunity, for the first time, to exploit high concentrations of NCAs **33** for RNA-aminoacylation, as well as a specific opportunity for aminoacylation of RNA-2',3'-diols.

Therefore, we next incubated 600 mM Ala-NCA (L-$\mathbf{33}_{\mathbf{Ala}}$) with coenzyme M (**3c**) and adenosine (**17A**) in water. After 16 hours the resultant solution was pH 6, and we observed ribonucleoside aminoacylation furnishing alanyl-adenosine $\mathbf{18}^{\mathbf{A}}_{\mathbf{Ala}}$ in 35% yield (Fig. 4c, spectrum iii). This effective thiol-mediated aminoacylation stands in stark contrast to the reaction in the absence of thiol; with no thiol, no nucleoside-aminoacylation was observed (Fig. 4c, spectrum ii) demonstrating the remarkable enhancement of aminoacylation by thiols **3**.

We next turned our attention to aspartate (Asp) that possesses two adjacent carboxylate moieties, and therefore poses the most challenging (proteinogenic) amino acid disposition for chemoselective activation and α-selective aminoacylation. Asp-NCA ($\mathbf{33}_{\mathbf{Asp}}$) yielded aminoacyl-thiol $\mathbf{5}^{\mathbf{c}}_{\mathbf{Asp}}$ in 49% yield with excellent α-selectivity (α/β 11:1). The in-situ reaction of $\mathbf{5}^{\mathbf{c}}_{\mathbf{Asp}}$ with adenosine (**17A**) yielded aspartyl-adenosine $\mathbf{18}^{\mathbf{A}}_{\mathbf{Asp}}$ (12%) again with excellent α-selectivity after 18 hours. Aspartyl-thioester $\mathbf{5}^{\mathbf{c}}_{\mathbf{Asp}}$ slowly isomerised to its β-isomer $\mathbf{5}^{\mathbf{c}}_{\mathbf{\beta Asp}}$, which led to selective β-aminoacylation after 5 days, however, initial α-selectivity was strongly favoured. Moreover, α-selectivity was not only favoured by the inherent (α-amine tethered) α-activation of NCA **33** but also for anhydride $\mathbf{35}_{\mathbf{Asp}}$, despite dual activation of both the α- and β-carboxylates. Remarkably, $\mathbf{35}_{\mathbf{Asp}}$ reacted with thiols **3** to selectively yield α-aspartyl-thioester ($\mathbf{5}^{\mathbf{c}}_{\mathbf{Asp}}$, 45%) in water with excellent α-selectivity (α/β 12:1). Therefore, incubating anhydride $\mathbf{35}_{\mathbf{Asp}}$, thiol **3c** and adenosine **17A** in water at pH 6.5 resulted in selective α-aminoacylation of the ribonucleoside-2',3'-diol (12% $\mathbf{18}_{\mathbf{\alpha Asp}}$) after 12 hours, demonstrating a profound kinetic selectivity for natural (proteinogenic) α-aminoacylation, irrespective of α-amine tethered-activation. Racemisation of Asp was not observed; $\mathbf{18}_{\mathbf{\alpha Asp}}$ and $\mathbf{18}_{\mathbf{\beta Asp}}$ were single diastereoisomers (Fig. 4e). Moreover, glutamate thioester $\mathbf{5}^{\mathbf{e}}_{\mathbf{Glu}}$ (and Glu-NCA $\mathbf{33}_{\mathbf{Glu}}$ + thiol **2c**) were observed to only yield α-aminoacylation of nucleosides; Glu-isomerisation rapidly quenched γ-activation via pyroglutamate (**36**) formation (Fig. 4e), which completely blocked γ-aminoacylation. The reaction of thiols **3** with activated amino acids (**32**–**33**), which have previously been suggested to be relevant for prebiotic aminoacylation, [44, 45] demonstrate there are multiple high-yielding prebiotic pathways to aminoacyl-thiols **5**, as well as in situ thioester-mediated ribonucleoside-2',3'-aminoacylation at near-neutral pH. Aminoacyl-thiols **5** appear to be remarkably well-suited to the spontaneous and selective aminoacylation of nucleosides in water.



