

# Selective prebiotic formation of RNA pyrimidine and DNA purine nucleosides

<https://doi.org/10.1038/s41586-020-2330-9>

Received: 12 December 2019

Accepted: 16 April 2020

Published online: 3 June 2020

 Check for updates

Jianfeng Xu<sup>1,6</sup>, Václav Chmela<sup>1,6</sup>, Nicholas J. Green<sup>1</sup>, David A. Russell<sup>1</sup>, Mikołaj J. Janicki<sup>2</sup>, Robert W. Góra<sup>2</sup>, Rafał Szabla<sup>3,4</sup>, Andrew D. Bond<sup>5</sup> & John D. Sutherland<sup>1</sup>✉

The nature of the first genetic polymer is the subject of major debate<sup>1</sup>. Although the ‘RNA world’ theory suggests that RNA was the first replicable information carrier of the prebiotic era—that is, prior to the dawn of life<sup>2,3</sup>—other evidence implies that life may have started with a heterogeneous nucleic acid genetic system that included both RNA and DNA<sup>4</sup>. Such a theory streamlines the eventual ‘genetic takeover’ of homogeneous DNA from RNA as the principal information-storage molecule, but requires a selective abiotic synthesis of both RNA and DNA building blocks in the same local primordial geochemical scenario. Here we demonstrate a high-yielding, completely stereo-, regio- and furanosyl-selective prebiotic synthesis of the purine deoxyribonucleosides: deoxyadenosine and deoxyinosine. Our synthesis uses key intermediates in the prebiotic synthesis of the canonical pyrimidine ribonucleosides (cytidine and uridine), and we show that, once generated, the pyrimidines persist throughout the synthesis of the purine deoxyribonucleosides, leading to a mixture of deoxyadenosine, deoxyinosine, cytidine and uridine. These results support the notion that purine deoxyribonucleosides and pyrimidine ribonucleosides may have coexisted before the emergence of life<sup>5</sup>.

Considerable progress in the prebiotic synthesis of the pyrimidine ribonucleosides of RNA: cytidine (C) **1** and uridine (U) **2**, and their 2-thio derivatives: **3** and **4**<sup>6,7</sup>, together with recent advances in non-enzymatic RNA replication<sup>8–10</sup> have given credence to the ‘RNA world’ theory. Progress towards the abiotic synthesis of purine nucleosides has been made, but only using routes that use chemically and enantiomerically pure sugars as starting materials<sup>11–15</sup>, which were probably not found on the primordial earth. Additionally, no prebiotically plausible route has been shown to provide a mixture containing a competent set of nucleosides for information storage at the polymeric level.

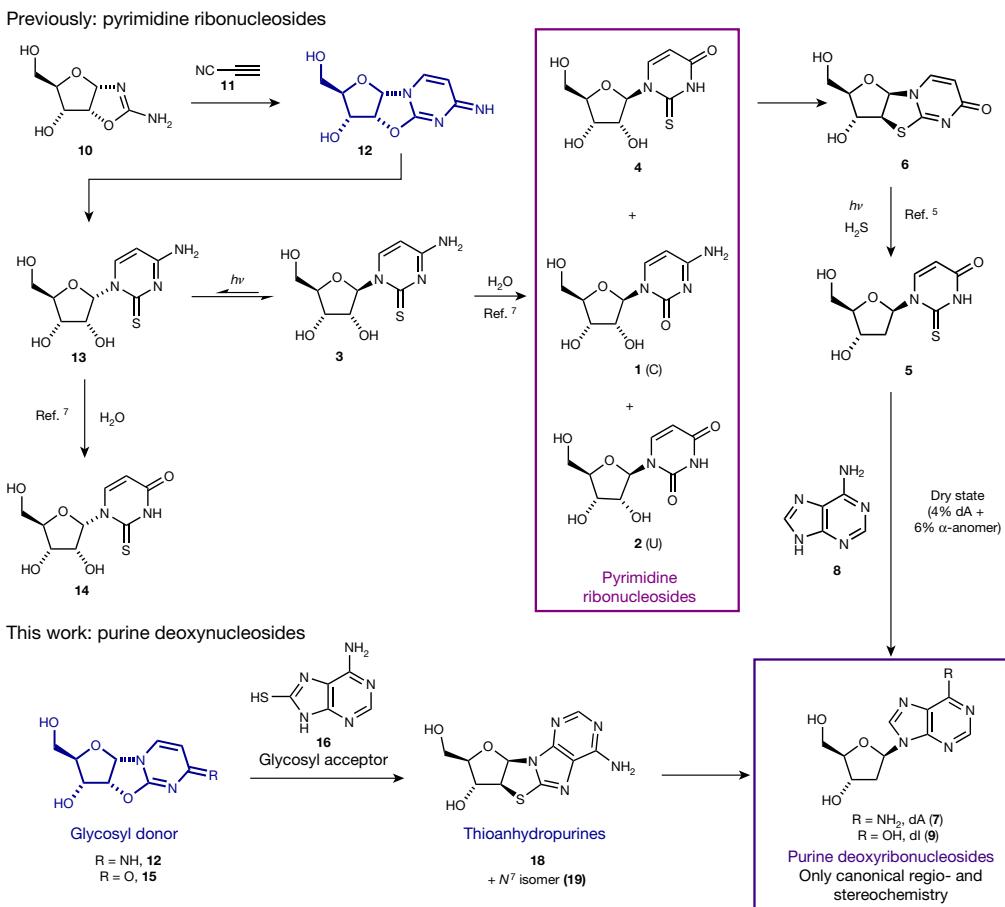
Extant biology, in contrast to the proposed RNA world theory, features DNA as the central information-carrying molecule. This discrepancy between the RNA world and modern biology requires a ‘genetic takeover’ that invokes the power of primitive biosynthetic machinery and natural selection operating over millions of years, ultimately resulting in an ancestral biosynthetic route to DNA<sup>16</sup>. The superior hydrolytic stability and replication fidelity<sup>17</sup> of DNA could have resulted in selection of primitive organisms capable of synthesizing DNA, and thus its rise to prominence, but the feasibility of this evolutionary process in a pre-DNA world is debated<sup>4</sup>. To circumvent this potentially problematic transition, an RNA/DNA world has been proposed, in which nascent biology had access to both RNA and DNA building blocks from the outset, without requiring elaborate biosynthesis<sup>18–20</sup>. In such a world, heterogeneous polymers would have initially been most common, but polymers with increased homogeneity—and hence properties closer to that of either RNA or DNA—would have been selected for over their mixed counterparts<sup>4</sup>.

