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

Saidul Islam,[1] Dejan-Krešimir Bučar[1] and Matthew W. Powner[1]

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

A central problem for prebiotic synthesis of the biological amino acids and nucleotides is avoiding the concomitant synthesis of undesired or irrelevant byproducts. Additionally, multi-step 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 isomerisation of glyceraldehyde to 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 synthesis of the non-proteinogenic α,α-disubstituted analogues. The common physicochemical mechanism that controls proteinogenic amino acid and ribonucleotide assembly from prebiotic mixtures suggests 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 recognised 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*"*.[4-7] For example, the most-efficient and specific proposed prebiotic pathway to the pyrimidine ribonucleotides requires synthesis of the key intermediate pentose aminooxazoline (**1**) (Fig. 1a).[2-6,22-24] However, the plausibility of this proposed prebiotic synthesis of pentose aminooxazoline (**1**) has been questioned because it is contingent upon the strictly controlled sequential delivery of pure glycolaldehyde (**2a**) to cyanamide (**3**) to yield 2-aminooxazole (**4**), followed by pure glyceraldehyde (**2b**) to 2-aminooxazole (**4**) to yield the desired product (Fig. 1a). This is a serious problem because both of these reactions lack the intrinsic selectivity required to exclusively yield their respective products (2-aminooxazole (**4**) and pentose aminooxazoline (**1**)) from mixtures of glycolaldehyde (**2a**) and glyceraldehyde (**2b**). The problem becomes increasingly worse in the presence of other sugars. Without a separate and sequential delivery of glycolaldehyde (**2a**) and glyceraldehyde (**2b**), a complex mixture of undesirable byproducts 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 the ketose isomer

dihydroxyacetone (**2q**) than the aldose isomer glyceraldehyde (**2b**) (**2q**/**2b** = 8.5:1; 25 °C, pH 7; Supplementary Fig. 1–4). Under the conditions required for the formation of 2-aminooxazole (**4**),[3] in aqueous solution aldose glyceraldehyde (**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 reaction of triose sugars with 2-aminooxazole (**4**) the predominant product is the non-natural branched apiose **5** rather than the desired product pentose aminooxazoline (**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 glycolaldehyde (**2a**) and glyceraldehyde (**2b**) (1:1) in phosphate buffer yields **2a**/**2b**/**2q** (1:0.11:0.89). Reaction of this mixture with 2-aminooxazole (**4**; 1 equiv.) returns pentose aminooxazoline (**1**; 11%) as very minor components 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 glycolaldehyde (**2a**) and glyceraldehyde (**2b**) and in 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 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 glycolaldehyde (**2a**) and glyceraldehyde (**2b**) is ostensively implausible, 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 glycolaldehyde (**2a**) and glyceraldehyde (**2b**) when we observed precipitation of glycolaldehyde (**2a**) and homochiral D-glyceraldehyde (D-**2b**) from water as aminals **8a** and D-**8b**.[14] As discussed below, 2-aminothiazole (**7**) is expected to be abundant in a cyanosulfidic environment.[29] We now report that 2-aminothiazole (**7**) induced crystallisation resolves several long-standing problems inherent to ribonucleotide selection. We demonstrate the time-resolved sequestration of glycolaldehyde (**2a**) and glyceraldehyde (**2b**) from complex sugar mixtures as separate stable crystalline reservoirs in the form of the corresponding aminals **8a** and **8b**. Aminals **8a** and **8b** allow for sequential 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 glyceraldehyde (**2b**) from the aldose/ketose equilibrium, which overturns the inherent thermodynamic bias towards the ketose isomer dihydroxyacetone (**2q**).[5,13,27]

For 2-aminothiazole (**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 2-aminothiazole (**7**) in complex sugar mixtures. 2-Aminothiazole (**7**) was synthesised in near-quantitative yield (>95%) from cysteine precursor β-mercaptoacetaldehyde (**9a**) and cyanamide (**3**) (Fig. 2). Very high conversion was observed even in stoichiometric competition with 26 other aldehyde, ketone and sugar species present in the mixture (including **2a-w** and C_2-, C_3-, C_4-, C_5- and C_6 sugars; Supplementary Fig. 21 and 22). Furthermore, because the addition of thiols to cyanamide (**3**) is reversible, we also explored the

synthesis of 2-aminothiazole (**7**) from β-mercaptoacetaldehyde (**9a**) in the presence of five other stoichiometric sulfides/thiols (**9b**-**f**; Fig. 2) and observed a highly selective reaction; 2-aminothiazole (**7**) was obtained in 86% yield. Moreover, isothiourea (**10**) undergoes efficient thiol exchange with β-mercaptoacetaldehyde (**9a**) to furnish 2-aminothiazole (**7**) in 72% yield under mild aqueous conditions (pH 7, r.t., 16 h; Fig. 2). Finally, we were drawn to investigate the synthesis of 2-aminothiazole (**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 2-aminothiazole (**7**; 14% yield), whereas the reaction of disulfide **2n** with cyanide and cyanamide (**3**) gave 2-aminothiazole (**7**) in up to 75% yield (Supplementary Fig. 23). The facile and predisposed assembly of 2-aminothiazole (**7**) from a complex mixture of aldehydes, ketones, sulfides and thiols in water suggests that 2-aminothiazole (**7**) is a highly apposite reagent in the search for prebiotic selectivity (Fig. 2).

