nature chemistry

ARTICLES

PUBLISHED ONLINE: 16 JANUARY 2017 | DOI: 10.1038/NCHEM.2703

Prebiotic selection and assembly of proteinogenic amino acids and natural nucleotides from complex mixtures

Saidul Islam, Dejan-Krešimir Bučar and Matthew W. Powner*

A central problem for the prebiotic synthesis of biological amino acids and nucleotides is to avoid the concomitant synthesis of undesired or irrelevant by-products. Additionally, multistep pathways require mechanisms that enable the sequential addition of reactants and purification of intermediates that are consistent with reasonable geochemical scenarios. Here, we show that 2-aminothiazole reacts selectively with two- and three-carbon sugars (glycolaldehyde and glyceraldehyde, respectively), which results in their accumulation and purification as stable crystalline aminals. This permits ribonucleotide synthesis, even from complex sugar mixtures. Remarkably, aminal formation also overcomes the thermodynamically favoured isomerization of glyceraldehyde into dihydroxyacetone because only the aminal of glyceraldehyde separates from the equilibrating mixture. Finally, we show that aminal formation provides a novel pathway to amino acids that avoids the synthesis of the non-proteinogenic α,α-disubstituted analogues. The common physicochemical mechanism that controls the proteinogenic amino acid and ribonucleotide assembly from prebiotic mixtures suggests that these essential classes of metabolite had a unified chemical origin.

The conservation of the genetic code, amino acids and nucleotides in biology suggests a single origin of life on Earth[1–13]. Proteins are built from a highly restricted set of about 20 amino acids according to a universal triplet code of four ribonucleotides. Therefore, it is essential to learn how this specific small constellation of molecules became irrevocably linked at the advent of life[1–21]. In contrast to the narrow distribution of universal metabolites observed in biology, typical prebiotic reactions are notorious for their complex product distributions. Accordingly, it has been recognized that "the chief obstacle to understanding the origin of RNA-based life is identifying a plausible mechanism for overcoming the clutter wrought by prebiotic chemistry"[4]. For example, the most-efficient and specific proposed prebiotic pathway to the pyrimidine ribonucleotides requires the synthesis of the key intermediate pentose aminooxazoline (**1**) (Fig. 1a)[2–6,22–24]. However, the plausibility of this proposed prebiotic synthesis of **1** has been questioned because it is contingent on the strictly controlled sequential delivery of pure glycolaldehyde (**2a**) to cyanamide (**3**) to yield 2-aminooxazole (**4**) and then of pure glyceraldehyde (**2b**) to **4** to yield the desired product (**1**) (Fig. 1a). This is a serious problem because both these reactions lack the intrinsic selectivity required to yield exclusively their respective products (**4** and **1**) from mixtures of **2a** and **2b**. The problem becomes increasingly worse in the presence of other sugars. Without a separate and sequential delivery of **2a** and **2b**, a complex mixture of undesirable by-products results en route to the canonical nucleotides[4–7,19,23,25–27]. To exacerbate matters further, aldose glyceraldehyde (**2b**) is thermodynamically unstable with respect to its ketose isomer dihydroxyacetone (**2q**). Triose equilibration is catalysed by specific base, general acid–base and metal-ion (such as Zn^{2+}) catalysis[13,27], and equilibration yields more of **2q** than the aldose isomer glyceraldehyde (**2b**) (**2q**/**2b** = 8.5:1; 25 °C, pH 7 (Supplementary Figs 1–4)). Under the conditions required for the formation of **4** (ref. 3), in aqueous solution **2b** equilibrates very rapidly (<0.5 h) with its ketose isomer dihydroxyacetone (**2q**) (11% **2b**, 89% **2q** (Supplementary Fig. 2)). If this equilibration occurs prior to the reaction of triose sugars with **4**, the predominant product is the non-natural branched apiose **5** rather than the desired product **1** (69% **5**, 31% **1** (Supplementary Fig. 17)). Moreover, the dominant ketose isomer problem is even more acute in C_2 and C_3 sugar mixtures. Incubation of **2a** and **2b** (1:1) in phosphate buffer yields **2a**/**2b**/**2q** (1:0.11:0.89). Reaction of this mixture with **4** (1 equiv.) returns **1** (11%) as a very minor component of a complex product mixture that consists predominantly of the undesirable apiose aminooxazoline (**5**, 12%)) and tetrose aminooxazoline (**6**, 70% (Supplementary Fig. 19)). These observations demonstrate the importance of identifying a process that could sequentially deliver **2a** and **2b** in a pure form to support the recently proposed prebiotic synthesis of activated pyrimidine ribonucleotides[3,24].

Results

Selective extraction of glycolaldehyde (2a) and glyceraldehyde (2b) from complex sugar mixtures. All the known prebiotic syntheses of sugars, including the formose reaction and the Kiliani–Fischer process, lead to mixtures of glycolaldehyde (**2a**) and glyceraldehyde (**2b**) together with other C_4, C_5 and C_6 sugars[1–7,13,15–21]. A separate synthesis of pure **2a** and **2b** is ostensibly implausible, and therefore a means of physical separation is better suited to promote the desired selective and sequential delivery of these simple sugars[6,14,28]. Our attention was first drawn to the possible role of 2-aminothiazole (**7**) in mediating the sequestration, separation and accumulation of **2a** and **2b** when we observed the precipitation of **2a** and homochiral D-glyceraldehyde (D-**2b**) from water as aminals **8a** and D-**8b** (ref. 14). As discussed below, **7** is expected to be abundant in a cyanosulfidic environment[29]. We now report that **7**-induced crystallization resolves several long-standing problems inherent to ribonucleotide selection. We demonstrate the time-resolved sequestration of **2a** and **2b** from complex sugar mixtures as separate stable crystal reservoirs in the form of the corresponding aminals **8a** and **8b**. Aminals **8a** and **8b** allow for sequential

