Multicomponent Assembly of Proposed DNA Precursors in Water

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Published Version</td>
<td>doi:10.1021/ja306176n</td>
</tr>
<tr>
<td>Citable link</td>
<td><a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:10473955">http://nrs.harvard.edu/urn-3:HUL.InstRepos:10473955</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA</a></td>
</tr>
</tbody>
</table>
Multicomponent Assembly of Proposed DNA Precursors in Water

Matthew W. Powner,§ Shao-Liang Zheng,§ and Jack W. Szostak§

†Department of Chemistry, University College London, Christopher Ingold Laboratories, 20 Gordon Street, London, WC1H 0AJ, U.K.
‡Harvard X-ray Crystallography Centre, 12 Oxford Street, Cambridge, Massachusetts 02138, United States
§Howard Hughes Medical Institute and Department of Molecular Biology and Center for Computational and Integrative Biology, Massachusetts General Hospital, 185 Cambridge Street, Boston, Massachusetts 02114, United States

Supporting Information

ABSTRACT: We propose a novel pathway for the prebiotic synthesis of 2′-deoxynucleotides. Consideration of the constitutional chemical relationships between glycolaldehyde and β-mercapto-acetaldehyde, and the corresponding proteinogenic amino acids, serine and cysteine, led us to explore the consequences of the corresponding sulfur substitution for our previously proposed pathways leading to the canonical ribonucleotides. We demonstrate that just as 2-aminooxazole—an important prebiotic ribonucleotide precursor—is readily formed from glycolaldehyde and cyanamide, so is 2-aminothiazole formed from β-mercapto-acetaldehyde and cyanamide in water at neutral pH. Indeed, both the oxazole and the thiazole can be formed together in a one-pot reaction, and can be copuriﬁed by crystallization or sublimation. We then show that 2-aminothiazole can take part in a 3-component carbon–carbon bond-forming reaction in water that leads to the diastereoselective synthesis of masked 2′-thiosugars regiospeciﬁcally tethered to purine precursors, which would lead to 2′-deoxynucleotides upon desulfurization. The possibility of an abiotic route to the 2′-deoxynucleotides provides a new perspective on the evolutionary origins of DNA. We also show that 2-aminothiazole is able to sequester, through reversible aminal formation, the important nucleotide precursors glycolaldehyde and glyceraldehyde in a stable, crystalline form.

INTRODUCTION

A plausible abiotic chemical route to the canonical nucleotides is a major goal in origins of life research.1–3 Recently, we demonstrated the first chemical steps toward a divergent pyrimidine and purine ribonucleotide synthesis.3 A one-pot multicomponent reaction was demonstrated to stereoselectively tether and consequently regiospeciﬁcally glycosylate purine precursors and masked pentose sugars, while concurrently furnishing known pyrimidine precursors (Scheme 1).

Although RNA has often been considered as a candidate for the ﬁrst biopolymer of life,1,2 extant biology utilizes two chemically distinct, but related, nucleotide polymers, RNA and DNA. DNA is usually viewed as a late evolutionary adaptation of earlier RNA-based life.4 However, it would not be possible to make DNA without deoxyribonucleotides, and yet in the absence of DNA there is no obvious reason for the evolution of the biochemical pathways for the synthesis of deoxyribonucleotides. Only after the emergence of DNA as an important cellular component would there have been a strong selective pressure favoring the emergence of biochemical pathways for the synthesis of deoxyribonucleotides from ribonucleotides. Although the advantages of DNA as a medium for information storage are clear, there has not until now been any reasonable hypothesis for how DNA could be ‘invented’ by primitive cells in the absence of pre-existing deoxyribonucleotides. In this paper, we propose a hypothesis for the prebiotic synthesis of 2′-deoxyribonucleotides.