For the RNA/DNA world to be plausible, an efficient prebiotic synthesis of DNA building blocks is required, and one that provides building blocks for both RNA and DNA in the same localized geochemical scenario is preferable. We recently demonstrated proof of this principle by showing that 2'-deoxy-2-thiouridine (**5**)—a non-canonical deoxynucleoside—can be synthesized from thioanhydrouridine (**6**)—an RNA derivative—by way of a prebiotically plausible hydrogen sulfide-mediated photoreduction<sup>5</sup>. Although this finding provides an important prebiotic link between RNA and DNA building blocks, the lability of **5** to hydrolysis may limit its phosphorylation and subsequent oligomerization<sup>21,22</sup>. Additionally, the synthesis of canonical deoxyadenosine (dA; **7**) from **5** and adenine (**8**) was low-yielding (4%), and generated a more abundant, undesired side product: the  $\alpha$ -anomer of **7** (6%). Using guidance from a geochemical scenario<sup>23</sup>, we now demonstrate a synthesis of purine deoxynucleosides that is based on prebiotically plausible reactions and substrates. We then evaluate our route at a systems level by enacting the synthesis on mixtures of materials that might have arisen in a primordial environment, culminating in the demonstration of multiple reaction sequences able to selectively furnish a mixture of U (**1**), C (**2**), dA (**7**) and deoxyinosine (dI, **9**) (Extended Data Fig. 1).

## Prebiotic route to purine deoxyribonucleosides

A route to purine nucleosides that diverges from a prebiotic RNA synthesis is attractive because it implies that the constituents of a set of nucleosides capable of storing information—pyrimidines and

<sup>1</sup>MRC Laboratory of Molecular Biology, Cambridge Biomedical Campus, Cambridge, UK. <sup>2</sup>Department of Physical and Quantum Chemistry, Faculty of Chemistry, Wrocław University of Science and Technology, Wrocław, Poland. <sup>3</sup>EaStCHEM, School of Chemistry, University of Edinburgh, Edinburgh, UK. <sup>4</sup>Institute of Physics, Polish Academy of Sciences, Warsaw, Poland. <sup>5</sup>Department of Chemistry, University of Cambridge, Cambridge, UK. <sup>6</sup>These authors contributed equally: Jianfeng Xu, Václav Chmela. <sup>✉</sup>e-mail: johns@mrc-lmb.cam.ac.uk



**Fig. 1 | Previous synthesis of RNA pyrimidine nucleosides C (1), U (2) and a deoxypyrimidine nucleoside (5), and the present work.** RAO (10) is a starting point in the network because it crystallizes in enantiopure form from minimally enantio-enriched solutions<sup>25,26</sup>. It can be elaborated via 12 and 3 to the pyrimidine nucleosides<sup>7</sup>. Although we have previously developed a low-yielding route to deoxyadenosine 7 (dA) from 6 via 5<sup>5</sup>, we recognized that 12 and 15 are ideal candidates for tethered glycosylation with 16. The products,