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

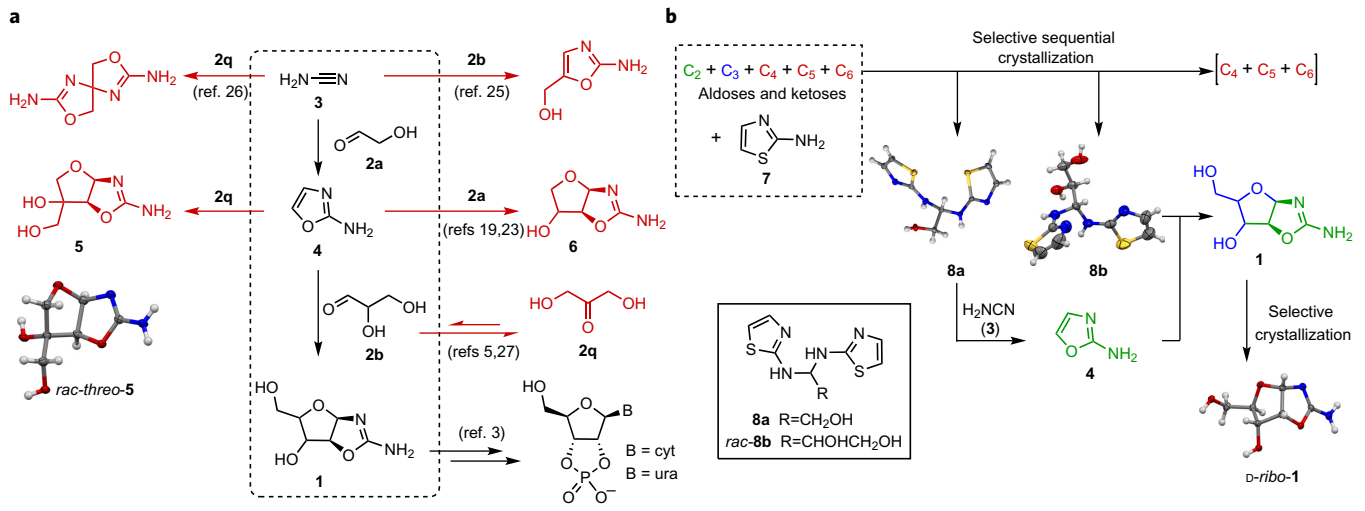


Figure 1 | Prebiotic ribonucleotide synthesis. a, Stepwise nucleotide synthesis from pure reagents (previous work). Pyrimidine ribonucleotide synthesis (black arrows) requires the stepwise addition of glycolaldehyde (**2a**) and glyceraldehyde (**2b**) to avoid deleterious by-product formation (red arrows). Bottom left: single-crystal X-ray structure of *rac*-apiose *threo*-furanosyl aminooxazoline (*rac*-*threo*-**5**). cyt, cytosine; ura, uracil. **b**, Nucleotide synthesis from complex sugar mixtures (this work). Crystallization-controlled synthesis of ribose aminooxazoline (*ribo*-**1**) from a sugar mixture that contains an equilibrating mixture of **2a**, **2b**, dihydroxyacetone (**2q**), L-erythrulose (**2w**), L-arabinose, D-lyxose, D-ribose, D-xylose, D-glucose, D-galactose, D-mannose, D-fructose and D-sorbose mediated by sequential and separate 2-aminothiazole (**7**)-induced crystallization of glycolaldehyde aminal **8a** and *rac*-glyceraldehyde aminal **8b**. Centre: single-crystal X-ray structures acquired directly from the sequential precipitates of C_2-aminal **8a** and racemic C_3-aminal *rac*-**8b** from a complex sugar mixture, and homochiral D-*α*-*ribo*-furanosyl aminooxazoline (D-*ribo*-**1**) isolated from bulk conglomerate crystals obtained from the reaction of C_3-aminal (*rac*-**8b**) with 2-aminooxazole (**4**), synthesized from the exposure of C_2-aminal **8a** to cyanamide (**3**).

formation of the key intermediates 2-aminooxazole (**4**) and pentose aminooxazoline (**1**) from a highly complex mixture of C_2, C_3, C_4, C_5 and C_6 sugars (Fig. 1b). Furthermore, the unprecedented sequestration of racemic aminal *rac*-**8b** removes **2b** from the aldose/ketose equilibrium, which overturns the inherent thermodynamic bias towards the ketose isomer dihydroxyacetone (**2q**)[5,13,27].

For **7** to play the role of 'chemical chaperone' for ribonucleotide synthesis in a chemically complex environment it must have a plausible prebiotic synthesis of its own. Therefore, we examined the synthesis of **7** in complex sugar mixtures. **7** was synthesized in

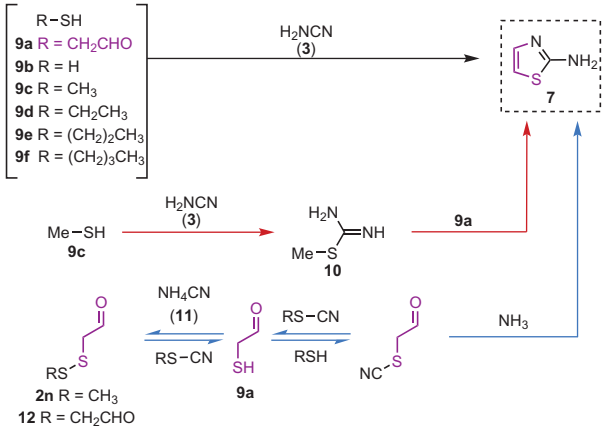


Figure 2 | Multiple pathways to 2-aminothiazole (7) in aqueous cyanosulfidic solution. Facile and selective synthesis of **7** at neutral pH from **9a** (magenta) and cyanamide (**3**) (black arrow) and selective assembly of **7** is observed in competition with prebiotically plausible thiols (**9b**–**9f**). **9a** reacts efficiently with methylisothiourea (**10**) to furnish **7** (red arrows), which demonstrates a reversible thiol-exchange during the synthesis of **7**. Finally, disulfides **2n** and **12** are observed to react with ammonium cyanide (**11**) to afford **7** (blue arrows) under Strecker conditions. These pathways collectively demonstrate the facile and selective assembly of **7** under mild, aqueous cyanosulfidic conditions.