Biochemically, 2′-deoxyribonucleotides are synthesized by chemo- and regio-speciﬁc reduction of ribonucleotides.5 In the absence of the complex and energetically costly enzymatic resources required to regiospeciﬁcally deoxygenate ribonucleotides, abiological 2′-deoxygenation of nucleotides requires site-
specific ribo- or arabino-nucleoside deoxygenation. In light of the difficulty of regioselective nucleotide deoxygenation in the absence of sophisticated enzymatic control, it would seem that regiospecific positioning of a group/atom with the appropriate latent potential reactivity could form the basis of an alternative pathway to the 2′-deoxygenated nucleotides. In principle, a sufficiently chemoselective reactivity difference could be exploited to generate both ribonucleotides and 2′-deoxygenated nucleotides, simultaneously.

RESULTS AND DISCUSSION

Constitutional analysis of DNA, with respect to RNA, in light of the low dissociation energy of C–S bonds,6 suggests that regiospecific positioning of sulfur at C2′ could result in the chemical differentiation required for the divergent synthesis of DNA and RNA monomers from common precursors (Scheme 2).

Scheme 2. Proposed Multicomponent Deoxyribonucleotide Syntheses

Oxazole 1—a key ribonucleotide precursor—is derived from cyanamide 2 and glycolaldehyde 3.1,3,7 The C2-carbon atom of aldehyde 3 is formally delivered to a ribonucleotide as the C2′ sugar carbon. It is of note that aldehyde 3 is also the aldehyde precursor of serine 4, a proteinogenic amino acid, via Strecker synthesis with ammonium cyanide (Scheme 3).8 The constitutional similarity of serine 4 and cysteine 5, and their aldehyde precursors 3 and β-mercaptoacetaldehyde 6, suggests that both 3 and 6 must be considered within our exploration of prebiotic azole synthesis.9 Specifically, we propose that the divergent reactivity of glycolaldehyde 3 and β-mercaptoacetaldehyde 6 with cyanamide 2 could lead, via 2-aminooxazole 1 and 2-aminothiazole 7, to ribonucleotides1,3 and 2′-deoxygenated nucleotides, by desulfurization of C2′ (Scheme 2).10,11

The importance of 2-aminothiazole 7, and its derivatives, has been recognized in medicinal chemistry,12 but though the synthesis of thiazole 7 in water has been reported,13 the potential relevance of thiazole 7 to the origins of life remains unexplored. The formation of oxazole 1 from cyanamide 2 and glycolaldehyde 3 at neutral pH requires stoichiometric (or excess) cyanamide 2 and general acid–base catalysis.1 Aldehyde 6, upon imine formation with cyanamide 2, has, like 3, the potential to undergo 5-exo-dig cyclization to furnish imine 14 (or hemiaminal 15) as shown in Scheme 3. Furthermore, we expected the reaction to proceed more rapidly at neutral pH and be less prone to stall at intermediates en route to thiazole 7 due to the increased nucleophilicity of sulfur and greater aromaticity of thiazoles with respect to oxazoles.

To test this supposition, β-mercaptoacetaldehyde 6 (8.5 mM) and 2 (15, 26, or 50 mM) were incubated in D2O at pD 7 and 20 °C and a rapid (ca. 10 min) quantitative conversion of aldehyde 6 to hemiaminal 15 was observed (see the Supporting Information Figure S1). Over the course of several hours, 15 dehydrated to furnish thiazole 7 (see Supporting Information Figures S2 and S3).14 It is of note that excess aldehyde 6 did not lead to further reaction of thiazole 7. Moreover, at elevated concentration (>150 mM) 2-aminothiazole 7 was observed to directly crystallize from water, and a crystal structure of thiazole 7 is shown in Scheme 3.15,16 Next we investigated the concomitant synthesis of oxazole 1 and thiazole 7. In the presence of excess cyanamide 2 (2.2–5 equiv), rapid and stoichiometric conversion of glycolaldehyde 3 (1 equiv) and β-mercaptoacetaldehyde 6 (1 equiv) to hemiaminals 11 and 15, respectively, was observed in water. At high pH/pD (pH/pD > 9) specific base catalysis or in 1 M phosphate solution (pH/pD 6.5–7.5) general acid–base catalysis furnish oxazole 1 and thiazole 7 over the course of 24–36 h at ambient temperature (Scheme 3; see Supporting Information Figure S8–S10).13 At neutral pD, selective dehydration of 15 to liberate thiazole 7 was observed. Limiting cyanamide 2 (2/3/6 ∼ 0.8:1:1) also resulted in the specific generation of thiazole 7, even in 1 M phosphate solution (pH 7; See Supporting Information Figure S9). It is likely that the rapid 5-exo-dig cyclization of imine 12/hemiaminal 13 leads to the selective sequestration of cyanamide 2 in thiazole 7.