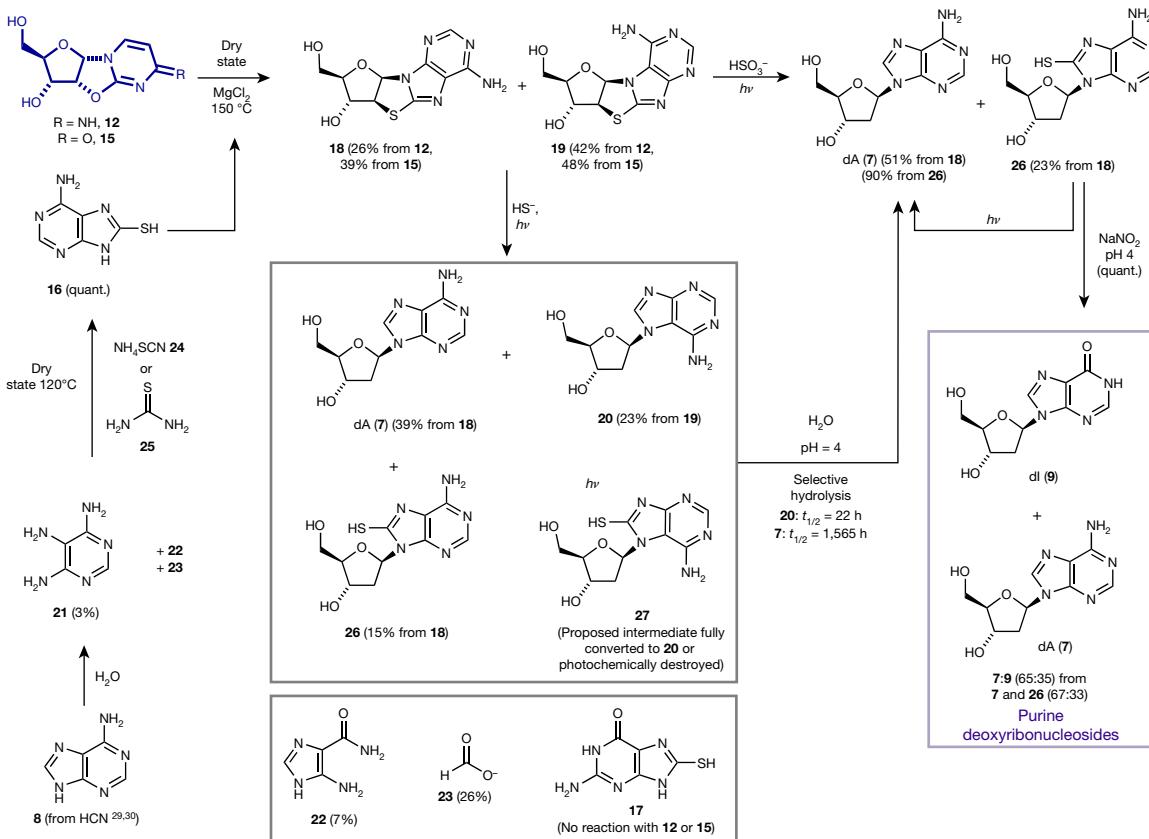
thioanhydropurines 18 and 19, are reduced photochemically in a similar way to 6, providing an efficient route to deoxynucleosides. Critically, once produced, pyrimidines 1(C) and 2(U) survive the sequence that produces purines 7 (dA) and 9 (dl), and we show that the four nucleosides 1(C), 2(U), 7 (dA) and 9 (dl) can be produced alongside one another. v, frequency of UV irradiation, centred at 254 nm.

purines—may have formed in the same location on a primordial Earth, and not necessarily brought together by environmental processes after their separate formation. To develop such a route, we evaluated intermediates in the prebiotic RNA pyrimidine nucleoside synthesis<sup>6,7</sup> as ribosyl donors (Fig. 1). The RNA synthesis proceeds from ribo-amino oxazoline (RAO, 10), which reacts with cyanoacetylene (11) to provide α-anhydrocytidine (12). Thiolytic cleavage of 12 in formamide produces α-2-thiocytidine (13), which undergoes efficient ultraviolet (UV)-mediated photoanomerisation to 2-thiocytidine (3), which hydrolyses to the canonical pyrimidines cytidine (1) and uridine (2), and the biologically important non-canonical pyrimidine (4). Alternatively, in the dark, 13 is hydrolysed to α-2-thiouridine (14)<sup>7</sup>. Although 14 was initially viewed as only a by-product that would be produced in the dark on the early Earth, it is readily cyclized to anhydrouridine (15) at 80 °C (63% yield in water or 89% yield in formamide; Extended Data Fig. 2). We recognized α-anhydropyrimidines (12 and 15) as ideal glycosyl donors for 1',2'-cis tethered glycosylation<sup>24</sup>. The sugar of 12 and 15 is fixed in its furanosyl form, and so the formation of pyranosyl nucleosides—one of the weaknesses of previous strategies—should be excluded. Additionally, the α-stereochemistries of C1' and C2' of 12 and 15 led us to expect transglycosylation to provide only β-anomers, the correct stereochemistry at C1' for all natural (deoxy)ribonucleosides. Finally, because 12 and 15 are ultimately derived from RAO

(10)—which crystallizes enantiopure from solutions of minimally enantioenriched carbohydrates or amino acids<sup>25,26</sup>—this route offered the as-yet-unmet potential to deliver enantio- and diastereomerically pure furanosyl-nucleosides by glycosylation.

Accordingly, we evaluated 8-mercaptopoadenine (16) and 8-mercaptoguanine (17) as potential nucleophiles to participate in trans-glycosylation with 12 and 15 (Fig. 2). Although 17 proved unreactive, 16 reacts with 12 and 15 at 150 °C in the dry state (Fig. 2) to provide two new β-configured nucleoside products in moderate yields (respectively, 14% and 16% from 15, trace amounts from 12). The minor product was determined to be N<sup>9</sup>-8,2'-anhydro-thioadenosine (18) by X-ray crystallography and <sup>1</sup>H NMR spiking experiments with a synthetic standard. The major product was inferred to be N<sup>7</sup>-8,2'-anhydro-thiadenosine (19)—the regioisomer of 18—by its subsequent conversion to 2'-deoxy-N<sup>7</sup>-adenosine (20). The presence of magnesium chloride in the reaction—presumably acting as a Lewis acid<sup>27</sup>—dramatically improved the yield of 18 and 19 to 39% and 48% respectively from 15 (combined yield 87%), and 26% and 42% respectively from 12 (combined yield 68%). Thus, in a prebiotic environment where 12 or 15, and 16 are brought together, perhaps by converging streams that then undergo evaporation, 18 and 19 could be readily generated, especially in the presence of magnesium ions<sup>28</sup>.