Having established that 2-aminothiazole (**7**) is 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 glycolaldehyde (**2a**; 0.25 M), glyceraldehyde (**2b**; 0.25 M) and 2-aminothiazole (**7**; 0.5 M) in phosphate buffer (pH 7, 25 ºC) and monitored the reaction over 24 h. We initially (0–6 h) observed a mixture of aminals **8a** and **8b**. However, surprisingly, after 12 h complete resolution of the C_2- and C_3 sugars was observed: **8a** was the only aminal isolated (77%, 24 h; Fig 3a & Supplementary Table 3) and all 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 glycolaldehyde (**2a**) with C_3-aminal **8b** (1:1, 25 °C, 3 d, pH 7) in phosphate buffer and observed accumulation of the C_2-aminal **8a** in 56% yield. Conversely, incubation of glyceraldehyde (**2b**) with C_2-aminal **8a** (1:1,

25 °C, 3 d, pH 7) in phosphate buffer C_3-aminal **8b** was not formed—only triose equilibration (**2b**↔**2q**) was observed, even with a two-fold excess of glyceraldehyde (**2b**). Therefore, it is clear that equilibration of glyceraldehyde (**2b**) to its more-stable ketose isomer dihydroxyacetone (**2q**) drives the selective accumulation of C_2-aminal **8a**. However, remarkably, incubation of dihydroxyacetone (**2q**) with either 2-aminothiazole (**7**) or β-mercaptoacetaldehyde (**9a**) and cyanamide (**3**) (to effect *in situ* synthesis of 2-aminothiazole (**7**)) delivered glyceraldehyde (**2b**) aminal *rac*-**8b** in up to 87% yield (Fig. 3a and Supplementary Table 6), which thereby deposited a reservoir of crystalline aldose isomer glyceraldehyde (**2b**) and completely overturned the thermodynamic C_3 ketose isomer preference observed in aqueous solution. In contrast, phosphate-induced isomerisation 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 dihydroxyacetone (**2q**) and erythrulose (**2w**) after incubation (23 d) with 2-aminothiazole (**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 crystallisation of aminal **8b** (Fig. 3b & Supplementary Table 8).

Selective synthesis of pentose aminooxazoline (1) from a complex sugar mixture

We recognised 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 spatially separate glycolaldehyde (**2a**) and glyceraldehyde (**2b**) that 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., 3 d; Fig. 4b) to allow triose equilibration, and then 2-aminothiazole (**7**; 6 equiv.) was added. Over the

course of 2 h the selective crystallisation of pure aminal **8a** (>95%) was observed (Fig. 4c). Further incubation of the supernatant, following isolation of pure glycolaldehyde-aminal **8a**, resulted in crystallisation of pure glyceraldehyde-aminal **8b** (61%, 2–17 d; Fig. 4d). No further crystallisation was observed. The onset of crystallisation of **8b** is slow (9% after 24 h), 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 2-aminooxazole (**4**) with aminal **8b**, which crystallised after **8a** from the same complex sugar solution, led to selective formation of pentose aminooxazoline (**1**), followed by direct crystallisation of pure *ribo*-**1** (Fig. 4f).

Having demonstrated the crystallisation-controlled synthesis of *ribo*-**1** for 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 previously been reported that the simplest sugars glycolaldehyde (**2a**) and glyceraldehyde (**2b**), due to their rapid rate of reaction, thwarted previous attempts to crystallise *ribo*-**1** by the action of cyanamide (**3**) on complex sugar mixtures.[22] However, our sequential 2-aminothiazole (**7**) induced crystallisation from an aqueous solution of C_2-, C_3-, C_4-, C_5- and C_6 sugars exploits the reactivity of glycolaldehyde (**2a**) and glyceraldehyde (**2b**) to facilitate their sequesteration from solution resulting in a residual sugar mixture in the supernatant devoid of glycolaldehyde (**2a**) and glyceraldehyde (**2b**). Interestingly, we observed that addition of cyanamide (**3**) to this supernatant C_4-, C_5- and C_6 sugar mixture directly induced further crystallisation 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 selective formation of aminal **8**. However, it is of note that the pathway described using aminals **8a** and **8b** to build the pentose aminooxazolines **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 under investigation 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, Sutherland and co-workers recently 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 assembly of the branched carbon framework of valine and leucine.[13] Ideally ketones would be excluded from prebiotic aminonitrile **13** synthesis because prebiotic ketones 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 recognised 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, thus uniting the chemical selection of amino acid and ribonucleotide precursors by a common mechanism. Accordingly, we examined the reaction of 2-aminothiazole (**7**) with aldehydes and ketones **2d-w**, which are all prebiotic amino acid precursors.[2,8-13] All of the aldehydes furnished their aminals (**8d-w**) in excellent yield (Table 1). In stark contrast, no ketone aminals precipitated, even after extended incubation. This suggests that 2-aminothiazole