Entry	33	pH	5c (%)	18^{A} (%)
1	L-Ala	3	29	-
2	L-Ala	4	70	-
3	L-Ala	5	86	35
4	L-Ala	6	63	-
5	Gly	5	> 95	21
6	L-Val	5	> 95	5
7	L-Phe	5	> 95	5
8	L-Leu	5	88	22
9	L-Glx	5	84	22
10	L-Glu	5	82	11
11	L-Asp	5	49	12[a]

Figure 4 | Thiol-mediated synthesis of aminoacyl-nucleosides. a) Thiol-mediated aminoacyl-transfer. **b)** ^{1}H NMR spectra to show: (i) alanine-NCA (**33_{Ala}**) in DMSO; (ii) the reaction of L-**33_{Ala}** (20 mM) in MES buffer (200 mM), at room temperature, which yields alanine (**1_{Ala}**, >90%); and (iii) the reaction of **33_{Ala}** (20 mM) and thiol **3c** (100 mM) (R^1=$(CH_2)_2SO_3H$) in MES buffer (200 mM), which yields alanine thioester **5^{c}_{Ala}** %). MES = 3.99 p.p.m. **c)** ^{1}H NMR spectra to show: (i) adenosine (**17A**) at pH 6.5; (ii) the reaction of **17A** (20 mM) and **33_{Ala}** (600 mM) at room temperature and at pH 6.5, where no nucleoside-aminoacylation was observed; and (iii) the reaction of **17A** (20 mM), **33_{Ala}** (600 mM), and thiol **3c** (1 M) at room temperature and at pH 6.5, which furnished alanyl-adenosine **18^{A}_{Ala}** (35%). **d)** NMR yield of aminoacyl-thiol **5c** (R^1=$(CH_2)_2SO_3H$) from NCA (**33**, 20 mM) and thiol **3c** (100 mM), and of aminoacyl-adenosine **18^{A}_{Aaa}** from **17A** (20 mM), **33** (600 mM) and thiol **3c** (1 M) at room temperature after 1 day. Glx = glutamine-γ-nitrile; a = 11/1 α:β. **e)** Anhydride **35_{Asp}** reacted with thiols **3** to yield α-aspartyl-thioester (**5^{c}_{Asp}**) and aspartyl-adenosine (**$18^{A}_{\alpha Asp}$**) with excellent α-selectivity. **5^{c}_{Asp}** isomerised slowly to its β-isomer **$5^{c}_{\beta Asp}$**, which then led to aspartyl-adenosine (**$18^{A}_{\beta Asp}$**) with excellent β-selectivity. ^{1}H NMR spectra to show the reaction of adenosine (**17A**, 20 mM), aspartate-anhydride **35_{Asp}** (600 mM) and thiol **3c** (1 M) at room temperature and at pH 6.5 after: (i) 0.5 days, (ii) 2 days, and (iii) 5 days.

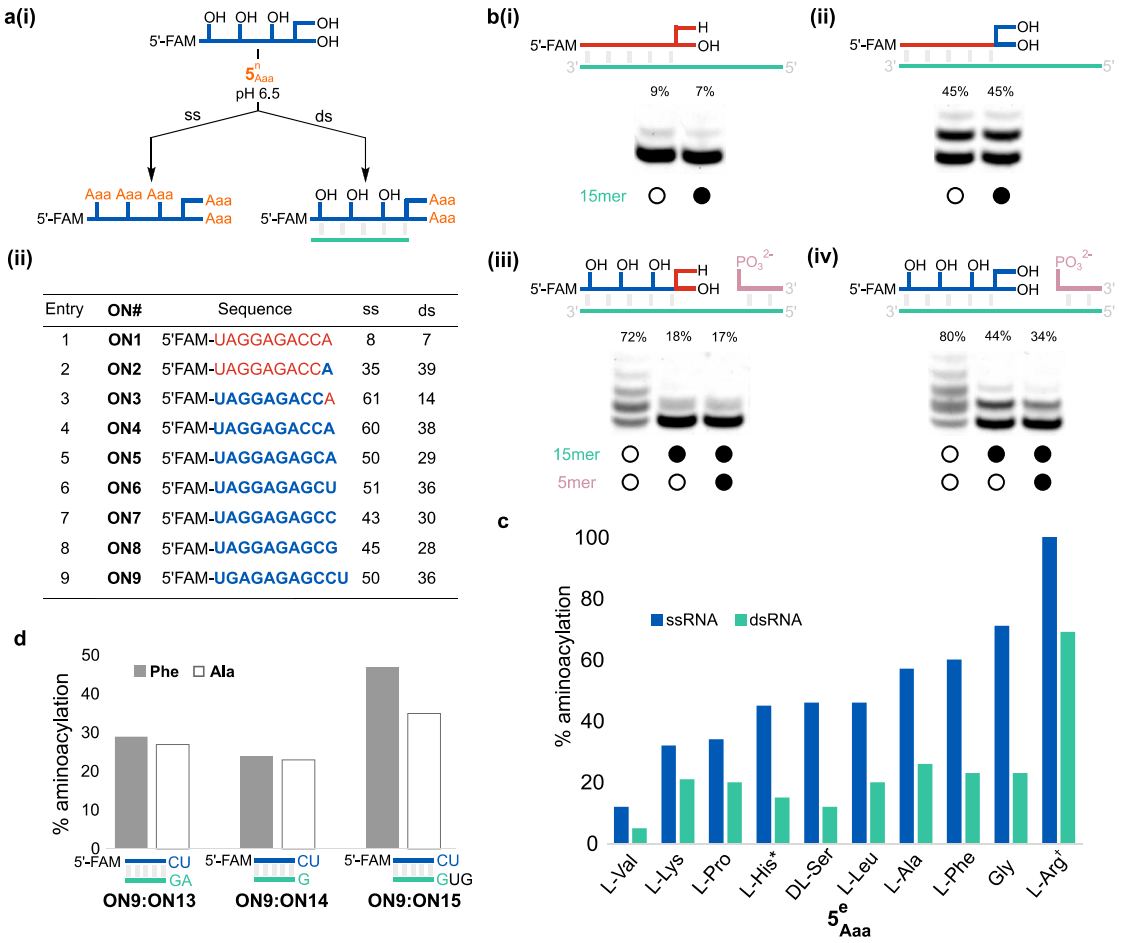
https://doi.org/10.26434/chemrxiv-2024-kbkvg **ORCID:** https://orcid.org/0000-0003-3380-7866