near-quantitative yield (>95%) from the cysteine precursor β-mercaptoacetaldehyde (**9a**) and cyanamide (**3**) (Fig. 2). A very high conversion was observed even in stoichiometric competition with 26 other aldehyde, ketone and sugar species present in the mixture (including **2a**–**2w** and C_2, C_3, C_4, C_5 and C_6 sugars (Supplementary Figs 21 and 22)). Furthermore, because the addition of thiols to **3** is reversible, we also explored the synthesis of **7** from **9a** in the presence of five other stoichiometric sulfides/thiols (**9b**–**9f** (Fig. 2)) and observed a highly selective reaction; **7** was obtained in 86% yield. Moreover, isothiourea (**10**) undergoes an efficient thiol exchange with **9a** to furnish **7** in 72% yield under mild aqueous conditions (pH 7, room temperature (r.t.), 16 hours (Fig. 2)). Finally, we were drawn to investigate the synthesis of **7** from prebiotically plausible disulfides[29]. It is known that cyanide efficiently reduces disulfides[30]. Thus, the reaction of disulfide **2n** with ammonium cyanide (**11**) furnishes **7** (14% yield), whereas the reaction of disulfide **2n** with cyanide and **3** gave **7** in up to 75% yield (Supplementary Fig. 23). The facile and predisposed assembly of **7** from a complex mixture of aldehydes, ketones, sulfides and thiols in water suggests that it is a highly apposite reagent in the search for prebiotic selectivity (Fig. 2).

With **7** established as a plausible and highly robust component of prebiotic cyanosulfidic chemical environments, we began to explore C_2 and C_3 sugar separation, which is required for selective ribonucleotide synthesis (Fig. 1a)[3]. We incubated **2a** (0.25 M), **2b** (0.25 M) and **7** (0.5 M) in phosphate buffer (pH 7, 25 °C) and monitored the reaction over 24 hours. Initially (0–6 hours), we observed a mixture of aminals **8a** and **8b**. However, surprisingly, after 12 hours a complete resolution of the C_2 and C_3 sugars was observed: **8a** was the only aminal isolated (77%, 24 hours (Fig. 3a and Supplementary Table 3)) and all the C_3 sugars remained in solution. To distinguish between kinetic and thermodynamic effects in the selective formation of the C_2-aminal **8a**, we incubated **2a** with the C_3-aminal **8b** (1:1, 25 °C, three days, pH 7) in phosphate buffer and observed an accumulation of the C_2-aminal **8a** in 56% yield. Conversely, incubation of **2b** with the C_2-aminal **8a** (1:1, 25 °C, three days, pH 7) in phosphate buffer did not form the C_3-aminal **8b**—only the triose

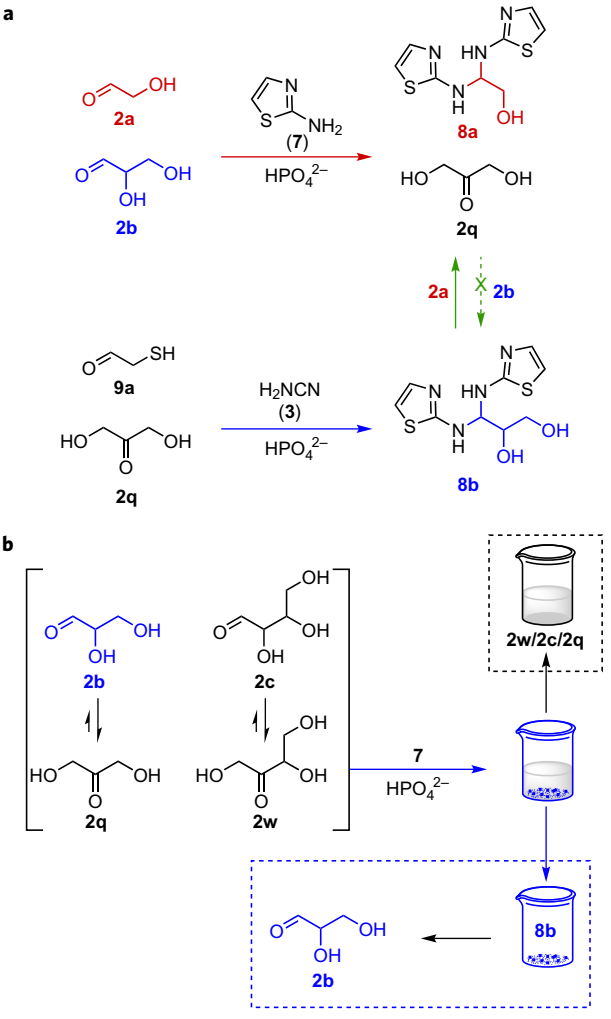


Figure 3 | Selective C_2-aminal 8 sequestration from a mixture of C_2 and C_3 sugars. a, Synthesis of glycolaldehyde aminal **8a** from a mixture of **2a** (1 equiv.), **2b** (1 equiv.) and **7** (2 equiv.) (red arrow). Selective multicomponent glyceraldehyde aminal **8b** synthesis from dihydroxyacetone (**2q**), β-mecaptoacetaldehyde (**9a**) and cyanamide (**3**) (blue arrow), and selective conversion of **8b** into **8a** on incubation with **2a** (1 equiv.) in pH 7 phosphate buffer (green arrows) demonstrate the thermodynamically controlled selection of **8a** and **2q** from mixtures of C_2 and C_3 sugars. **b**, C_3-specific aldose sequestration by the crystallization of aminal **8b** from an equilibrating C_3- and C_4-sugar mixture (**2b**, **2c**, **2q** and **2w**) on reaction with **7** in phosphate buffer (blue arrows) provides a facile and highly selective separation of C_3 and C_4 sugars.