We did not observe significant homo- and hetero-aldol reaction during the reaction of glycolaldehyde 3 and mercaptoacetaldehyde 6 with either excess or limiting cyanamide 2 between ambient temperature and 60 °C in 1 M phosphate solution at pH 7. Interestingly, we also found that, like oxazole 1, thiazole 7 is easily sublimed (see Supporting Information Figure S4), suggesting possible geo-chemical routes to the distribution, separation, and purification of thiazole 7 under abiotically plausible conditions. Sublimation from a mixture of oxazole 1 and thiazole 7 resulted in the isolation of both azoles as colorless crystalline solids. Partial sublimation of a mixture of oxazole 1 and thiazole 7 (1:1) resulted in the isolation 4:1–2:1 oxazole 1/thiazole 7, at 50 °C between 1 and 24 h at ambient pressure, indicating the more rapid but comparable sublimation rate of oxazole 1 relative to thiazole 7. Complete sublimation was observed after 2 d at 60 °C to afford ∼1:1 mixture of crystalline azoles at ambient pressure.
We then examined the multicomponent reaction of thiazole 7, 4-amino-imidazole-5-carboxamide 16, and thiazole and a series of aldehydes. Initially model aliphatic aldehydes were investigated to avoid issues of annulation control. Equimolar quantities of thiazole 7 and aliphatic aldehydes (including formaldehyde, acetaldehyde, propionaldehyde, and cyclobutaldehyde) with excess amino-imidazole 16 (1.2−3 equiv) furnished high yields (75−90%) of thiazolines 17, 18, 19, and 20, respectively, at pH 5. The multicomponent reactions were found to be robust at multi-gram scale and comparable diastereoselectivity was observed in the reaction of aliphatic aldehydes with oxazole 1 or thiazole 7 and amino-imidazole 16 (via imine 21 R = alkyl; d.r.= 2.5:1 thrice/erythro).

Compounds 17, threeo-18 and erythro-18, and the major isomers of 19 and 20 were isolated by fractional crystallization, and the crystal structures are shown in Scheme 4.19

To continue our investigation into the potential of multicomponent assembly of nucleotides, we studied the participation of thiazole 7 in carbon−carbon bond-forming reaction with α-hydroxyldehydes and imines derived from α-hydroxyldehydes in water.20 Incubation of equimolar aldehyde 3, 2-aminothiazole 7, and amino-imidazole 16 for 5 d in water at pH 4−6 at 20 °C gave a high yield of C7-thiazoline 23 (80% isolated yield) as 1:2:1 erythro/threo mixture of diastereomeric products. Surprisingly, this signified that the facial selectivity of thiazole 7 is significantly relaxed with respect to oxazole 1 or imine 21 (R = CH₃OH). However, both erythro-23 and threeo-23 were formed with complete annulation control (>99%, see ¹H NMR spectrum shown in Supporting Information Figure S28; single crystal X-ray structures of erythro-23 and threeo-23 are shown in Scheme 5).22