Any prebiotic synthesis requires a viable route to all reagents from plausible early-Earth feedstocks. We considered adenine (8) as a



**Fig. 2 | Prebiotic route to purine deoxyribonucleosides, 7 (dA) and 9 (dl).** The route starts with  $\alpha$ -anhydropurines **12** and **15**, which are intermediates in the RNA pyrimidine synthesis, and 8-mercaptopadenine (**16**), which is available from adenine (**8**) via hydrolysis and reaction with ammonium thiocyanate or thiourea. Dry-state tethered glycosylation of **16** and **12** or **15** provides thioanhydropurines **18** and **19**, which can be photochemically reduced by two routes. If bisulfite is the reductant, only  $N^{\text{o}}$ -configured

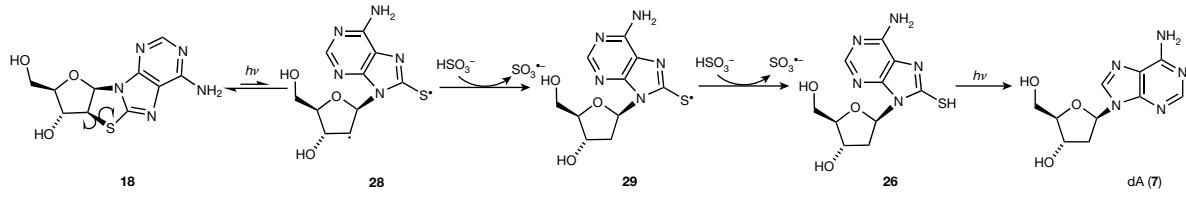
products **7** (dA) and **26** are formed. **26** can be converted to **7** (dA) by further irradiation or by nitrosation. If hydrosulfide is used as the reductant, both  $N^{\text{o}}$ -configured **7** (dA) and **26** as well as  $N^{\text{l}}$ -configured **20** is formed. **20** has a half-time of hydrolysis nearly two orders of magnitude lower than that of **7** (dA) and so is selectively degraded. To generate **9** (dl) alongside **7** (dA), the products of either photoreduction are treated with nitrous acid at pH 4.

starting point for the provision of 8-mercaptopadenine (**16**) owing to its widely accepted prebiotic plausibility as a relatively stable pentamer of hydrogen cyanide<sup>29,30</sup>. Notably, despite the reactivity of related purines<sup>31</sup>, adenine did not react with elemental sulfur at temperatures up to 300 °C. However, adenine does undergo (slow) hydrolysis in aqueous media. Previous reports provided a half-time for hydrolysis of adenine of about one year at 100 °C, and identified (but did not quantify) 4,5,6-triaminopyrimidine (**21**) among the products of hydrolysis<sup>32</sup>. We reinvestigated this hydrolysis of **8** under conditions more suited to a laboratory timescale (138 °C, phosphate buffer pH 8), and at partial conversion after 10–12 d we confirmed the presence of TAP in yields of 2–3% (8–9% based on recovered adenine) (Fig. 2). Owing to the differential solubilities of adenine and TAP, the supernatants of adenine hydrolysis reactions are enriched in TAP after cooling. A typical supernatant contains 5-aminoimidazole-4-carboxamide (**22**), TAP (**21**) and **8** in a 4:2:1 ratio, and formate (**23**) as the only other major component (see Supplementary Figs. 1–5 for full details). We found that TAP (either commercially supplied or in the crude adenine hydrolysate) is converted to 8-mercaptopadenine (**16**) by heating in the dry state with either ammonium thiocyanate (**24**) or thiourea (**25**). **24** is an inevitable by-product of the photochemistry of hydrogen cyanide and hydrogen sulfide<sup>33</sup>, two precursors that were probably abundant on the primordial earth, and that are heavily implicated in the origin of life by our cyanosulfidic chemical network<sup>23</sup>. **25** has also been

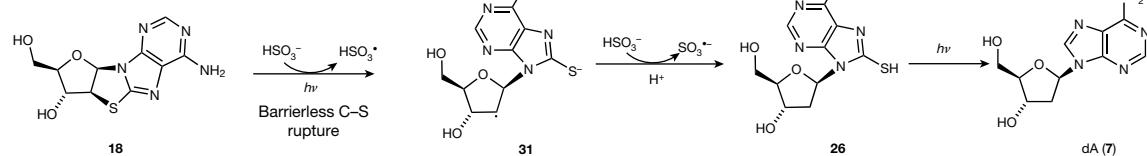
widely invoked as a prebiotically plausible reagent<sup>34</sup>. Thus, we envision that a primordial environment supplied with adenine and water would continuously generate TAP, which can be enriched in aqueous solution by moving down a thermal gradient. **24** can be mixed with TAP at any stage, and eventual evaporation and dry-state reaction leads to **16**. This method of accumulation of TAP also improves the plausibility of some intermediates of other prebiotic syntheses<sup>12</sup>.

With thioanhydropurine nucleosides (**18** and **19**) in hand, we moved on to evaluate their photoreduction chemistry to see if we might directly generate deoxyadenosine. Our previous synthesis of a deoxypyrimidine via a thioanhydropurine (**6**) (Fig. 1) proceeded by the reduction of a C–S to a C–H bond mediated by a hydrated electron, generated by UV irradiation of hydrosulfide<sup>5,33</sup>. **18** and **19** were separately subjected to UV irradiation at 254 nm in water with hydrogen sulfide ( $\text{H}_2\text{S}$ ) as the reductant (Fig. 2). In the photoreduction of **18**, the natural regio-isomer dA (**7**) was detected in 39% yield, along with 15% of 8-mercaptop-deoxyadenosine (**26**). **26** was demonstrated to be a competent intermediate in the reaction by desulfurization to give dA (**7**) either by UV irradiation<sup>35</sup> or treatment with nitrous acid, which is produced from common atmospheric gases, nitrogen and carbon dioxide<sup>36</sup>. Nucleobase loss was also apparent (**16** in 10% yield and **8** in 17% yield). The same reaction starting with **19** gave  $N^{\text{l}}$ -deoxyadenosine (**20**) in 23% yield with no other nucleoside products. Our proposed intermediate in this process, 8-mercato- $N^{\text{l}}$ -deoxyadenosine (**27**), is either fully converted to **20** or photochemically destroyed.