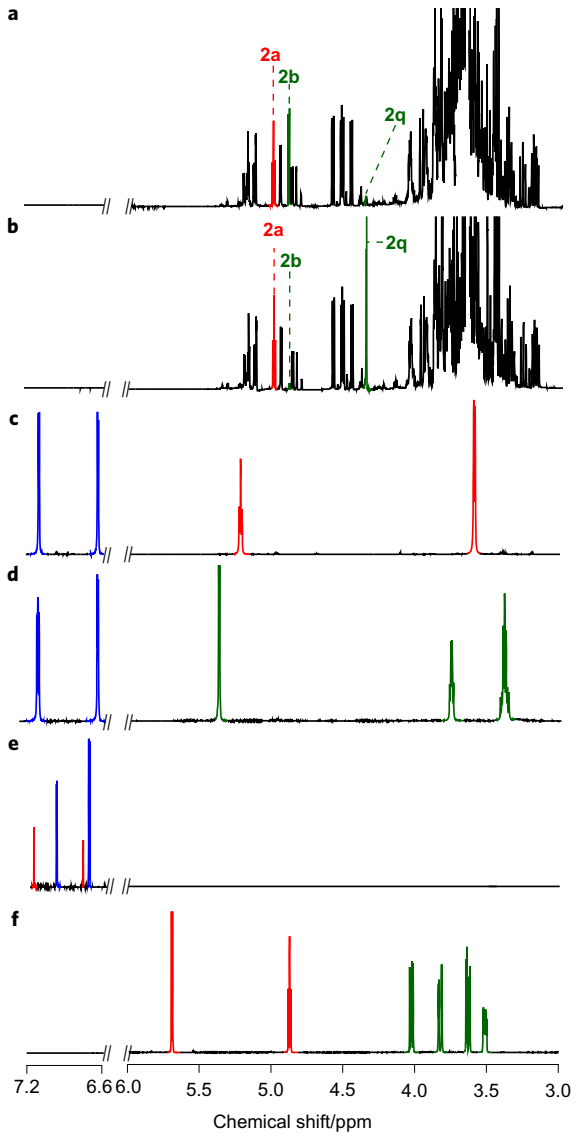
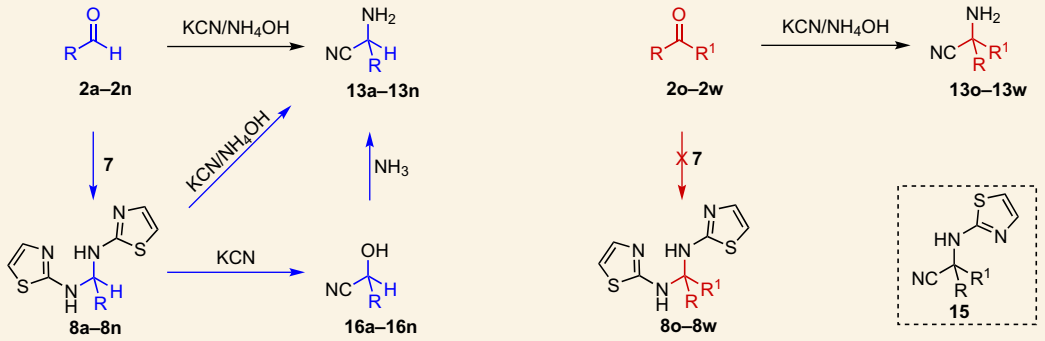


Figure 4 | Convergent crystallization-controlled synthesis of pure *ribo*-1 from a complex mixture of sugars. a–f, ^{1}H NMR spectra (600 MHz) of the reaction of one equivalent each of **2a** (red), *rac*-**2b** (green), L-erythrulose, L-arabinose, D-lyxose, D-ribose, D-xylose, D-glucose, D-galactose, D-mannose, D-fructose and D-sorbose incubated in phosphate buffer (1 M, pH 7, 25 °C) with **7** (6 equiv.). The initial sugar mixture is shown in **a**, and the equilibrated sugar mixture after three days (**b**) shows the conversion of **2b** (green, 4.90 ppm, doublet, J = 5.4 Hz) into **2q** (green, 4.30 ppm, singlet). The subsequent four spectra are from the crystalline precipitate **8a** that accumulated 0–2 h after the addition of **7** (blue) (**c**), the crystalline precipitate **8b** that accumulated 48–480 h after the addition of **7** (**d**), the quantitative transformation of aminal **8a** and cyanamide (**3**) to give a 1:2 mixture of **4** (red) and **7** (blue) (**e**) and the crystalline conglomerate D-*ribo*-**1** isolated on exposure of precipitate **8b** to a 1:2 mixture of **4** and **7** (**f**).

equilibration (**2b** ↔ **2q**) was observed, even with a twofold excess of **2b**. Therefore, clearly the equilibration of **2b** with its more-stable ketose isomer dihydroxyacetone (**2q**) drives the selective accumulation of the C_2-aminal **8a**. However, remarkably, incubation of **2q** with either **7** or **9a** plus **3** (to effect an *in situ* synthesis of **7**) delivered aminal *rac*-**8b** in up to 87% yield (Fig. 3a and Supplementary Table 6), which thereby deposited a reservoir of the crystalline aldose isomer **2b** and completely overturned the thermodynamic preference for the C_3 ketose isomer observed in aqueous solution. In contrast, phosphate-induced isomerization of tetrose **2c** to erythrulose (**2w**) completely inhibited precipitation of the C_4-aminal **8c**. Consequently, only C_3-aminal **8b** (56%) was observed to precipitate from a stoichiometric mixture of **2q** and **2w** after incubation (23 days) with **7** (6 equiv.) in phosphate buffer (1 M, 25 °C, pH 7), which resulted in a robust resolution of C_3 and C_4 sugars through the unique crystallization of aminal **8b** (Fig. 3b and Supplementary Table 8).