Upon investigation of the parallel MCR with the C3-synthon glyceraldehyde 22, comparable efficiency in the synthesis of C7-thiazoline 24 (70% isolated yield) was observed. Again, exclusive 6-exo-trig annulation is observed; however, in this case, 24 is furnished with significant diastereoselectivity (d.r.= 9:5:1:1, see Scheme 5 and Supporting Information Figure S35). Isolation of the two major isomers, crystallization, and X-ray diffraction proved their xyllo- and lyxo-configuration.22 The diastereoselectivity obtained, upon carbon−carbon bond formation is proposed to result from intramolecular hydrogen-bonding and concomitant steric impedance of thiazole addition to the Si-face of imine 21 (see Supporting Information Scheme S1).3,23 Although the facial selectivity of the nucleophile 7 has been relaxed with respect to the comparable reaction of 1, the facial selectivity of the glyceraldehyde-imine 21 (R = CH₃OH) is retained.3 This is important with respect to purine nucleotide synthesis by 3′-purine tethers, which necessarily requires N3′-displacement and C3′-inversion. Furthermore, as well as regiospecific tethering of 16 and consequent delivery of N1 to the anomic carbon atom, reaction of thiazole 7 with imine 21 regiospecifically delivers sulfur to C2′, which is of significant interest with regards to the potential synthesis of 2′-deoxynucleotides by C2′-desulfurization.

The multicomponent tricyclic thiazoline products 17−24 were uncontaminated with homoaldolization and homomannich byproducts. Additionally, oxazole 1 is known to react efficiently with α-hydroxyldehydes in a pH-dependent bimolecular domino-reaction.3 In striking contrast, no significant carbon−carbon bond formation was observed by ¹H NMR analysis upon incubation of thiazole 7 with either glycolaldehyde 3 or rac-glyceraldehyde 22 in D₂O (pD 3−9, 4−100 °C). Instead, over a period of 4 days at pD 7, crystal formation was observed from an aqueous solution containing 3 (100 mM) and thiazole 7 (100 mM) at 20 °C.24 Single crystal X-ray diffraction of a crystal isolated directly from the reaction supernatant proved that the structure was aminal 25 (Scheme 6).26

Though NMR studies confirmed that rac-22 formed hemiaminal 26 and aminal 27 with thiazole 7 in D₂O, no aminal crystallization was observed in solutions of rac-22 and 7, even at significantly elevated concentrations (500 mM to 2M) or with substoichiometric quantities of rac-22. To test whether
the change in solubility properties was due to the increased number of hydroxyl moieties or perhaps due to the differential solubility and packing properties of rac-27 with respect to the achiral aminal 25, we investigated the solubility properties of the aminal structures formed in a mixture of d-glyceraldehyde 22 and thiazole 7. Intriguingly, we found that at 100 mM concentration, d-27 now crystallized from water at 20 °C. NMR analysis showed a 1:2 ratio of 22/7 and recrystallization (10:1 H2O/EtOH) allowed for crystallographic confirmation of the aminal structure (Scheme 6).25

**SUMMARY**

In summary, our results suggest that in a sulfur-rich prebiotic environment, oxazole 1 and thiazole 7 could be formed together under the same conditions, as long as both glycolaldehyde 3 and its thiol analogue, β-mercapto-acetaldehyde 6, were present along with cyanamide 2. In such mixed reactions, 2-aminothiazole 7 forms first, followed by 2-aminooxazole 1. The absence of any interference between oxazole and thiazole synthesis shows that these potential precursors of deoxyxynucleotides and ribonucleotides could indeed be formed at the same time in the same conditions. We have also found that two abiotically plausible methods for the purification of oxazole 1, namely, crystallization and sublimation, also apply to thiazole 7. However, the chemical reactivities of oxazole 1 and thiazole 7 are quite different, suggesting that the subsequent pathways would diverge.

The most striking difference between the reactivity of oxazole 1 and thiazole 7 is the failure of the thiazole to react with aldehydes to form the thiazoline precursors of the (thio)pyrimidine nucleotides. The most likely explanation for this difference is the increased aromaticity of the thiazole relative to the oxazole, and thus decreased nucleophilicity of C5. We are currently investigating an alternative synthesis of 2′-deoxyribopyrimidine nucleotides mediated by C2′-inversion of anhydronucleotide 28 (Scheme 1).10,11 However, during the course of efforts to observe two-component reactivity of thiazole 7 with aldehydes, we observed instead the crystallization of glycolaldehyde 3 and homochiral (but not racemic) glyceraldehyde 22 (two known ribonucleotide precursors)1,3 as aminals directly from water. The facile reversibility of aminal formation suggests that the chemical sequestration and protection of C2′ and C3′-aldehyde nucleotide synthons by the change in solubility properties was due to the increased number of hydroxyl moieties or perhaps due to the differential solubility and packing properties of rac-27 with respect to the achiral aminal 25, we investigated the solubility properties of the aminal structures formed in a mixture of d-glyceraldehyde 22 and thiazole 7. Intriguingly, we found that at 100 mM concentration, d-27 now crystallized from water at 20 °C. NMR analysis showed a 1:2 ratio of 22/7 and recrystallization (10:1 H2O/EtOH) allowed for crystallographic confirmation of the aminal structure (Scheme 6).25