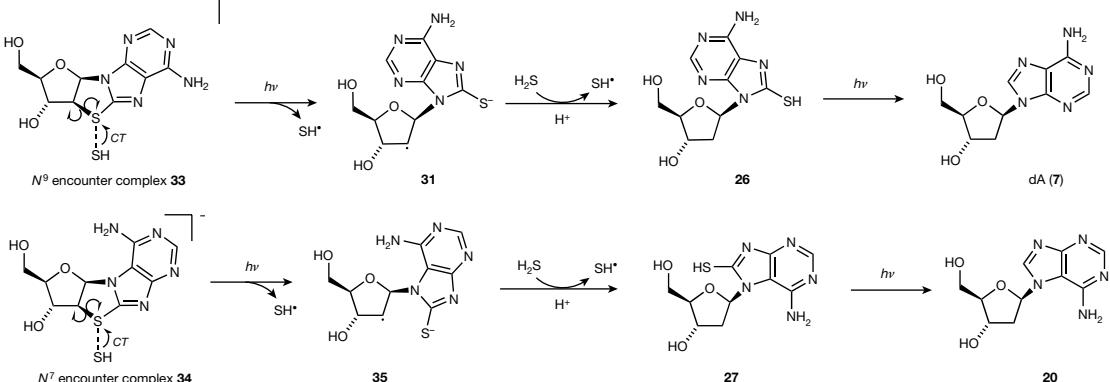
**a Photocleavage first**



**b Reduction first**



**c**



**Fig. 3 | Proposed mechanism of photoreduction of  $N^7$ -2',8,2'-anhydro-thioadenosine (**18**) and  $N^9$ -2',8,2'-anhydro-thioadenosine (**19**) nucleosides.** **a**, A potential mechanism involving bisulfite, proceeding with initial photoexcitation of the thioanhydronucleosides to **28**, followed by reduction of C2', sulfur and C8. Photoexcitation of the  $N^7$  isomer **19** to **30** leads to decomposition. **b**, A potential mechanism involving bisulfite, proceeding via intial reduction of ground-state thioanhydronucleosides, followed by

desulfurisation of **26**. Reduction of **19** gives **32**, which leads to decomposition.

**c**, A distinct mechanism involving reduction of thioanhydronucleoside–hydrosulfide encounter complexes, **33** and **34**, which both undergo charge transfer and concomitant C–S bond cleavage to produce **31** and **35**. **31** and **35** undergo reduction at C2' and desulfurisation to furnish **7** (dA) and **20**. CT, charge transfer; Decomp., decomposition.

Photoreduction was also carried out on a mixture of **18** and **19** compatible with our synthesis by tethered transglycosylation. The ratio of  $N^9:N^7$  regioisomers was increased from 38:62 of **18:19** in the starting mixture to 56:44 of **7:20** after photoreduction (31% yield for **7**, 17% yield for **20**), indicating, compared to  $N^7$  isomers, an enhanced stability of intermediates or products bearing the natural  $N^9$  glycosidic linkage. Replacing hydrosulfide as the electron donor with bisulfite ( $\text{HSO}_3^-$ , pH 7)<sup>37</sup>, which is readily formed by the dissolution of atmospheric  $\text{SO}_2$  in water<sup>38</sup>, improved both the yield and selectivity of the photoreduction. Photoreduction with bisulfite of **18** alone provided dA (**7**) in 51% yield and **26** in 23% yield, whereas a similar reaction with the  $N^7$ -regio-isomer (**19**) led only to its photochemical destruction. Photoreduction of a mixture of **18** and **19** with bisulfite led only to  $N^9$ -linked products: dA (**7**) and **26** in 44% and 18% yield, respectively (Extended Data Fig. 3). Separate experiments probing the stability of starting materials and

products under the reaction conditions indicated that the relative stabilities of intermediates are the cause of this selectivity. This strikingly selective destruction is highly suggestive of a potential mechanism by which primordial nucleosides may have been restricted to a near-canonical set<sup>39,40</sup>. We found further evidence for such restriction in the hydrolysis rates of the  $N^9$  and  $N^7$  isomers of deoxyadenosine. In acetate buffer (pH 4, 25 °C), the natural isomer dA (**7**) is more than 70 times more stable than **20** (half-lives of 1,565 and 22 h, respectively), which is consistent with the reported difference in stabilities towards acid hydrolysis between the corresponding isomers of adenosine<sup>41,42</sup>.