Selective synthesis of pentose aminooxazoline (1) from a complex sugar mixture. We recognized that incubation of glycolaldehyde (**2a**) and glyceraldehyde (**2b**) in phosphate buffer prior to aminal precipitation would result in a very large differential accumulation of C_2-aminal **8a** and C_3-aminal **8b**, which would provide the direct physical mechanism to separate **2a** and **2b** spatially, which is so crucial to the synthesis of pentose aminooxazoline (**1**). To test our hypothesis, a model prebiotic sugar mixture containing C_2, C_3, C_4, C_5 and C_6 sugars (Fig. 4a) was incubated in phosphate buffer (1 M, pH 7, r.t., three days (Fig. 4b)) to allow triose equilibration, and then

Table 1 | Prebiotic aminonitrile synthesis.



2	R	R^1	8 (%)	13 (%)	Amino acid
a	CH_2OH	H	95	>95	Ser
b	$CH(OH)CH_2OH$ (*rac*)	H	90*	–	–
c	*R,R*-$CH(OH)CH(OH)CH_2OH$	H	32	–	–
d	H	H	96	>95	Gly
e	CH_3	H	99	>95	Ala
f	CH_2CH_3	H	85	>95	-
g	$CH(OH)CH_3$	H	70	91	Thr
h	$CH(CH_3)CH_3$	H	94	>95	Val
i	$CH(CH_3)CH_2CH_3$	H	92	72	Ile
j	$CH_2CH(CH_3)CH_3$	H	89	70	Leu
k	CH_2CH_2CN	H	89	90†	Glu/Gln
l	$CH_2CH_2CH_2NHAc$	H	90	>95	(Pro/Arg)‡
m	$CH_2CH_2SCH_3$	H	75	70	Met
n	CH_2SSCH_3	H	79	–	Cys§
o	CH_3	CH_3	0	83	–
p	CH_2OH	CH_3	0	>95	–
q	CH_2OH	CH_2OH	0	>95	–
r	CH_2CH_3	CH_3	0	88	–
s	$CH_2CH_2CH_3$	CH_3	0	78	–
t	$CH(CH_3)CH_3$	CH_3	0	87	–
u	$CH(CH_3)CH_2CH_3$	CH_3	0	84	–
v	$CH_2CH_2CH_2CH_2CH_3$	CH_3	0	90	–
w	$CHOHCH_2OH$	CH_2OH	0	–	–

Non-selective synthesis of aminonitriles 13a–13w from aldehydes and ketones by the traditional Strecker pathway (black arrows; previous work). Aldehyde aminal 8 selective synthesis of aminonitriles 13a and 13d–13m (blue arrows; this work); aminonitriles 15 were not observed. Isolated yields of aminals 8a–8w from the reaction of aldehydes/ketones 2a–2w (500 mM) and 2-aminothiazole (7 (1 M, pH 7, 24 h)) are reported. Aminonitriles 13a–13n were obtained by equilibration of aminals 8a–8n with potassium cyanide (1.5 equiv.) and ammonium hydroxide (5 equiv.) at pH 9.2; NMR yields are reported. Aminonitriles 13o–13w were acquired by the equilibration of ketones 2o–2w (1 equiv.), potassium cyanide (1.5 equiv.) and ammonium hydroxide (5 equiv.) at pH 9.2; NMR yields are reported. *Homochiral D-8b has previously been crystallized from an ethanol/water mixture[14]. †A mixture of α-aminonitrile 13k (62%) and α-aminoamide (28%) was obtained. ‡8l contains the four contiguous carbon/nitrogen atoms of the proline/arginine precursors and can be elaborated to the canonical amino acids after amide hydrolysis[13]. §8n is converted into the Strecker precursor of cysteine 9a on disulfide reduction.

2-aminothiazole (7 (6 equiv.)) was added. Over the course of two hours the selective crystallization of pure aminal **8a** (>95%) was observed (Fig. 4c). Further incubation of the supernatant, after the isolation of pure glycolaldehyde aminal **8a**, resulted in the crystallization of pure glyceraldehyde aminal **8b** (61%, 2–17 days (Fig. 4d)). No further crystallization was observed. The onset of the crystallization of **8b** is slow (9% after 24 hours), and the large observed time resolution for aminal precipitation has significant implications for the spatial resolution, accumulation, concentration and subsequent reactions of aminals **8a** and **8b**. Indeed, exposure of aminal **8a**—sequestered from a mixture of 12 homologous sugars—to cyanamide (**3** (0.65 M)) gave a quantitative yield of 2-aminoxazole (**4** (Fig. 4e)). Incubation of **4** with aminal **8b**, which crystallized after **8a** from the same complex sugar solution, led to the selective formation of pentose aminooxazoline (**1**), followed by the direct crystallization of pure *ribo*-**1** (Fig. 4f).

Having demonstrated the crystallization-controlled synthesis of *ribo*-**1** from the simplest C_2 and C_3 components of a complex sugar mixture, we next turned our attention to the residual (C_4, C_5 and C_6) sugars. It has been reported previously that, because of their rapid rate of reaction, the simplest sugars, **2a** and **2b**, thwarted previous attempts to crystallize *ribo*-**1** by the action of **3** on complex sugar mixtures[22]. However, our sequential **7**-induced crystallization from an aqueous solution of C_2, C_3, C_4, C_5 and C_6 sugars exploits the reactivity of **2a** and **2b** to facilitate their sequestration from the solution, which results in a residual sugar mixture in the supernatant devoid of **2a** and **2b**. Interestingly, we observed that the addition of **3** to this supernatant C_4, C_5 and C_6 sugar mixture directly induced further crystallization of *ribo*-**1** (Supplementary Fig. 35). Therefore, we have demonstrated two complementary pathways to accrue pure crystalline *ribo*-**1** en route to ribonucleotides from the most-complex sugar mixtures explored in prebiotic chemistry, and both pathways are only enabled by the selective formation of aminal **8**. Of note, the pathway described using aminals **8a** and **8b** to build **1** obviates the requirement to access the notoriously unstable sugar ribose[1–7,15].