Selective nucleic acid aminoacylation

Control over the site of aminoacylation in oligomers would be essential for optimal positioning of aminoacyl-esters within biomolecules in the same manner that life aminoacylates at the 3'-terminus of tRNAs to mediate protein synthesis. We therefore next turned our attention to nucleic acid aminoacylation, where internal 2'-aminoacylation introduced a further vector to selectivity.[46] We specifically evaluated the aminoacylation of FAM-labelled RNA, DNA, and chimeric nucleic acids (Fig. 5). Minimal DNA-aminoacylation was observed, as a single-gel band, consistent with a single site of aminoacylation at the 3'-terminus of DNA (Fig. 5bi); the reactivity of DNA-oligomers was comparable to 2'-deoxyribonucleoside monomers (**19**; Fig. 2). DNA-oligomer aminoacylation was enhanced 6-fold with a terminal 2',3'-diol, and a second minor bis-aminoacylation product was observed (Fig. 5bii), exhibiting the greatly enhanced reactivity of the 2',3'-diol with respect to the 3'-hydroxyl of DNA. On the other hand, extensive and non-specific single strand (ss) RNA-aminoacylation was observed, with aminoacylation at multiple nucleotide hydroxyls (up to 5) throughout ss-RNA oligonucleotides. Moreover, the multiple-site aminoacylation of ss-RNA was not blocked by a 3'-terminal deoxynucleotide (Fig. 5biii). Indeed, little or no selectivity was observed for aminoacylation at the 2',3'-diol of ss-RNA even in a short nucleic acid (Fig. 5aii Entry 3, 4) – longer RNAs have more inter-nucleotide 2'-hydroxyls that will further diminish this selectivity, which in principle presents a significant problem for selective ss-RNA aminoacylation. However, the formation of Watson-Crick duplex was observed to resurrect selective aminoacylation. The yield of 2',3'-diol aminoacylation was equal in double stranded (ds) RNA (**ON4**) and in ss-oligomer **ON2**, a chimeric nucleic acid with a single RNA-diol at the 3'-terminus, providing a striking demonstration that duplex formation did not significantly inhibit 2',3'-diol aminoacylation of RNA. At the same time, duplex formation inhibits aminoacylation of internal 2'-hydroxyls[44] – ds-RNA aminoacylates less than ss-RNA (Fig. 5aii, Entry 4). The same effect was seen for chimeric nucleic acids with internal ribonucleotides and a terminal 3'-terminal deoxynucleotide. These results indicate that aminoacylation is selective for the 3'-terminal 2',3'-diol in double stranded RNA.