Until now made it difficult to conceive of means for accumulating large reservoirs of these essential starting materials in a prebiotic context.

By exploring the multicomponent reactivity of thiazole 7, together with aldehydes and the purine precursor aminoimidazole, we have demonstrated a selective and high yielding 3-component carbon—carbon bond-forming reaction that chemospecifically furnishes masked 2′-thiosugars regiospecifically tethered to an amino-imidazole purine-precursor. The facial selectivity of thiazole 7, upon addition to α-hydroxymethylaldehyde-imines, is quite relaxed (relative to oxazole 1), resulting in low diastereoselectivity at C2′; however, this is inconsequential with regards to the proposed deoxynucleotide synthesis because C2′ would be rendered achiral following desulfurization. In contrast, we observed very high diastereoselectivity of nucleophilic addition of thiazole 7 to imine 21, generating almost exclusively the lyxo and xylo isomers with controlled N9–C1′ annulation. The conversion of the resulting C2-thiazolines to 2′-deoxyribo-purine nucleotides requires three additional steps: C3′-inversion to release the purine-precursor from its tethered attachment to C3′, completion of the purine heterocycle, and chemospecific desulfurization of C2′.23 These steps are currently under investigation in our laboratories as part of our continuing studies aimed at the abiotic synthesis of 2′-deoxyribonucleotides.29

Finally, it is of note that if life emerged in an environment containing both ribo- and 2′-deoxyribonucleotides, it is likely that the first biopolymers were heterogeneous in composition, that is, composed of a mixture of ribo- and deoxyribonucleotides, and perhaps other compatible nucleotides. Recent work has shown that functional biopolymers, in the form of aptamers with highly specific molecular recognition properties, can be derived by in vitro evolution from libraries of polynucleotides composed of randomly interspersed ribo- and deoxyribo-nucleotides.29 Thus, chimeric RNA/DNA polymers may have been sufficient for the emergence of life. The primordial biochemical exploitation of a mixed RNA and DNA genetic system could eliminate the requirement for a genetic takeover (of RNA by DNA), and would arguably result in a simplification of the transition from chemistry to biology. However, such mixed-composition polymers have neither the advantageous stability of DNA, nor the optimal functional characteristics of RNA. Thus, RNA and DNA may have emerged as early and contemporaneous specializations of a primitive mixed biopolymer.

**EXPERIMENTAL SECTION**

**General Methods.** Reagents and solvents were purchased from Sigma-Aldrich, TCI America, Frontier Scientific, or Cambridge Isotope Laboratories. Flash column chromatography was carried out on Merck 9385 silica gel 60 (230–400 mesh). NMR spectroscopy was carried out on a Varian NMR spectrometer (Oxford AS-400) or a Bruker NMR spectrometer (AvanceIII 600) operating at 25 °C probe temperature (unless otherwise specified). When possible, the chemical shift of the corresponding solvent was used as a reference. Chemical shift values are reported in parts per million (ppm) and J-couplings are recorded in hertz (Hz). Electrospray mass spectrometry was recorded on a Bruker Daltonics Esquire 6000 ESI-MS. High resolution mass spectrometry was carried out on a Waters Q-ToF micro LC/MS/MS system. Infrared spectra were recorded as manually pressed KBr discs on a PerkinElmer spectrum 100 series FT-IR spectrometer. Single crystal X-ray spectroscopy was carried out with a Bruker Apex II CCD diffractometer (Mo Kα radiation, λ = 0.71073 Å), equipped with an Oxford Cryosystems nitrogen flow apparatus, at 100 K. Data
integration down to 0.76 Å was resolved using SAINT V7.46 A (Bruker diffractometer, 2009) with reflection spot size optimization. Absorption corrections were made with the program SADABS (Bruker diffractometer, 2009). The structure was solved by the direct methods procedure and refined by least-squares methods against F² by using SHELX-97 and SHELXL-97 (Sheldrick, 2008). Non-hydrogen atoms were refined anisotropically.