(**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 incubation of ammonium cyanide (**11**) with aminals **8a** and **8d-m** gave smooth conversion to aminonitriles **13a** and **13d-m**, respectively (Table 1). Even in the absence of ammonia, *N*-thiazolyl aminonitriles **15** were not observed; only quantitative formation of the relevant cyanohydrin **16** was noted (pH 3–10). Although 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 due to rapid imidolactone formation (Supplementary Fig. 85),[31] suggesting their respective amino acids were not genetically encoded due to 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 2-aminothiazole (**7**) induced selective amino acid synthesis, a mixture of eight prebiotic aldehydes and ketones **2** (Fig. 5) was incubated with 2-aminothiazole (**7**; 10 equiv.). The quantitative crystallisation of aminals **8a**, **8d**, **8e** and **8h** was observed after 24 h, which allowed all unwanted ketones **2** (and excess 2-aminothiazole (**7**)) to be removed with the supernatant leaving behind only pure proteinogenic amino acid precursors. The precipitate, a mixture of aminals **8a**, **8d**, **8e** and **8h**, was then smoothly converted to the corresponding aminonitriles **13a**, **13d**, **13e** and **13h** upon 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

Avoiding the concomitant synthesis of undesired or irrelevant byproducts 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 upon the rapid exploitation of reagents as and when they form. Accordingly, these models rely upon achieving of a specific and controlled order of synthetic steps under geochemical constraints, and are also 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 a 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 glyceraldehyde (**2b**) to dihydroxyacetone (**2q**) is not only highly detrimental to stepwise ribonucleotide synthesis, but 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 dihydroxyacetone (**2q**).[27] Moreover, the formation of aminal **8b** from dihydroxyacetone (**2q**) drives the recovery of the (first) chirogenic centre (en route to sugars and ribonucleotides), and this recovery is coupled to crystallisation. Therefore, we suggest that the crystallisations outlined here may have further implications for the generation of homochiral ribonucleotides from achiral precursors, and investigations to exploit the crystallisation 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 crystallisation of novel conglomerates of *ribo*-**1** from aminal **8** mediated synthesis (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** to *arabino*-**1** (ref. 24)—is mechanistically equivalent to the conversion of glyceraldehyde (**2b**) to dihydroxyacetone (**2q**), which maximises the resolution of C_2- and C_3 sugars by 2-aminothiazole (**7**) induced crystallisation.

Given the close biological and generational relationship between the proteinogenic amino acids and ribonucleotides, it is highly significant that 2-aminothiazole (**7**) induced crystallisation provides absolute chemical selection for the natural amino acids (i.e. 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, thus creating a puzzling dichotomy. This has now been resolved by selective sequestration of the natural α-amino acid precursors by accumulation of aminals **8**. Restricting peptide sequence space is thought to have been advantageous during the origins of life,[2,4] and 2-aminothiazole (**7**) provides a simple chemical-selection tool for α-hydrogen moieties that may underpin the later implementation of biological selection that would have refined (e.g. the exclusion of **13f**) and expanded the amino acid repertoire to provide further selective advantages.[4,34]

Finally, and importantly, it is of particular note that 2-aminothiazole (**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 Strecker-type reactivity. Accordingly, it is remarkable that 2-aminothiazole (**7**) can react with ribonucleotide and amino acid aldehyde precursors and facilitate their purification and accumulation from prebiotic mixtures. These features make 2-aminothiazole (**7**) an ideal chemical chaperone for prebiotic multi-step syntheses.

Data availability statement

The authors declare that the data supporting the findings of this study are available within the paper and its supplementary information files. X-ray crystallographic data was 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), **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. Supplementary information and compound

characterisation data are available in the online version of the paper. Reprints and permissions information is available online at www.nature.com/reprints. Correspondence should be addressed to M.W.P. (matthew.powner@ucl.ac.uk).

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**, 125–32 (1992).

[10] Miller, S. L. Production of amino acids under possible primitive Earth conditions. *Science* **117**, 528–29 (1953).

[11] Pizzarello, S., Cooper, G. & Flynn, G. The nature and distribution of the organic material in carbonaceous chondrites and interplanetary dust particles. In Meteorites and the early solar system II, (Eds. Lauretta, D. S. & McSween, Jr., H. Y.), Univ. of Arizona Press, 625–651 (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. U.S.A.* **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. *Nature 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. U.S.A.* **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. *Nature Chem.* **3**, 704–706 (2011).

Acknowledgments

This work was supported by the Simons Foundation (31881), 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 Dr. K. Karu for assistance with Mass Spectrometry and Dr. 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.

Competing Financial Interests Statement

The authors declare no competing financial interests.