Struck by this simple and inherent mechanism by which RNA-duplex directs selective 2',3'-aminoacylation, we next investigated a more congested nicked duplex with an adjacent downstream 5'-phosphorylated oligomer. The presence of this downstream oligomer did not block the observed diol-aminoacylation (Fig. 5b). Aminoacylation in this nicked duplex demonstrates that selective 2',3'-aminoacylation can occur despite a proximal 5'-phosphate. Conversely, random RNA hydrolysis will furnish oligomers (or nicks) that terminate in a 2'- or 3'-phosphate (rather than a 5'-phosphate), and due to the high pK_a of their 3'- or 2'-hydroxyl moieties, these hydrolysis products are not expected to aminoacylate (Fig. 2c). Interestingly this would offer an innate mechanism for hydrolysed RNA to block adventitious aminoacylation.

Given the remarkable efficiency of selective ds-RNA-aminoacylation, we next investigated whether changing the aminoacyl-thiol side chain would enable a broad scope of chemical aminoacylation of RNAs in water with proteinogenic amino acids. We observed that selective aminoacylation was readily achieved in good-to-excellent yields across a range of non-polar, polar and charged amino thioesters (Fig. 5c). Due to the intramolecular catalytic effect of its side chain, aminoacylation with L-$\mathbf{5^{e}_{Arg}}$ led to quantitative ss-RNA aminoacylation, and 3'-selective aminoacylation of ds-RNA in 64% yield. Unprecedented scope for (non-enzymatic) nucleic acid aminoacylation in water was observed. Furthermore, comparable aminoacylation yields were observed for ds-RNAs terminating in all four canonical nucleotides (A, U, C and G; **ON5-8**) in the presence of the same thioester (e.g., $\mathbf{5^{e}_{Ala}}$, $\mathbf{5^{e}_{Arg}}$, $\mathbf{5^{e}_{Val}}$ or $\mathbf{5^{e}_{Gly}}$) indicating the propensity of thioester-mediated aminoacylation to load a variety of proteinogenic amino acids in a relatively sequence independent manner.



(ii)

Entry	ON#	Sequence	ss	ds
1	ON1	5'FAM-UAGGAGACCA	8	7
2	ON2	5'FAM-UAGGAGACCA	35	39
3	ON3	5'FAM-UAGGAGACCA	61	14
4	ON4	5'FAM-UAGGAGACCA	60	38
5	ON5	5'FAM-UAGGAGAGCA	50	29
6	ON6	5'FAM-UAGGAGAGCU	51	36
7	ON7	5'FAM-UAGGAGAGCC	43	30
8	ON8	5'FAM-UAGGAGAGCG	45	28
9	ON9	5'FAM-UGAGAGAGCCU	50	36

Figure 5 | Selective aminoacylation of nucleic acids. a) (i) Chemoselective 2',3'-aminoacylation of (double strand) ds-RNA oligomers is observed, whereas (single strand) ss-RNA undergoes non-selective aminoacylation. (ii) Aminoacylation (%) for the reaction of ss-/ds-**ON1–9** (0.5 μM) with L-$\mathbf{5^{e}_{Ala}}$(200 mM) in MES buffer (200 mM, pH 6.5) and KCl (500 mM) after 16 hours at room temperature. FAM = Fluorescein; blue = RNA; red = DNA; green = complementary oligomer; pink = downstream oligomer. **b)** PAGE showing the reactions of ss-/ds-**ON1–4** (0.5 μM) with L-$\mathbf{5^{a}_{Ala}}$(600 mM) in MES buffer (200 mM, pH 6.5) and KCl (500 mM) after 16 hours at room temperature: (i) **ON1** ± DNA 15-mer; (ii) **ON2** ± DNA 15-mer; (iii) **ON3** ± RNA 15-mer and ± 5-mer; (iv) **ON4** ± RNA 15-mer and ± 5-mer. Filled circles denote presence of the specified oligonucleotide. 15-mer sequence = 5'-GCAGUUGGUCUCCUA. 5-mer sequence = 5'-phos-ACUGC. **c)** Aminoacylation (%) for the reaction of ss-/ds-**ON4** (0.5 μM) with L-$\mathbf{5^{e}_{Aaa}}$(200 mM) in MES buffer (400 mM, pH 6.5) and KCl (1 M) after 16 hours at room temperature. *MES buffer (200 mM, pH 6.5), KCl (500 mM). †L-$\mathbf{5^{e}_{Arg}}$ (60 mM). **d)** Aminoacylation (%) for the reaction of **ON9** present as a blunt end (**ON9:ON13**, 0.5 μM), underhang (**ON9:ON14**, 0.5 μM) or 'minizyme' (**ON9:ON15**, 0.5 μM) with either L-$\mathbf{5^{e}_{Phe}}$ or L-$\mathbf{5^{e}_{Ala}}$ (100 mM) in MES buffer (200 mM, pH 6.5) and KCl (500 mM) after 16 hours at room temperature. **ON9:ON15** – a previously reported RNA-minizyme, led to a modestly increased yield of aminoacylation.