2-Aminothiazole (7). Mercaptoacetaldehyde 6 and cyanamide 2 (0.5–1.5 equiv) were dissolved in H₂O, D₂O or phosphate buffer (1 mL) at pH/pD 7.0. Reaction progress was monitored by the ¹H NMR spectroscopy (see the Supporting Information, Figure S1). An initial concentration of 100 mM led to direct crystallization of 7 (60% isolated yield of crystals). Alternatively, upon complete conversion (by NMR spectroscopy), the solution was concentrated and 7 recovered by sublimation. Compound 7 (1–5 g) was placed in a covered beaker or Erlenmeyer flask and warmed from beneath to 40 °C. Sublimate was collected on flask/beaker sides and cover at ambient temperature (see Figure S4) or on a water-cooled coldfinger (5 °C). Crystalline 7 was isolated. IR (solid, cm⁻¹): 3407, 3285, 3120, 3086, 1622, 1518, 1488. ¹H NMR (400 MHz; DMSO): δH 6.00 (d, J = 3.7, 1H, H−(C4)), 6.84 (s, 2H, NH₂), 6.51 (d, J = 3.7, 1H, H−(C−)), 13C NMR (101 MHz; DMSO): δC 167.7, 159.3, 129.7, 115.1, 79.1, 54.8, 46.7, 22.7. HRMS ESI [M + Na+] calc. for C₇H₈N₃O₃SNa 275.0691; obs. 275.0685. Cambridge Crystallographic Data Centre deposition number 864227.

One-Pot Synthesis of 2-Aminothiazole (7) and 2-Amino-thiophene (8). Glyceraldehyde Bis-thiazole Aminal (50–250 mM, 1 eq) was added to a solution of 7 (1.1 equiv) and 5-aminoimidazole+carboxamide (1.5–3.0 equiv) in H₂O at pH 5.0 ± 0.2. The reaction was stirred under an argon atmosphere until no residual aldehyde was detectable following hydropolysis and analysis by ¹H NMR. The reaction was lyophilized and purified by silica gel column chromatography (CH₂Cl₂:MeOH, 10:0 to 6:4). The major product(s) were concentrated and recrystallized from ethanol/water or dichloromethane/methanol mixtures.

**General Procedure for Isolation of Bis-thiazole Aminal Crystals.** Aldehyde (0.49–1.0 equiv) was added to a solution of 7 in H₂O at pH 7 ± 0.2. The solution was then incubated at room temperature for 5–10 d; the resultant solid was then isolated by filtration or trituration and washed twice with ethanol and once with ice cold water. 2-Glyceraldehyde 22/thiazole 7 crystals were then recrystallized from aqueous ethanol (10:1 H₂O/ErOH).

**Glyceraldehyde Bis-thiazole Aminal 17.** IR (solid, cm⁻¹): 3213 (br), 2966, 2934, 2846, 1589, 1539, 1497. Cambridge Crystallographic Data Centre deposition number 864235.

**Glyceraldehyde Bis-thiazole Aminal 19.** IR (solid, cm⁻¹): 3326 (br), 3244, 3113, 2948, 2933, 2901, 2882, 1536, 1489. Cambridge Crystallographic Data Centre deposition number 864236.