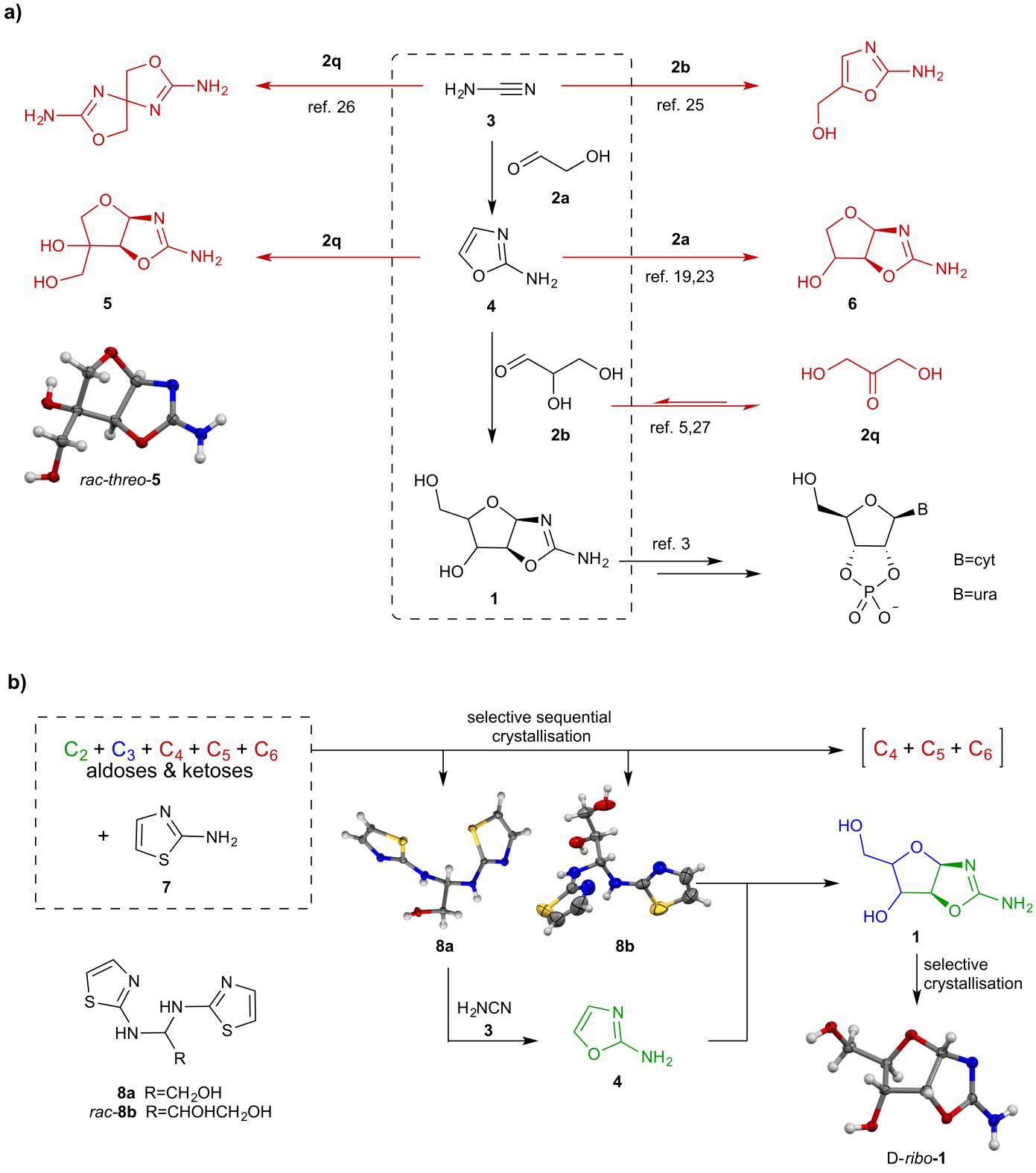

a)
2q
ref. 26
2b
ref. 25
3
2a
2q
2a
ref. 19,23
5
4
6
rac-threo-5
2b
ref. 5,27
2q
1
ref. 3
B=cyt
B=ura
b)
C2 + C3 + C4 + C5 + C6
aldoses & ketoses
+
7
selective sequential
crystallisation
[C4 + C5 + C6]
8a
8b
1
selective
crystallisation
H2NCN
3
4
8a R=CH2OH
rac-8b R=CHOHCH2OH
D-ribo-1

Figure 1. Prebiotic ribonucleotide synthesis. a) Stepwise nucleotide synthesis from pure reagents (previous work). Pyrimidine ribonucleotide synthesis (black arrows) requires stepwise addition of glycolaldehyde (**2a**) and glyceraldehyde (**2b**) to avoid deleterious byproduct formation (red arrows). 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).** Crystallisation-controlled synthesis of ribose aminooxazoline (*ribo*-**1**) from a sugar mixture containing an equilibrating mixture of glycolaldehyde (**2a**), glyceraldehyde (**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 crystallisation of glycolaldehyde-aminal **8a** and *rac*-glyceraldehyde aminal **8b**. Single-crystal X-ray structures acquired directly from 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 crystallised from the reaction of C_3-aminal (*rac*-**8b**) with 2-aminooxazole (**4**), synthesised from the exposure of C_2-aminal **8a** to cyanamide (**3**).