Ultimately, RNA-aminoacylation and the genetic code are catalytically controlled by a unique family of aminoacyl-tRNA synthetases.[47] These enzymes implement the rules of the genetic code whilst being translated following the same rules; this paradox makes the origins of catalytically controlled aminoacylation particularly intriguing. We therefore finally turned our attention to the thiol-catalysed transfer of other activated amino acids to RNA. As expected, upon incubation of oligonucleotides with NCAs **33** in water no aminoacylation of RNA-oligomers was observed, even with a proximal 5'-phosphate.[38,39] However, the

addition of thiol **3b** alongside NCA $\mathbf{33_{Ala}}$ led to effective oligonucleotide-aminoacylation, via in situ formation of aminoacyl-thiol $\mathbf{5^{b}_{Ala}}$ illustrating the role that thiol cofactors or catalysts could have potentially played in enabling catalytically-controlled aminoacylation during the inception of coded protein synthesis. Ribozyme catalysis was also found to augment thioester-mediated aminoacylation. For example, an RNA-duplex with a mis-paired 3'-U/5'-U and 5'-G overhang, (**ON9**:**ON15**), which is a previously identified catalytic motif from a 5-nucleotide ribozyme ('minizyme'),[41] amplified aminoacylation yields (Figure 5d) and enabled significant aminoacylation (10%) even at extremely low (2 mM) concentrations of aminoacyl-thiol ($\mathbf{5_{Phe}}$). These results suggest that enhancing the diol-selective reactivity of amino thioesters (**5**) by matching with specific RNA catalysts has the potential to boost selectivity for amino acid-sequence pairings, paving the way for a (primitive) coding system. In a preliminary investigation of the diastereoselectivity of thioester-mediated 2',3'-aminoacylation, oligonucleotides terminating in pyrimidines were observed to react with negligible D-Ala selectivity (**ON6-7**: **<** 0.06 d.e.) whilst those ending with a purine led to a slightly increased d.e. (**ON5**, **ON8**: ~0.1 d.e.). This selectivity was not affected by duplex formation and FAM-$(dN)_9N$ chimeric nucleic acids (e.g. **ON2**) underwent aminoacylation with similar diastereoselectivity (< 0.2 d.e.). This selectivity appears to be governed at the monomeric level, with comparable selectivity observed during the aminoacylation of nucleosides; pyrimidines **17C**/**17U** = <0.05 d.e., purines **17A**/**17G** = 0.3 d.e. with $\mathbf{5^{e}_{Ala}}$. Given the transient nature of aminoacyl esters, as only one intermediate of nucleic acid-mediated peptide synthesis, the impact of this selectivity on the overall multi-step scheme cannot yet be known, but it is inevitable that (multiple) subsequent steps will impact diastereoselectivity.[48] Even extant evolved synthetase enzymes may not achieve strictly L-selective aminoacylation,[49,50] despite the presence of D-amino acids and D-peptides in the cell. The encoding of aminoacylation will inevitably affect stereochemical selection; indeed, it is possible that stereochemically-agnostic aminoacylation is important to enable the comprehensive side-chain compatibility necessary for the generation of (diverse) proteins. Therefore, we have not attempted to augment or reverse the innate selectivity observed here, but catalytic aminoacylation paves the way to address amino acid coding and selection in the future.