**General Procedure for Multicomponent Reaction.** Aldehyde (50–250 mM, 1 eq) was added to a solution of 7 (1.1 equiv) and 5-aminoimidazole+carboxamide (1.5–3.0 equiv) in H₂O at pH 5.0 ± 0.2. The reaction was stirred under an argon atmosphere until no residual aldehyde was detectable following hydropolysis and analysis by ¹H NMR. The reaction was lyophilized and purified by silica gel column chromatography (CH₂Cl₂:MeOH, 10:0 to 6:4). The major product(s) were concentrated and recrystallized from ethanol/water or dichloromethane/methanol mixtures.
HRMS ESI [M+H]+ calc. for C$_{10}$H$_{14}$N$_{6}$O$_3$SNa 321.0746; obs. 321.0745.

$^{50.5}$HRMS ESI [M + Na+] calc. for C$_{10}$H$_{14}$N$_{6}$O$_3$SNa 321.0746; obs. 321.0742. Cambridge Crystallographic Data Centre deposition number 864234.

**xylo-24:** $R_2 = 0.20$ (CH$_3$Cl/MeOH 7:3). IR (solid, cm$^{-1}$): 3344, 3106, 2971, 2843, 1627, 1533. $^{1}$H NMR (400 MHz; DMSO): $\delta$ = 7.26 (s, 1H, Ar), 5.94 (d, $J = 5.8$, 1H, H$^\prime$), 4.21 (t, $J = 7.6$, 1H, H$^\prime$), 3.73 (m, 1H, H$^\prime$). $^{13}$C NMR (100 MHz; DMSO): 166.0, 147.1, 141.2, 128.6, 73.8, 73.5, 53.3. HRMS ESI [M + Na+] calc. for C$_{10}$H$_{14}$N$_{6}$O$_3$SNa 321.0746; obs. 321.0742. Cambridge Crystallographic Data Centre deposition number 864234.

**ASSOCIATED CONTENT**

[Supporting Information](http://pubs.acs.org)

**AUTHOR INFORMATION**

Corresponding Author
matthew.powner@ucl.ac.uk

Notes
The authors declare no competing financial interest.

**ACKNOWLEDGMENTS**

M.W.P. was supported by a fellowship from the Harvard Orgins Initiative, the Howard Hughes Medical Institute and University College London. This research was funded in part by Grant CHE-0809413 from the NSF. J.W.S. is an Investigator of the Howard Hughes Medical Institute. The authors thank J.C. Blain, Dr. R.B. Moreno, Dr. A. Ricardo, and Prof. J.D. Sutherland for helpful discussions and The Mass Spectrometry Facility at the FAS Center for Systems Biology for mass spectrometry support.

**REFERENCES**


5. (a) Schirm, A.; Schlemp, R. M.; Calvin, M. Science 1965, 147, 149.


10. (a) Carbone, R. B.; Brutel, R.; Beutel, R. H. 1957, US PAT. 2,605,004. (b) Significant H−(C4) exchange was observed in D$_2$O.


12. Cambridge Crystallographic Data Centre deposition number for 7 is 864225.

13. (a) 16 is a product of HCN-oligomerization and known nucleobase precursor.18 Notably, HCN also is required for Streecker synthesis.18 (b) Criegee, M. Biochim. Biophys. Res. Commun. 1960, 2, 407.

14. (a) Brand, K. U.; Rees, M. W.; Markham, R. Nature 1963, 199, 219. (b) Cambridge Crystallographic Data Centre deposition numbers for 17, ethryro-18, threreo-18, and 19 and 20 are 864226, 864227, 864228, 864229, and 864230, respectively.

15. (a) Proposed prebiotic nucleotide syntheses must be evaluated for their potential to yield alternative structures capable of base-pairing. Constitutionally, 3 (C$_3$) and 22 (C$_2$) can react with azoles (C$_3$ sugar-synths) to furnish threose nucleic acid (TNA; C$_3$sugar) and DNA/RNA (C$_2$sugar) precursors, respectively.1


17. Cambridge Crystallographic Data Centre deposition numbers for ethryro-23, threreo-23, xylo-24, and xylo-24 are 864231, 864232, 864233, and 864234, respectively.


(28) N.B. C₄-thiazoline desulfurization would generate monovalent nucleosides incapable of forming polymeric phosphodiester.