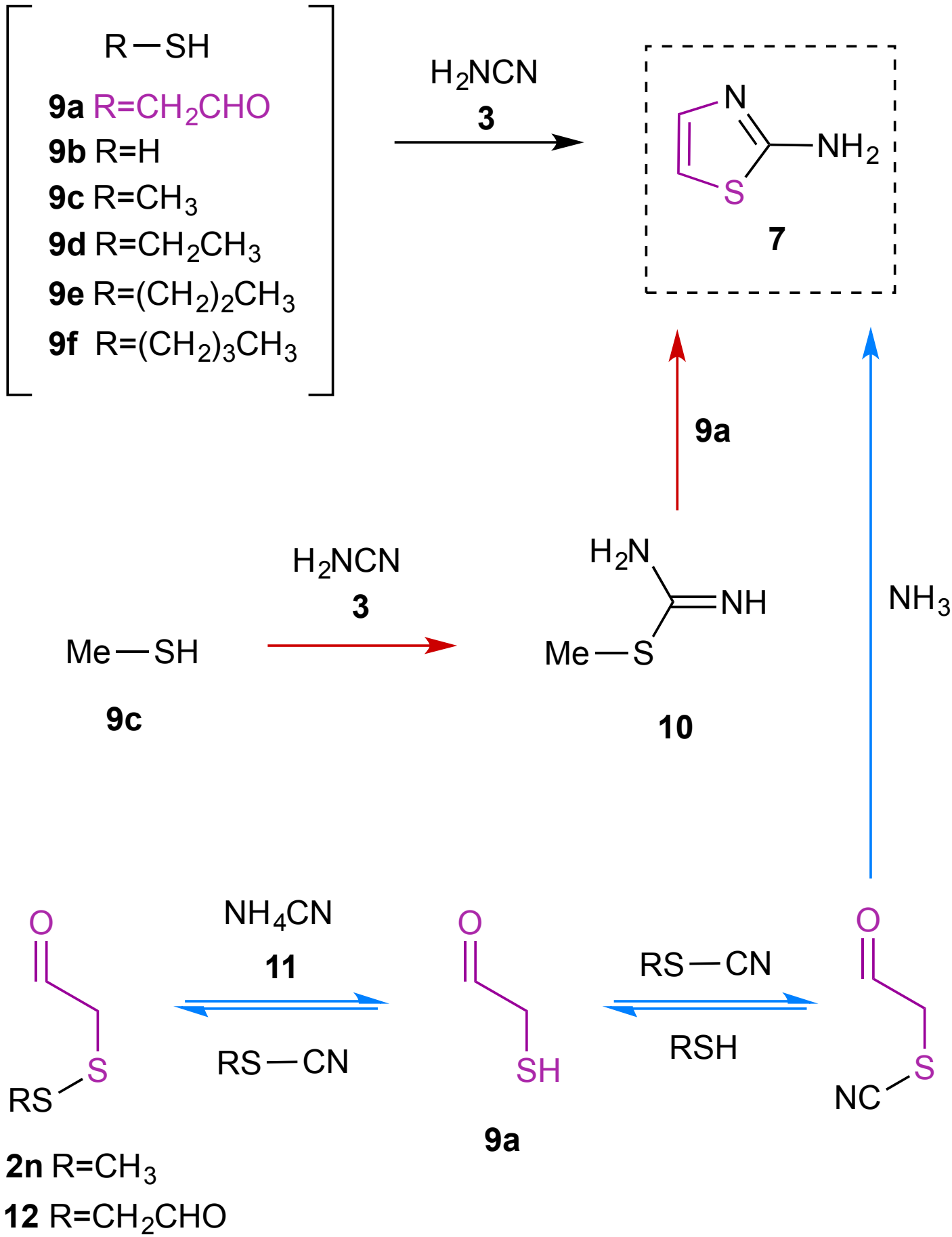

R—SH
9a R=CH2CHO
9b R=H
9c R=CH3
9d R=CH2CH3
9e R=(CH2)2CH3
9f R=(CH2)3CH3
H2NCN
3
N
NH2
S
7
9a
H2NCN
3
Me—SH
9c
H2N
NH
Me—S
10
NH3
O
S
RS
NH4CN
11
RS—CN
O
SH
9a
RS—CN
RSH
O
S
NC
2n R=CH3
12 R=CH2CHO