Conclusion

We have discovered a direct route from prebiotically plausible aminonitriles **2**, via thioesters **5**, to aminoacyl-RNA, setting the stage for chemically controlled nucleic acid-mediated synthesis of peptides. The spontaneous formation of aminoacyl-RNAs is a necessary precondition of coded peptide synthesis but so far both thioesters and aminoacyl ribonucleosides have been extremely challenging to form under prebiotically realistic conditions.[24] Our work shows that aminoacyl thioesters **5**, and thiol catalysis, make RNA-aminoacylation a predisposed chemical process in water. It is of note that this pathway unites the intermediates of non-ribosomal peptide synthesis with ribosomal peptide synthesis (Fig. 1a,b). Thioester mediated RNA-aminoacylation occurs with unprecedented efficiency and sidechain tolerance in water at neutral pH. The next stage of this journey will be to understand how aminoacyl-RNAs were chemically exploited under the conditions of their formation in a primordial context. Recent work has shown that (flexizyme) aminoacylation of phosphorimidazolide-activated oligonucleotide nicked-loops can lead to chimeric loop-closure, which can yield functional ribozymes,[51] and (dimethyl)aminoacylation of RNA has been shown to protect RNA strands against hydrolysis and enzymatic degradation,[36] which suggests that aminoacylation could be extremely beneficial for early nucleosides in multiple ways that have not yet been sufficiently explored. Of course, the ideal next step would be to understand how oligonucleotide-aminoacylation can enable nucleic acids to control and impact sequence specific peptide synthesis with (ideally) a breadth of amino acid sidechain functional groups exploited in proteinogenic peptides, and to elucidate which stereochemical interactions within structured (non-duplex) nucleic acids direct, augment or constrain thioester-mediated aminoacylation.

Contributions

J.S., B.T., D.W., and M.W.P. conceptualisation, methodology and analysis. J.S., B.T., D.W., Y.Y. and M.W.P. investigation. J.S., B.T. and D.W. contributed equally. D.W. and M.W.P. acquisition of funding. M.W.P. supervision. J.S., B.T., D.W. and M.W.P wrote the paper, and all authors approved the final submission.

Acknowledgements

The authors thank the Engineering and Physical Sciences Research Council (EP/X011755/1 to M.W.P.), the Simons Foundation (1154101 to M.W.P.) and the Royal Society (URF\R1\231450 to D.W.) for financial support, and A. E. Aliev, J. Attwater, K. Karu (UCL, Chemistry) and F. Werner (UCL, Structural & Molecular Biology) for technical support, access to equipment and facilities, and for helpful discussions.

References

[1] Spiegelman, S., Haruna, I., Holland, I.B., Beaudreau, G. & Mills, D. The Synthesis of a Self-propagating and Infectious Nucleic Acid with a Purified Enzyme. *Proc. Natl Acad. Sci. U.S.A.* **54**, 919–927 (1965).
[2] Bartel, D. P. & Szostak, J. W. Isolation of new ribozymes from a large pool of random sequences. *Science* **261**, 1411–1418 (1993).
[3] Voytek, S. B. & Joyce, G. F. Emergence of a fast-reacting ribozyme that is capable of undergoing continuous evolution *Proc. Natl Acad. Sci. U.S.A.* **104** 15288–15293 (2007).
[4] Ichihashi, N., Usui, K., Kazuta, Y. Sunami, T. Matsuura, T. & Yomo, T. Darwinian evolution in a translation-coupled RNA replication system within a cell-like compartment. *Nat. Commun.* **4**, 2494 (2013).
[5] Lee, D. H., Granja, J. R., Martinez, J. A., Severin, K. & Ghadiri, M. R. A self-replicating peptide. *Nature* **382**, 525–528 (1996).
[6] Rout, S.K., *et al.* A prebiotic template-directed peptide synthesis based on amyloids. *Nat. Commun.* **9**, 234 (2018).
[7] Breaker, R. R. & Joyce, G. F. The Expanding View of RNA and DNA Function, *Chem. Biol.* **21**, 1059–1065 (2014).
[8] Rodnina, M. V. Translation in Prokaryotes *Cold Spring Harb Perspect Biol.* **10** a032664 (2018).
[9] Fischer, E. & Speier, A. Darsellung der Ester *Ber. Dtsch. Chem. Ges.* **28**, 3252 (1895).
[10] Janssen, M. J. Thiolo, thiono and dithio acids and esters. In S. Patai (Ed.), *Carboxylic Acids and Esters* pp. 705–764 John Wiley & Sons (1969).
[11] Iimura, S. Manabe, K. & Kobayashi, S. Direct thioesterification from carboxylic acids and thiols catalyzed by a Brønsted acid. *Chem. Commun.* 94–95 (2002).
[12] Benner, S. A., Kim, H.-J., & Carrigan, M. A. Asphalt, Water, and the Prebiotic Synthesis of Ribose, Ribonucleosides, and RNA. *Acc. Chem. Res.* **45**, 12, 2025–2034 (2012).
[13] Weber, A. L. Aqueous synthesis of peptide thioesters from amino acids and a thiol using 1,1'-carbonyldiimidazole *Orig. Life Evol. Biosph.* **35** 421–427 (2005).
[14] Amblard, M., Fehrentz, J.-A., Martinez, J. & Subra, G. Methods and protocols of modern solid phase peptide synthesis *Mol. Biotechnol.* **33** 239-354 (2006).
[15] Hadjichristidis, N., Iatrous, H., Pitsikalis, M. & Sakellariou, G. Synthesis of Well-Defined Polypeptide-Based Materials via the Ring-Opening Polymerization of α-Amino Acid N-Carboxyanhydrides *Chem. Rev.,* **109**, 5528–5578 (2009).
[16] Shannon, C. E. A Mathematical Theory of Communication *Bell Syst. Tech. J.* **27** 379-423 (1948).
[17] Foden, C. S., Islam, S., Fernández-García, C., Maugeri, L., Sheppard, T. D., & Powner, M. W. Prebiotic synthesis of cysteine peptides that catalyze peptide ligation in neutral water. *Science* **370**, 865–869 (2020).
[18] Singh, J., Whitaker, D., Thoma, B., Islam, S., Foden, C. S., Aliev, A. E., Sheppard, T. D., & Powner, M. W. Prebiotic Catalytic Peptide Ligation Yields Proteinogenic Peptides by Intramolecular Amide Catalyzed