### Photoreduction mechanism

To provide a mechanistic rationale for the observed photochemical selectivity, we performed quantum chemical calculations using density

functional theory and algebraic diagrammatic construction to the second order (ADC(2)) methods<sup>43,44</sup>. These calculations revealed, in the case of bisulfite, two possible competing mechanisms that explain the difference in reactivity of the two regioisomers. **18** and **19** can both undergo photoexcitation, but generate dissimilar biradical species (Fig. 3a). Photoexcitation of **18** leads to rupture of the C2'-S bond on the surface of the lowest excited singlet ( $S_1$ ) state, generating biradical **28** (Fig. 3a,  $N^9$ ; Extended Data Fig. 4a). Reduction of this species by intermolecular hydrogen atom transfer (HAT) or proton-coupled electron transfer (PCET) probably leads to the C2'-reduced species **29**, and ultimately—via a second HAT or PCET and subsequent photolysis of the C8-S bond of **26**<sup>35</sup>—to dA (**7**) (Fig. 3a,  $N^9$ ). By contrast, photoexcitation of **19** leads to N7-C8 bond rupture through the  $S_1/S_0$  state crossing (Fig. 3a,  $N^9$ ; Extended Data Fig. 4b), generating **30**, which probably undergoes decomposition without C2'-S reduction. Bisulfite is well-known to provide a hydrated electron upon irradiation<sup>45</sup>, and so a second possibility is the reduction by hydrated electrons of **18** and **19** in the ground state. Again, calculations suggest different fates of **18** and **19** upon reduction. Reduction of **18** is predicted to proceed with concomitant barrierless C2'-S bond rupture to give the radical anion intermediate **31** (Fig. 3b,  $N^9$ ; Extended Data Fig. 5), whereas reduction of **19** is predicted to lead to formation of a C8, N9 radical anion (**32**) which also probably undergoes decomposition instead of C2' reduction (Fig. 3b  $N^9$ ; Extended Data Fig. 5). In the absence of any reducing agent, both **18** and **19** undergo (equally) slow photochemical decomposition, presumably via the calculated biradical structures **28** and **30**, but in the presence of bisulfite, reduction of the ground state or photochemically generated intermediates results in remarkably different fates.

The successful reduction of **19** alongside **18** when using hydrosulfide as the reducing agent is explained by a distinct mechanism. Calculations located stable encounter complexes, **33** and **34**, between HS<sup>-</sup> and thioanhydronucleosides **18** and **19**, respectively (Fig. 3c, Extended Data Fig. 4c, d). This interaction is predominantly stabilized by electrostatic and dispersion interactions and our interaction energy decomposition demonstrates its stability in aqueous solution (see Supplementary Information for details). Similar S-S interactions were recently identified in intramolecular complexes and classified as chalcogen bonds<sup>46</sup>. Such an encounter complex facilitates charge transfer from the hydrosulfide anion to the thioanhydropurine fragment almost immediately after UV absorption by the complex to the  $S_1$  state. Subsequent relaxation on the  $S_1$  surface enables practically barrierless C2'-S bond breaking completed by a peaked  $S_1/S_0$  state crossing for both intermediates **31** and **35**, thus facilitating C2'-S reduction of both **18** and **19** (Extended Data Fig. 4c, d). The products of this photochemical transformation, **26** and **27**, may further undergo photochemical sulfur cleavage through the mechanism described previously<sup>35</sup> (Fig. 3c). Thus, a HS<sup>-</sup> thioanhydropurine encounter complex facilitates C-S bond cleavage and partially protects  $N^7$  isomer **19** from the photodestruction observed in the presence of bisulfite. This finding not only explains the distinctive outcomes of photoreduction between the two reducing agents, but also points towards a potentially important stabilizing role for hydrosulfide in prebiotic chemistry and photochemistry in general.

## Prebiotic route to a purine/pyrimidine genetic system

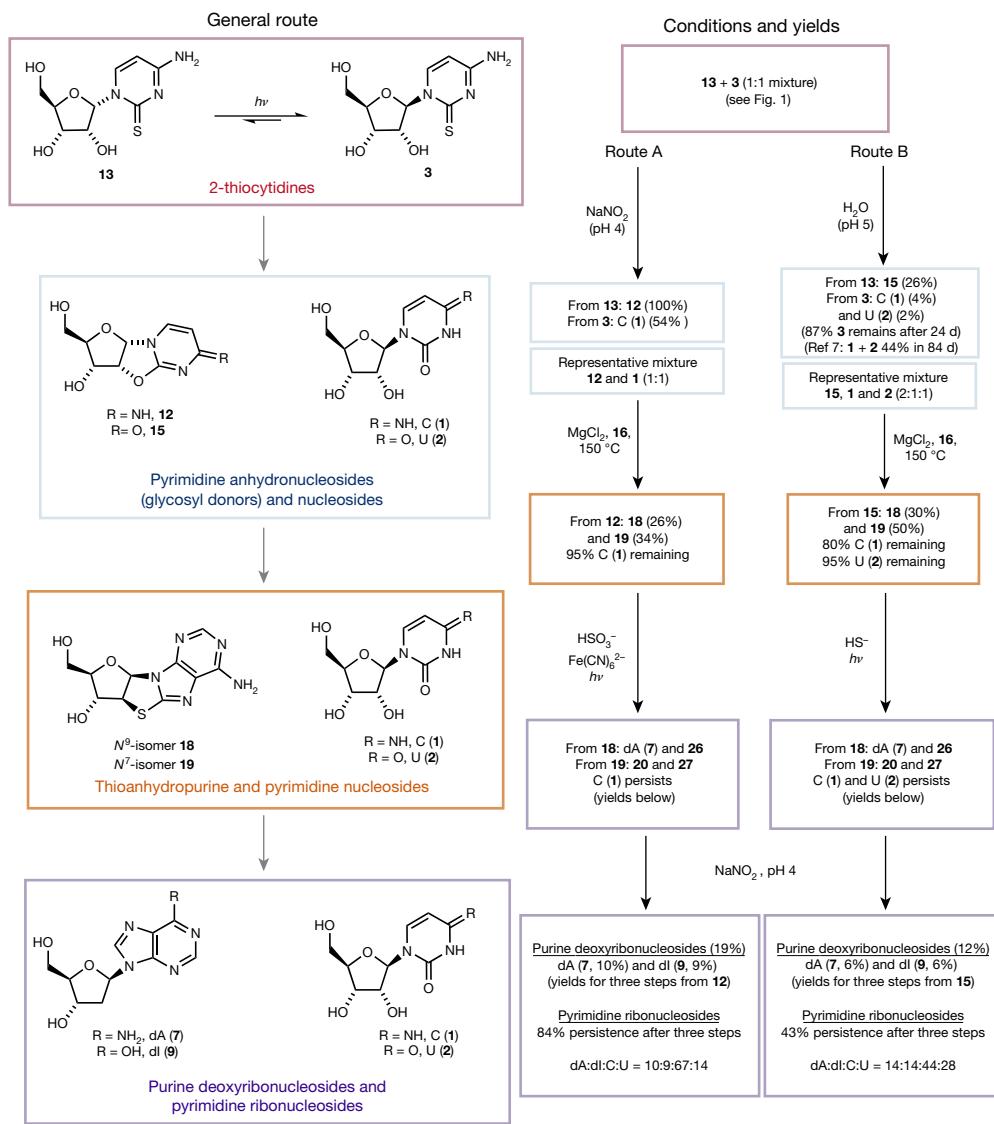
Although our attempts to glycosylate 8-mercaptoguanine (**17**) to provide thioanhydroguanosine (and ultimately deoxyguanosine) failed, the triple selectivity and high yield of our route to deoxyadenosine combined with recent results<sup>47</sup> suggest a possible alternative genetic alphabet that does not include (deoxy)guanosine. Guanosine is yet to be accounted for in a plausible prebiotic synthesis, but it has recently been shown<sup>47</sup> that inosine (**I**), which is capable of base-pairing with cytosine, can replace guanosine in non-enzymatic RNA replication systems with no loss of rate or fidelity. **7** (dA) is readily converted to **9** (dI) (Fig. 2) by deaminative hydrolysis, which spontaneously occurs