Prebiotic selection of proteinogenic amino acids. The prebiotic origins of amino acids have been investigated for over 60 years. However, no reported prebiotic synthesis or meteoritic amino acid sample provides the restricted set of amino acids assigned to the genetic code[2,8–13]. For example, recently Sutherland and co-workers demonstrated the stepwise prebiotic syntheses of 12 aminonitrile (**13**) proteinogenic amino acid precursors, but, paradoxically, essential ketones—such as acetone (**2o**), monohydroxyacetone (**2p**) and dihydroxyacetone (**2q**)—are required during the assembly of the branched carbon framework of valine and leucine[13]. Ideally, ketones would be excluded from a prebiotic aminonitrile **13** synthesis because prebiotic ketones

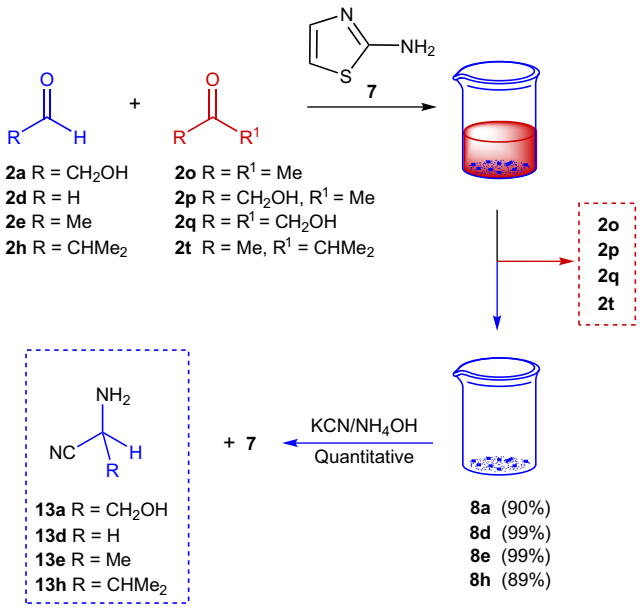


Figure 5 | High-yielding 2-aminothiazole (7)-controlled selective aldehyde reactivity. Selective precipitation of four aldehyde aminals (**8a**, **8d**, **8e** and **8h**) from a mixture of eight aldehydes and ketones (**2a**, **2d**, **2e**, **2h**, **2o**, **2p**, **2q** and **2t**) on reaction with **7** followed by the directed aminonitrile **13** synthesis on equilibration of aminals **8a**, **8d**, **8e** and **8h** with ammonium cyanide (1.5 equiv.) and ammonium hydroxide (5 equiv.) at pH 9.2 demonstrates a direct physicochemical mechanism for proteinogenic amino acid selection.

undergo aminonitrile formation just as effectively as aldehydes, but α,α-disubstituted amino acids are not genetically encoded (Table 1). To the best of our knowledge, there are no previously described mechanisms to discriminate between aldehydes and ketones during aminonitrile synthesis. However, we recognized that aminal **8a** is not only a C_2 sugar, but also an amino acid (serine) precursor, which prompted us to investigate whether 2-aminothiazole (**7**) could facilitate the separation of natural and non-natural amino acid precursors, and thus unite the chemical selection of amino acid and ribonucleotide precursors by a common mechanism. Accordingly, we examined the reaction of **7** with the aldehydes and ketones **2d**–**2w**, which are all prebiotic amino acid precursors[2,8–13]. All the aldehydes furnished their aminals (**8d**–**8n**) in excellent yield (Table 1). In stark contrast, no ketone aminals precipitated, even after extended incubation. This suggests that **7** could be ideally suited to facilitate the facile separation of natural α-substituted amino acid precursors (aldehydes) and non-natural (α,α-disubstituted) amino acid precursors (ketones).

A novel route to amino acids requires aminals **8** to participate directly in the synthesis of aminonitriles **13** without affording *N*-thiazolyl aminonitriles **15**. Therefore, we were pleased to observe that the incubation of ammonium cyanide (**11**) with aminals **8a** and **8d**–**8m** gave a smooth conversion into aminonitriles **13a** and **13d**–**13m**, respectively (Table 1). Even in the absence of ammonia, *N*-thiazolyl aminonitriles **15** were not observed; only a quantitative formation of the relevant cyanohydrin **16** was noted (pH 3–10). Although the incubation of aminal **8b** or **8c** with cyanide yielded a smooth conversion to cyanohydrins **16b** or **16c**, respectively (>95%, 25 °C, pD 7), aminonitriles **13b** and **13c** were not observed because of the rapid imidolactone formation (Supplementary Fig. 85)[31], which suggests their respective amino acids were not genetically encoded because of imidolactone-prohibited aminonitrile synthesis.

The breadth of amino acid precursors available on the early Earth remains unknown, but it is clear that both aldehydes and ketones would have been present in prebiotic mixtures[8–13]. Thus, to investigate **7**-induced selective amino acid synthesis, a mixture of eight prebiotic aldehydes and ketones **2** (Fig. 5) was incubated with **7** (10 equiv.). The quantitative crystallization of aminals **8a**, **8d**, **8e** and **8h** was observed after 24 hours, which allowed all the unwanted ketones **2** (and excess **7**) to be removed with the supernatant to leave behind only pure proteinogenic amino acid precursors. The precipitate, a mixture of aminals **8a**, **8d**, **8e** and **8h**, was then converted smoothly into the corresponding aminonitriles **13a**, **13d**, **13e** and **13h** on the addition of ammonium cyanide (**11**) (Supplementary Fig. 48). Therefore, this reaction is the first selective synthesis of solely natural α-hydrogen amino acids from prebiotic aldehyde/ketone mixtures.