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

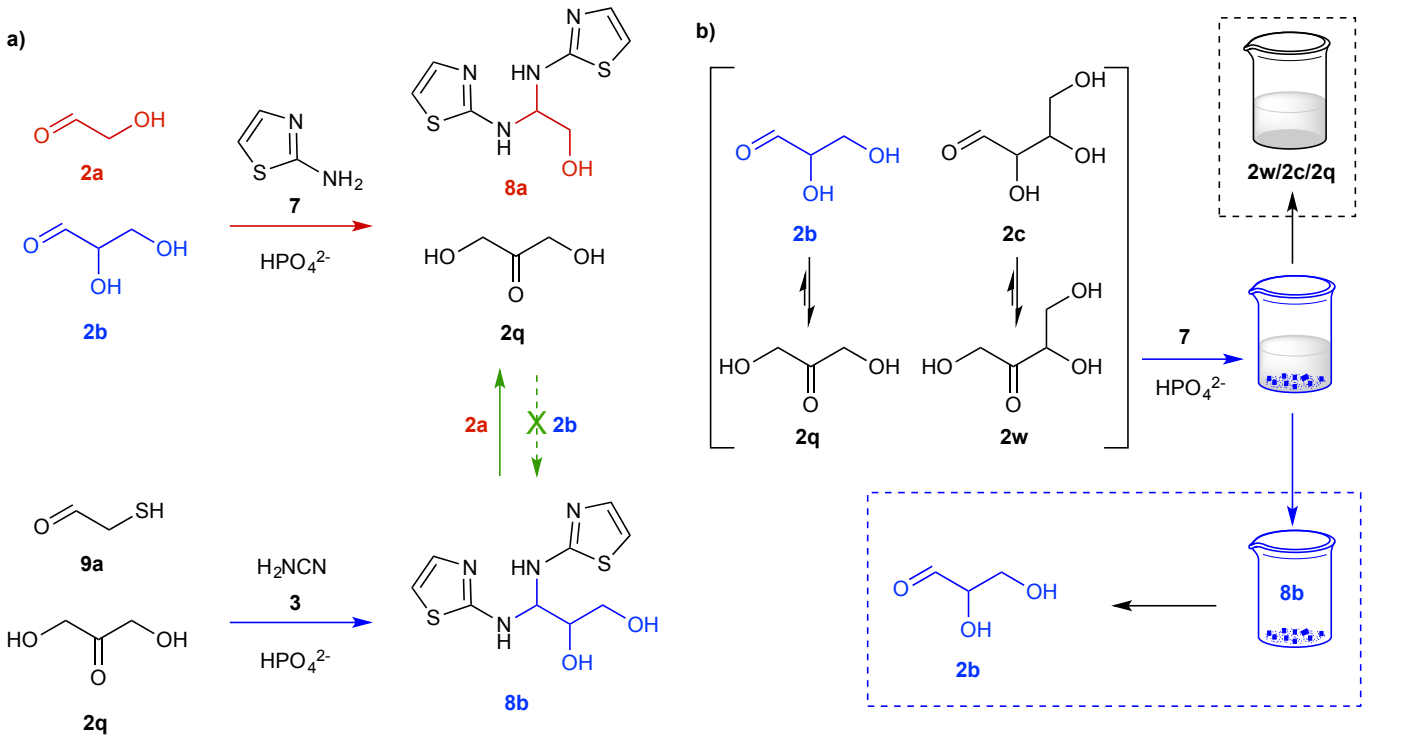

a)
2a
2b
7
HPO4 2-
8a
2q
2a
2b
9a
2q
H2NCN
3
HPO4 2-
8b
b)
2b
2c
2q
2w
7
HPO4 2-
2w/2c/2q
8b
2b

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 glycolaldehyde (**2a**, 1 equiv.), glyceraldehyde (**2b**, 1 equiv.) and 2-aminothiazole (**7**, 2 equiv.) (red arrows). Selective multicomponent glyceraldehyde-aminal **8b** synthesis from dihydroxyacetone (**2q**), β-mecaptoacetaldehyde (**9a**) and cyanamide (**3**) (blue arrows), and selective conversion of glyceraldehyde-aminal **8b** to glycolaldehyde aminal **8a** upon incubation with glycolaldehyde (**2a**, 1 equiv.) in pH 7 phosphate buffer (green arrows) demonstrating the thermodynamically controlled selection of glycolaldehyde aminal **8a** and dihydroxyacetone (**2q**) from mixtures of C_2- and C_3-sugars. **b)** C_3-Specific aldose sequestration by crystallisation of aminal **8b** from an equilibrating C_3- and C_4-sugar mixture (**2b**, **2c**, **2q** and **2w**) upon reaction with 2-aminothiazole (**7**) in phosphate buffer providing a facile and highly selective separation of C_3- and C_4-sugars.