very slowly in nucleic acid polymers<sup>48</sup>, and is greatly accelerated by the presence of nitrous acid<sup>49</sup>. To demonstrate that this conversion can occur under mild conditions consistent with our primordial geochemical scenario<sup>50</sup>, we treated **7** (dA) with nitrous acid at pH 4 (the same conditions in which we were able to effect desulfurization of **26**). After four days at room temperature, approximately 40% of **7** (dA) had been converted to **9** (dI), providing a 60:40 mixture of **7** (dA) and **9** (dI) (Fig. 2). A control experiment monitoring the decomposition of deoxyadenosine **7** (dA) at pH 4 without nitrous acid showed only a trace of depurination (half-life,  $t_{1/2} = 1,600$  h). When a 67:33 mixture of **7** and **26** (representative of the outcome of photoreduction) was submitted to the reaction conditions, **26** underwent relatively rapid desulfurization first, with deoxyadenosine **7** (dA) undergoing slower deaminative hydrolysis to ultimately provide a 65:35 mixture of **7** (dA) and **9** (dI). Thus, mixtures of **7** (dA) and **9** (dI) are readily obtainable from the partial deaminative hydrolysis of **7** (dA) or its precursor **26**, thereby supplying half of a potential primordial alphabet. Despite the potential for a mismatch in reactivity between deoxypurines and pyrimidines, a 47:53 mixture of **7** (dA) and cytidine (**1**; C) underwent nitrous acid-promoted deamination to provide all four (deoxy)nucleosides **7** (dA), **9** (dI), **1** (C) and uridine (**2**, U) in a 30:17:42:11 ratio (Extended Data Fig. 6). A similar primordial mixture may have been a starting point for genetic information storage. Furthermore, in the absence of geochemically plausible sources of pyrimidine deoxynucleotides and purine ribonucleotides, heteropolymers made from a mixture of purine deoxyribonucleotides and pyrimidine ribonucleotides should possess heritable backbone heterogeneity and thus a 1:1 phenotype-to-genotype correspondence, which is potentially advantageous in the evolution of catalytic activity<sup>18</sup>.

## Systems-level prebiotic plausibility

Having demonstrated the potential of a divergent route to yield a local mixture of **7** (dA), **9** (dI), **1** (C) and **2** (U), we sought to evaluate the key question of whether all four nucleosides could persist after divergence in the sequence. We chose a 1:1 mixture of  $\alpha$ - and  $\beta$ -2-thiouridines (**13** and **3**) as our starting point, obtained from the partial photoanomerisation of **13**, and evaluated two particular combinations of reactions as representative permutations of a primordial geochemical process (Fig. 4, routes A and B). In route A, exposure of the mixture to nitrous acid (pH 4) generates a mixture of **12** and **1** (100% yield for **12** from **13**, 54% yield for **1** from **3**). **12** is formed from **13** potentially by intramolecular addition of the C2' hydroxyl to C2 of an S-nitrosoyl intermediate, and subsequent elimination of SNO<sup>-</sup>. Dry-state glycosylation of **16** and a 1:1 mixture of **12** and **1** (C) in the presence of MgCl<sub>2</sub> leads to a mixture of **18** and **19** as described in our route development above; however, critically, 95% of **1** persists in this mixture. Subsequent photoreduction in the presence of ferrocyanide and bisulfite generates the expected mixture of purine nucleosides **7** (dA), **26**, **20** and **27** alongside **1** (C). Finally, a second exposure to nitrous acid converts this mixture into the components of a competent genetic system, **7** (dA), **9** (dI) (10% and 9% yield respectively from **12** for three steps), **1** (C) and **2** (U) (84% combined persistence after three steps) with no major nucleoside impurities. Products derived from **19**—with the wrong  $N^7$  regiochemistry—are hydrolysed in the last step. It is noteworthy that this route is only viable from a systems-level approach—for instance, the pyrimidines are fairly rapidly destroyed in the photoreduction step in the absence of the thioanhydropurines (Extended Data Fig. 7).