Discussion

To avoid the concomitant synthesis of undesired or irrelevant by-products alongside the desired biologically relevant molecules is one of the central challenges to the development of plausible prebiotic chemistry[1–7,12,13,15,19,22,32,33]. Previous models have advocated that kinetically controlled, segregated syntheses (under different local geochemical conditions) are required to overcome the incompatibility of distinct reactions[1–7]. However, these models are necessarily highly contingent on the rapid exploitation of reagents as and when they form. Accordingly, they are reliant on achieving a specific and controlled order of synthetic steps under geochemical constraints and also they are incompatible with the accumulation or purification of intermediates. Therefore, it is striking that we have discovered a common physicochemical mechanism that controls both α-hydrogen amino acid formation and ribonucleotide assembly from prebiotically plausible mixtures. 2-Aminothiazole (**7**), in conjunction with phosphate acting as a general acid–base catalyst, facilitates an unprecedented sequential accumulation and purification of C_2 and C_3 aldoses (as stable crystalline aminals **8a** and **8b**) from complex mixtures in the specific sequence required for ribonucleotide assembly[3]. The time-resolved separation of **8a** and **8b** observed in our experiments suggests that simple flow systems, such as a river, could provide a model prebiotic environment for the physical separation, accumulation and purification of **8a** and **8b**. The accumulation of **8b** is achieved by dynamic resolution of the C_3 sugars, such that the aldose isomer glyceraldehyde (**2b**) is sequestered from an unfavourable equilibrium with its thermodynamically more-stable ketose isomer dihydroxyacetone (**2q**)[27]. This is remarkable because the equilibration of **2b** to **2q** is not only highly detrimental to stepwise ribonucleotide synthesis, but is also especially rapid (<0.5 h (Supplementary Fig. 2)) under the conditions required for ribonucleotide assembly[3], and there are no previously reported conditions that overcome the unfavourable equilibrium that favours **2q** (ref. 27). Moreover, the formation of aminal **8b** from **2q** drives the recovery of the (first) chirogenic centre (en route to sugars and ribonucleotides), and this recovery is coupled to crystallization. Therefore, we suggest that the crystallizations outlined here may have further implications for the generation of homochiral ribonucleotides from achiral precursors, and investigations to exploit the crystallization of aminals **8** in chiral amplification strategies are currently underway in our laboratory. Indeed, although *rac-ribo*-**1** has only previously been observed to form enantiomorphously twinned crystals (in which individual crystals remain racemic but may contain homochiral domains)[22], we observed the crystallization of novel conglomerates of *ribo*-**1** from a synthesis mediated by aminal **8** (Fig. 1b), in which individual crystals are purely homochiral. It is also striking that a previously described mechanism for amplification and chirality transfer to ribonucleotides—the phosphate-mediated conversion of *ribo*-**1** in to *arabino*-**1** (ref. 24)—is mechanistically equivalent to the conversion of **2b** into **2q**, which maximizes the resolution of C_2 and C_3 sugars by **7**-induced crystallization.

Given the close biological and generational relationship between the proteinogenic amino acids and ribonucleotides, it is highly significant that 7-induced crystallization provides an absolute chemical selection for the natural amino acids (that is, those bearing an α-hydrogen atom) from a complex mixture of Strecker aldehydes and ketones **2** selected for their prevalence in abiotic chemistry. Ketones such as dihydroxyacetone (**2q**), monohydroxyacetone (**2p**) and acetone (**2o**) are required for the prebiotic synthesis of branched-chain amino acids (such as valine and leucine) and lipid precursors[13]. These ketones, and many others, all undergo effective aminonitrile formation (Table 1), which leads to the prebiotic synthesis of non-proteinogenic α,α-disubstituted amino acids alongside the natural α-amino acids, and thus creates a puzzling dichotomy. This has now been resolved by selective sequestration of the natural α-amino acid precursors by accumulation of aminals **8**. Restricting the peptide sequence space is thought to have been advantageous during the origins of life[2,4], and **7** provides a simple chemical-selection tool for α-hydrogen moieties that may underpin the later implementation of biological selection that would have refined (for example, the exclusion of **13f**) and expanded the amino acid repertoire to provide further selective advantages[4,34].

Finally, and importantly, in particular **7** does not inhibit the formation of pentose aminooxazoline (**1**) by Mannich-type reactivity[14,25,35] or the formation of the canonical α-amino acids through participation in a Strecker-type reactivity. Accordingly, it is remarkable that **7** can react with ribonucleotide and amino acid aldehyde precursors and facilitate their purification and accumulation from prebiotic mixtures. These features make **7** an ideal chemical chaperone for prebiotic multistep syntheses.

Data availability statement. The authors declare the data that support the findings of this study are available within the paper and its Supplementary Information files. X-ray crystallographic data were also deposited at the Cambridge Crystallographic Data Centre (CCDC) under the following CCDC deposition numbers: D-*ribo*-**1** (1477052, conglomerate), *rac*-*threo*-**5** (1477054), aminals **8a** (1477040), D-**8b** (1477041), L-**8b** (1477042), *rac*-**8b** (1477045), D-**8c** (1477043), **8d** (1477044), **8e** (1477046), **8f** (1477047), **8g** (1477051) and **8m** (1477048). These can be obtained free of charge from CCDC via www.ccdc.cam.ac.uk/data_request/cif.

Received 18 September 2016; accepted 22 November 2016; published online 16 January 2017