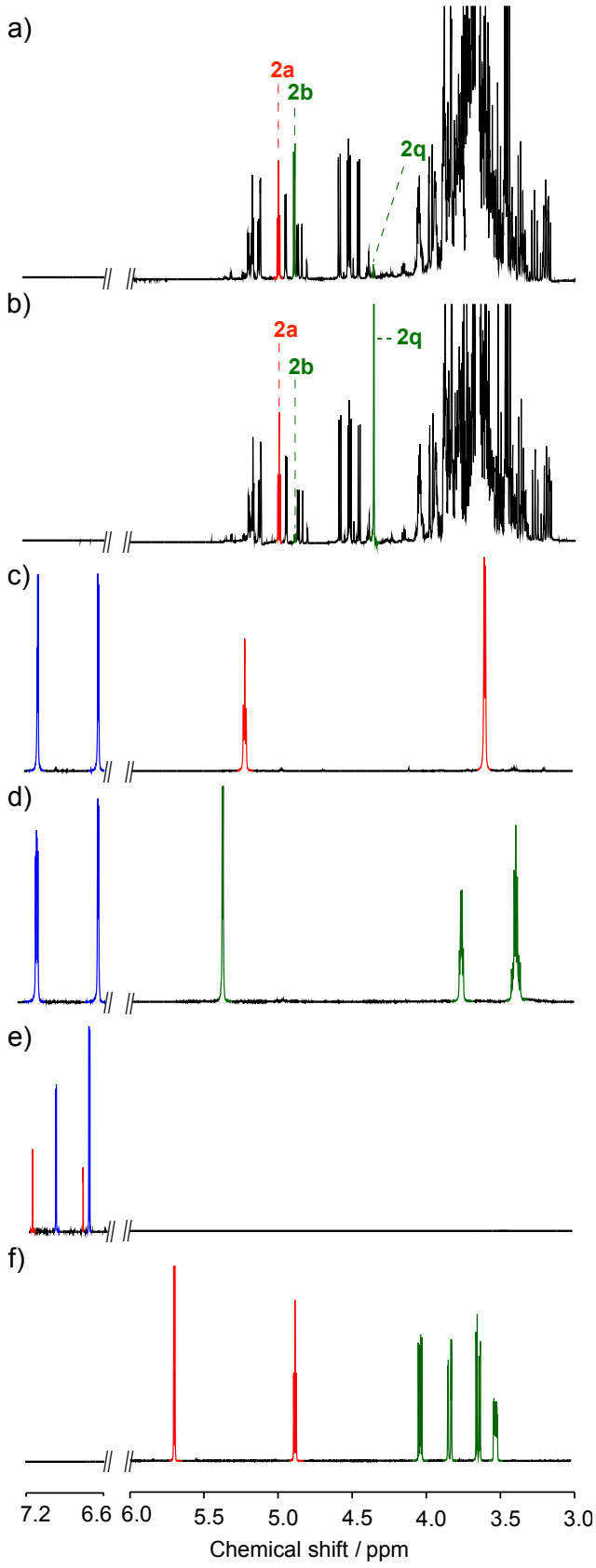

a)
2a
2b
2q
b)
2a
2b
2q
c)
d)
e)
f)
7.2
6.6
6.0
5.5
5.0
4.5
4.0
3.5
3.0
Chemical shift / ppm

Figure 4. Convergent crystallisation-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 glycolaldehyde (**2a**; red), *rac*-glyceraldehyde (**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.). **a)** Initial sugar mixture. **b)** Equilibrated sugar mixture after three days showing conversion of glyceraldehyde (**2b**; green, 4.90 ppm, doublet, J = 5.4 Hz) to dihydroxyacetone (**2q**; green, 4.30 ppm, singlet). **c)** Crystalline precipitate **8a** accumulated 0–2 h after addition of **7** (blue). **d)** Crystalline precipitate **8b** accumulated 48–480 h after addition of **7**. **e)** Quantative transformation of aminal **8a** and cyanamide (**3**) to a 1:2 mixture of 2-aminooxazole (**4**; red) and 2-aminothiazole (**7**; blue). **f)** Crystalline conglomerate D-*ribo*-**1** isolated upon exposure of precipitate **8b** to a 1:2 mixture of **4**/**7**.

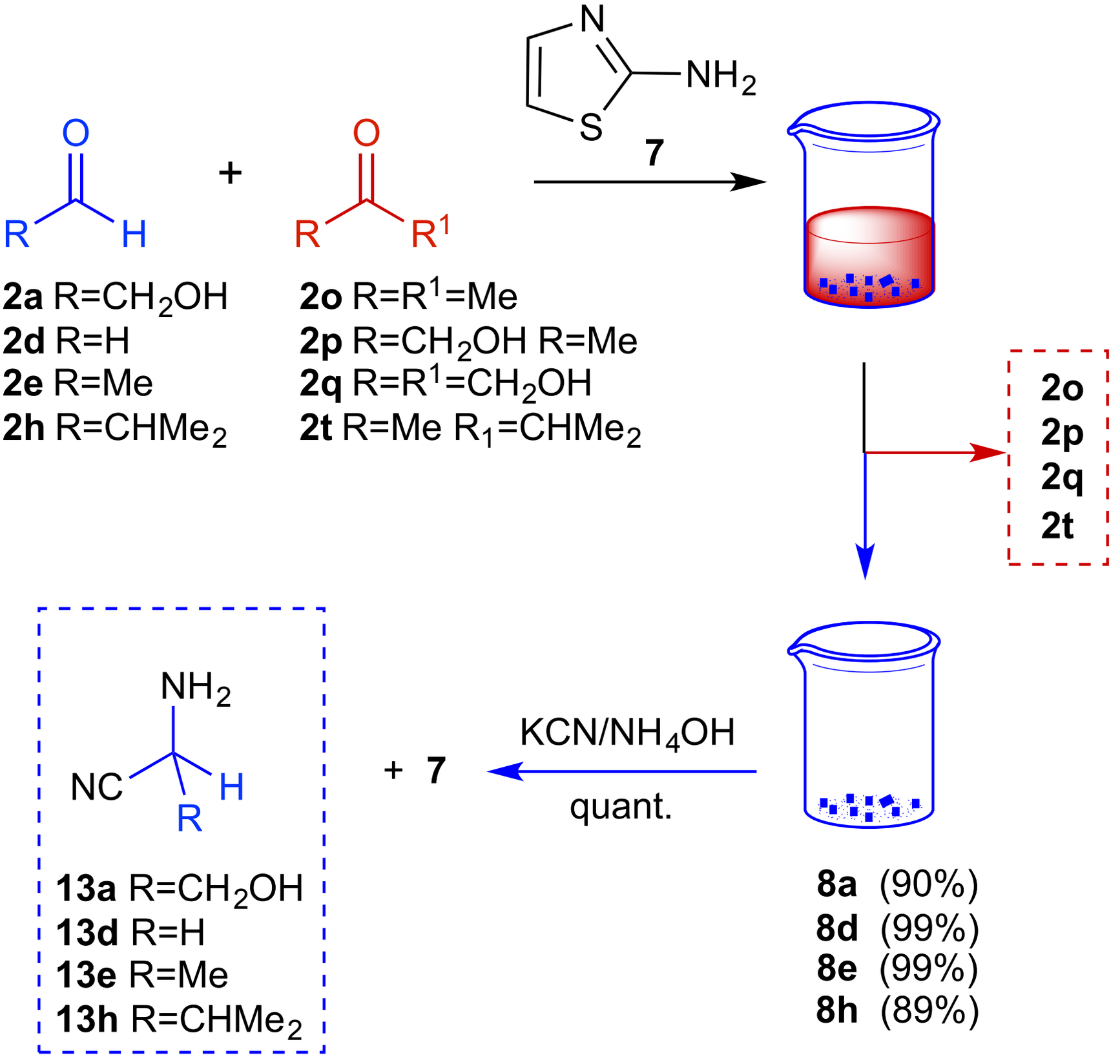
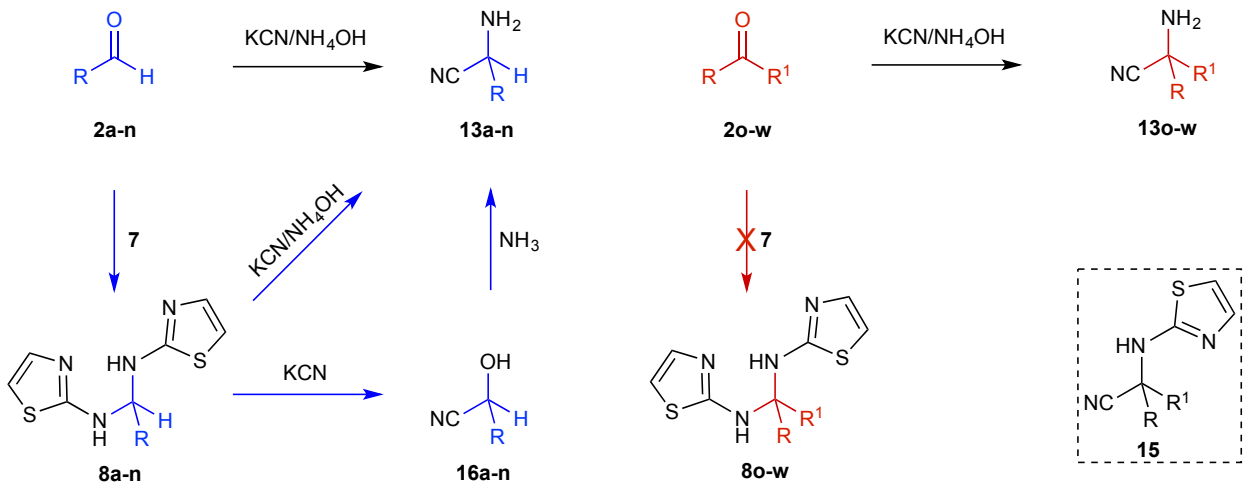