Hydrolysis Facilitating Regioselective Lysine Ligation in Neutral Water. *J. Am. Chem. Soc.*, **144**, 10151–10155 (2022).
[19] Barat, A. & Powner, M. W. Spontaneous Peptide Ligation Mediated by Cysteamine. *JACS Au* **4**, 1752–1757 (2024).
[20] Fairchild, J., Islam, S., Singh, J., Bučar, D. K. & Powner, M. W. Prebiotically plausible chemoselective pantetheine synthesis in water. *Science* **383**, 911–918 (2024).
[21] Chaturvedi, R. K., MacMahon, A. E. & Schmir, G. L. The Hydrolysis of Thioimidate Esters. Tetrahedral Intermediates and General Acid Catalysis. *J. Am. Chem. Soc.* **89**, 6984–6993 (1967).
[22] Chaturvedi, R. K. & Schmir, G. L. Hydrolysis of thioimidate esters. II. Evidence for the formation of three species of the tetrahedral intermediate. *J. Am. Chem. Soc.* **91**, 737–746 (1969).
[23] Weber, A. L. & Lacey, J. C. Aminoacyl transfer from an adenylate anhydride to polyribonucleotides. *J. Mol. Evol.* 6, 309–320 (1975).
[24] Weber, A. L., & Orgel, L. E. Amino acid activation with adenosine 5'-phosphorimidazolide. *J. Mol. Evol.* 11, 9–16 (1978).
[25] Radakovic, A. et al. Structure-guided aminoacylation and assembly of chimeric RNAs. bioRxiv 2024.03.02.583109; doi: https://doi.org/10.1101/2024.03.02.583109.
[26] Van den Berg, L. The effect of addition of sodium and potassium chloride to the reciprocal system: KH_2PO_4-Na_2HPO_4-H_2O on pH and composition during freezing. *Arch. Biochem. Biophys.* **84**, 305–315 (1959).
[27] Thoma, B., & Powner, M. W. Selective Synthesis of Lysine Peptides and the Prebiotically Plausible Synthesis of Catalytically Active Diaminopropionic Acid Peptide Nitriles in Water. *J. Am. Chem. Soc.* **145**, 3121–3130 (2023).
[28] Kvenvolden, K. A., Lawless, J. G. & Ponnamperuma, C. Nonprotein Amino Acids in the Murchison Meteorite. *Proc. Natl Acad. Sci. U.S.A.* **68**, 486–490 (1971).
[29] Gatto, G. J., Boyne, M. T., Kelleher, N. L. & Walsh, C. T. Biosynthesis of Pipecolic Acid by RapL, a Lysine Cyclodeaminase Encoded in the Rapamycin Gene Cluster. *J. Am. Chem. Soc.* **128**, 3838–3847 (2006).
[30] Frenkel-Pinter, M., Samanta, M., Ashkenasy, G., & Leman, L. J. Prebiotic Peptides: Molecular Hubs in the Origin of Life. *Chem. Rev.* **120**, 4707–4765 (2020).
[31] Canavelli, P., Islam, S., & Powner, M. W. Peptide ligation by chemoselective aminonitrile coupling in water. *Nature* **571**, 546–549 (2019).
[32] Porcher, S., Meyyappan, M., & Pitsch, S. Spontaneous Aminoacylation of a RNA Sequence Containing a 3'-Terminal 2'-Thioadenosine. *Helv. Chim. Act.*, **88**, 2897–2909 (2005).
[33] Ninomiya, K., Minohata, T., Nishimura, M. & Sisido, M. In Situ Chemical Aminoacylation with Amino Acid Thioesters Linked to a Peptide Nucleic Acid. *J. Am. Chem. Soc.* **126**, 15984–15989 (2004).
[34] Li, N. & Huang, F. Ribozyme-Catalyzed Aminoacylation from CoA Thioesters. *Biochemistry* **44**, 4582–4590 (2005).
[35] Gottikh, B. P., Krayevsky, A. A., Tarussova, N. B., Purygin, P. P. & Tsilevich, T. L. The general synthetic route to amino acid esters of nucleotides and nucleoside-5-triphosphates and some properties of these compounds. *Tetrahedron* **26**, 4419–4433 (1970).
[36] Fang, L., Xiao, L., Jun, Y. W., Onishi, Y. & Kool, E. T. Reversible 2'-OH acylation enhances RNA stability. *Nat. Chem.* **15** 1296–1305 (2023).
[37] Bowler, F. R. et al Prebiotically plausible oligoribonucleotide ligation facilitated by chemoselective acetylation. *Nat. Chem.* **5** 383–389 (2013).
[38] Tamura, K. & Schimmel, P. Chiral-Selective Aminoacylation of an RNA Minihelix. *Science* **305**, 1253–1253 (2004).
[39] Wu, L.-F., Su, M., Liu, Z., Bjork, S. J. & Sutherland, J. D. Interstrand aminoacyl transfer in a tRNA acceptor stem-overhang mimic. *J. Am. Chem. Soc.* **143** 11836–11842 (2021).
[40] Peacock, J. R. *et al.* Amino acid–dependent stability of the acyl linkage in aminoacyl-tRNA. *RNA* **20**, 758–764 (2014).
[41] Turk, R. M., Illangasekare, M. & Yarus, M. Catalyzed and Spontaneous Reactions on Ribozyme Ribose. *J. Am. Chem. Soc.* **133**, 6044–6050 (2011).
[42] Biron, J.-P., Parkes, A. L., Pascal, R. & Sutherland, J. D. Expeditious, potentially primordial, aminoacylation of nucleotides. *Angew. Chem. Int. Ed.* **44**, 6731–6734 (2005).