References

1. Eschenmoser, A. Etiology of potentially primordial biomolecular structures: from vitamin B_{12} to the nucleic acids and an inquiry into the chemistry of life's origin: a retrospective. *Angew. Chem. Int. Ed.* **50,** 12412–12472 (2011).
2. Sutherland, J. D. The origin of life—out of the blue. *Angew. Chem. Int. Ed.* **55,** 104–121 (2016).
3. Powner, M. W., Gerland, B. & Sutherland, J. D. Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. *Nature* **459,** 239–242 (2009).
4. Joyce, G. F. The antiquity of RNA-based evolution. *Nature* **418,** 214–221 (2002).
5. Benner, S. A., Kim, H.-Y. & Yang, Z. Setting the stage: the history, chemistry and geology behind RNA. *Cold Spring Harb. Perspect. Biol.* **4,** a003541 (2012).
6. Robertson, M. P. & Joyce, G. F. The origins of the RNA world. *Cold Spring Harb. Perspect. Biol.* **4,** a003608 (2012).
7. Krishnamurthy, R. On the emergence of RNA. *Isr. J. Chem.* **55,** 837–850 (2015).
8. Johnson, A. P. *et al.* The Miller volcanic spark discharge experiment. *Science* **322,** 404 (2008).
9. Chyba, C. & Sagan, C. Endogenous production, exogenous delivery and impact-shock synthesis of organic molecules: an inventory for the origins of life. *Nature* **355,** 125–132 (1992).
10. Miller, S. L. Production of amino acids under possible primitive Earth conditions. *Science* **117,** 528–529 (1953).
11. Pizzarello, S., Cooper, G. & Flynn, G. in *Meteorites and the Early Solar System II* (eds Lauretta, D. S. & McSween, H. Y. Jr) 625–651 (Univ. Arizona Press, 2006).
12. Wolman, Y., Haverland, W. J. & Miller, S. L. Nonprotein amino acids from spark discharges and their comparison with the Murchison meteorite amino acids. *Proc. Natl Acad. Sci. USA* **69,** 809–811 (1972).
13. 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).
14. Powner, M. W., Zheng, S. L. & Szostak, J. W. Multicomponent assembly of proposed DNA precursors in water. *J. Am. Chem. Soc.* **134,** 13889–13895 (2012).
15. Shapiro, R. Prebiotic ribose synthesis: a critical analysis. *Origin Life Evol. Biosphere* **18,** 71–85 (1988).
16. Ricardo, A., Carrigan, M. A., Olcott, A. N. & Benner, S. A. Borate minerals stabilize ribose. *Science* **303,** 196 (2004).
17. Breslow, R. & Cheng, Z. L. L-Amino acids catalyze the formation of an excess of D-glyceraldehyde, and thus of other D-sugars, under credible prebiotic conditions. *Proc. Natl Acad. Sci. USA* **107,** 5723–5725 (2010).
18. Sagi, V. S. *et al.* Exploratory experiments on the chemistry of the 'glyoxylate scenario': formation of ketosugars from dihydroxyfumarate. *J. Am. Chem. Soc.* **134,** 3577–3589 (2012).
19. Islam, S. *et al.* Detection of potential TNA and RNA nucleoside precursors in a prebiotic mixture by pure shift diffusion-ordered NMR spectroscopy. *Chem. Eur. J.* **19,** 4586–4595 (2013).
20. Meinert, C. *et al.* Ribose and related sugars from ultraviolet irradiation of interstellar ice analogs. *Science* **352,** 208–212 (2016).
21. Becker, S. *et al.* A high-yielding, strictly regioselective prebiotic purine nucleoside formation pathway. *Science* **352,** 833–836 (2016).
22. Springsteen, G. & Joyce, G. F. Selective derivatization and sequestration of ribose from a prebiotic mix. *J. Am. Chem. Soc.* **126,** 9578–9583 (2004).
23. Anastasi, C., Crowe, M. A., Powner, M. W. & Sutherland, J. D. Direct assembly of nucleoside precursors from two- and three-carbon units. *Angew. Chem. Int. Ed.* **45,** 6176–6179 (2006).
24. Powner, M. W. & Sutherland, J. D. Phosphate-mediated interconversion of *ribo*- and *arabino*-configured prebiotic nucleotide intermediates. *Angew. Chem. Int. Ed.* **49,** 4641–4643 (2010).
25. Powner, M. W., Sutherland, J. D. & Szostak, J. W. Chemoselective multicomponent one-pot assembly of purine precursors in water. *J. Am. Chem. Soc.* **132,** 16677–16688 (2010).
26. Saul, R. *et al.* Reaction of 1,3-disubstituted acetone derivatives with pseudohalides: a simple approach to spiro[4.4]nonane-type bis-oxazolidines and -imidazolidines (bicyclic carbamates, thiocarbamates, ureas, and thioureas). *Eur. J. Org. Chem.* **1,** 205–209 (2000).
27. Nagorski, R. W. & Richard, J. P. Mechanistic imperatives for aldose-ketose isomerization in water: specific, general base- and metal-ion-catalyzed isomerization of glyceraldehyde with proton and hydride transfer. *J. Am. Chem. Soc.* **123,** 794–802 (2001).
28. Budin, I. & Szostak, J. W. Expanding roles for diverse physical phenomena during the origin of life. *Annu. Rev. Biophys.* **39,** 245–263 (2010).
29. Sagan, C. & Khare, B. N. Long-wavelength ultraviolet photoproduction of amino acids on the primitive Earth. *Science* **173,** 417–420 (1971).
30. Catsimpoolas, N. & Wood, J. L. Specific cleavage of cystine peptides by cyanide. *J. Biol. Chem.* **241,** 1790–1796 (1966).
31. Serianni, A. S., Nunez, H. A. & Barker, R. Cyanohydrin synthesis: studies with [^{13}C]cyanide. *J. Org. Chem.* **45,** 3329–3341 (1980).
32. Schwartz, A. W. Intractable mixtures and the origin of life. *Chem. Biodiv.* **4,** 656–664 (2007).
33. Schwartz, A. W. Evaluating the plausibility of prebiotic multistage syntheses. *Astrobiology* **13,** 784–789 (2013).
34. Knight, R. D. & Landweber, L. F. The early evolution of the genetic code. *Cell* **101,** 569–572 (2000).
35. 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).

Acknowledgements

This work was supported by the Simons Foundation (318881), the Engineering and Physical Sciences Research Council (EP/K004980/1), the Leverhulme Trust (RGP-2013-189) and through an award from the Origin of Life Challenge (M.W.P.) and a UCL Excellence Fellowship (D.-K.B.). The authors thank K. Karu for assistance with mass spectrometry and A. E. Aliev for assistance with NMR spectroscopy.

Author contributions

M.W.P. conceived the research. M.W.P. and S.I. designed and analysed the experiments. S.I. conducted the experiments. D.-K.B. performed the crystallographic analyses. M.W.P. and S.I. wrote the paper.

Additional information

Supplementary information and chemical compound information are available in the online version of the paper. Reprints and permissions information is available online at www.nature.com/reprints. Correspondence and requests for materials should be addressed to M.W.P.

Competing financial interests

The authors declare no competing financial interests.