N
NH2
S
7
O
R
H
+
O
R
R1
2a R=CH2OH
2d R=H
2e R=Me
2h R=CHMe2
2o R=R1=Me
2p R=CH2OH R=Me
2q R=R1=CH2OH
2t R=Me R1=CHMe2
2o
2p
2q
2t
KCN/NH4OH
quant.
+ 7
NH2
NC
H
R
13a R=CH2OH
13d R=H
13e R=Me
13h R=CHMe2
8a (90%)
8d (99%)
8e (99%)
8h (89%)

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**) upon reaction with 2-aminothiazole (**7**), followed by directed aminonitrile **13** synthesis upon equilibration of aminals **8a**, **8d**, **8e** and **8h** with ammonium cyanide (1.5 equiv.) and ammonium hydroxide (5 equiv.) at pH 9.2 demonstrating a direct physicochemical mechanism for proteinogentic amino acid selection.

Tables and Table Captions



2	R	R¹	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	-	-

Table 1. Prebiotic aminonitrile synthesis. Non-selective synthesis of aminonitriles **13a-w** from aldehydes and ketones by the traditional Strecker pathway (black arrows; previous work). Aldehyde aminal **8** selective synthesis of aminonitriles **13a** and **3d-m** (blue arrows; this work); aminonitriles **15** were not observed. Isolated yields of aminals **8a**-**w** from the reaction of aldehydes/ketones **2a**-**w** (500 mM) and 2-aminothiazole (**7**; 1 M, pH 7, 24 h) are reported. Aminonitriles **13a**-**n** were obtained by equilibration of aminals **8a**-**n** with potassium cyanide (1.5 equiv.) and ammonium hydroxide (5 equiv.) at pH 9.2; isolated yields are reported. Aminonitriles **13o**-**w** were acquired by equilibration of ketones **2o**-**w** (1 equiv.), potassium cyanide (1.5 equiv.) and ammonium hydroxide (5 equiv.) at pH 9.2; isolated yields are reported. * Homochiral D-**8b** has previously been crystallised 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 following amide hydrolysis.[13] ‡ **8n** is converted to the Strecker precursor of cysteine **9a** upon disulfide reduction.

TOC summary: 2-Aminothiazole, a hybrid of prebiotic amino acid and nucleotide precursors, sequentially accumulates and purifies glycolaldehyde and glyceraldehyde from complex mixtures in the order required for ribonucleotide synthesis, dynamically resolves glyceraldehyde from its ketose-isomer dihydroxyacetone and provides the first strategy to select natural amino acids from abiotic aldehydes and ketones.