[43] Leman, L. J., Orgel, L. E. & Ghadiri, M. R. Amino Acid Dependent Formation of Phosphate Anhydrides in Water Mediated by Carbonyl Sulfide. *J. Am. Chem. Soc.* **128**, 1, 20–21 (2006)
[44] Biron, J.-P., Parkes, A. L., Pascal, R. & Sutherland, J. D. Expeditious, potentially primordial, aminoacylation of nucleotides. *Angew. Chem. Int. Ed.* **44**, 6731–6734 (2005).
[45] Leman, L. J., Orgel, L. E. & Ghadiri, M. R. Amino Acid Dependent Formation of Phosphate Anhydrides in Water Mediated by Carbonyl Sulfide. *J. Am. Chem. Soc.* **128**, 1, 20–21 (2006)
[46] Merino, E.J., Wilkinson, K.A., Coughlan, J.L. & Weeks, K. M. RNA structure analysis at single nucleotide resolution by selective 2'-hydroxyl acylation and primer extension (SHAPE). *J. Am. Chem. Soc.* **127** 4223–4231 (2005).
[47] Gomez, M. A. R. & Ibba, M. Aminoacyl-tRNA synthetases *RNA* **26**, 910-936 (2020).
[48] Deng, M., Yu, J. & Blackmond, D.G. Symmetry breaking and chiral amplification in prebiotic ligation reactions. *Nature* **626**, 1019–1024 (2024).
[49] Calendar, R. & Berg, P. D-Tyrosyl RNA: Formation, hydrolysis and utilization for protein synthesis. *J. Mol. Biol.* **26**, 39–54 (1967).
[50] Pawar, K. I. et al. Role of D-aminoacyl-tRNA deacylase beyond chiral proofreading as a cellular defense against glycine mischarging by AlaRS. *eLife* **6**:e24001 (2017).
[51] Radakovic, A. et al Nonenzymatic assembly of active chimeric ribozymes from aminoacylated RNA oligonucleotides. *Proc. Natl Acad. Sci. U.S.A.* **119**, e2116840119 (2022).