Route B presents an alternative in which initial hydrolysis of the mixture of **13** and **3** generates glycosyl donor **15** (26% yield) alongside the pyrimidine nucleosides (4% of **1** (C), 2% of **2** (U), 92% of **3** remaining). **3** has previously been shown to hydrolyse to **1** (C) and **2** (U) in greater yields (44%) over longer periods<sup>7</sup>. A representative mixture of **15**, **1** (C) and **2** (U) (2:1:1) was then subjected to tethered glycosylation, resulting in **18** and **19** as above (30% and 50% yield respectively), with 80% and 95% persistence of **1** (C) and **2** (U), respectively. Photoreduction of the mixture—this time with hydrogen sulfide—provides purine



**Fig. 4 | A systems-level approach to a potential primordial genetic alphabet composed of 1(C), 2(U), 7(dA) and 9(dl).** A mixture of the  $\alpha$ - and  $\beta$ -epimers of the 2-thiocytidines 13 and 3 (which interconvert in UV light), can generate a mixture containing 1(C), 2(U), 7(dA) and 9(dl). A general route is shown at left. The thiopyrimidines are initially converted into the canonical pyrimidines (cytidine (1) and uridine (2)) and the  $\alpha$ -anhydropurimidines 12 and 15. The latter

undergo tethered glycosylation and then photoreduction to selectively provide purine deoxyribonucleosides 7(dA) and 9(dl) as depicted in Fig. 2. The pyrimidines 1(C) and 2(U) persist through each step of this sequence, ultimately generating a mixture of all four nucleosides. Specific conditions and yields for two possible particular routes (routes A and B) are shown at right.

products 7, 26, 20 and 27 alongside the pyrimidines 1(C) and 2(U). Finally, nitrosation furnished the key mixture of 7(dA) and 9(dl) (6% for each from 15 for three steps) alongside pyrimidine nucleosides (43% persistence over three steps; the final ratio of dA:dl:C:U in the mixture is 14:13:45:28, Extended Data Fig. 8). Thus, sequences comprised of various orders of operations and various photoreduction conditions—which might plausibly emulate a terrestrial geochemical scenario—simultaneously generate the components of a mixed genetic system. The exact ratio of 1(C) and 2(U) (ribosylpyrimidines) to 7(dA) and 9(dl) (deoxyribosylpurines) in the final mixture will depend on the ratio of  $\alpha$ -(anhydro)pyrimidines (13, 12 and 15) to  $\beta$ -(thio)pyrimidines (1, 2 and 3) earlier in the sequence, which in turn varies with environmental conditions.

In conclusion, a highly efficient synthesis of both deoxyadenosine 7(dA) and deoxynosine 9(dl), requiring only prebiotically plausible reagents and conditions, is reported. In contrast to all previous attempts to synthesize purine nucleosides, our synthesis is both prebiotically plausible and strictly stereo-, regio- and furanosyl-selective for the only

isomer of the deoxypurine nucleosides used in modern biology. The pathway proceeds mostly via simple hydrolysis or dry-state processes, with a key reduction step promoted by UV irradiation supported by distinct mechanisms. The (photo)chemical selection exhibited by this route hints at an explanation for the biological importance of one isomer of nucleic acid from the many that are conceivable. We have demonstrated that co-occurring sequences leading selectively to both RNA pyrimidine and DNA purine nucleosides can result in mixtures that could conceivably complete a genetic alphabet. DNA building blocks can thus be co-produced with the RNA pyrimidine nucleosides—which is consistent with, and perhaps evidence for, the coexistence of RNA and DNA building blocks at the dawn of life.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions

and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-2330-9>.

1. Samanta, B. & Joyce, G. F. A reverse transcriptase ribozyme. *eLife* **6**, e31153 (2017).
2. Gilbert, W. Origin of life: the RNA world. *Nature* **319**, 618 (1986).
3. Joyce, G. F. The antiquity of RNA-based evolution. *Nature* **418**, 214–221 (2002).
4. Bhowmik, S. & Krishnamurthy, R. The role of sugar-backbone heterogeneity and chimeras in the simultaneous emergence of RNA and DNA. *Nat. Chem.* **11**, 1009–1018 (2019).
5. Xu, J., Green, N. J., Gibard, C., Krishnamurthy, R. & Sutherland, J. D. Prebiotic phosphorylation of 2-thiouridine provides either nucleotides or DNA building blocks via photoreduction. *Nat. Chem.* **11**, 457–462 (2019).
6. Powner, M. W., Gerland, B. & Sutherland, J. D. Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. *Nature* **459**, 239–242 (2009).
7. Xu, J. et al. A prebiotically plausible synthesis of pyrimidine  $\beta$ -ribonucleosides and their phosphate derivatives involving photoanomerization. *Nat. Chem.* **9**, 303–309 (2017).
8. Heuberger, B. D., Pal, A., Del Frate, F., Topkar, V. V. & Szostak, J. W. Replacing uridine with 2-thiouridine enhances the rate and fidelity of nonenzymatic RNA primer extension. *J. Am. Chem. Soc.* **137**, 2769–2775 (2015).
9. Walton, T. & Szostak, J. W. A highly reactive imidazolium-bridged dinucleotide intermediate in nonenzymatic RNA primer extension. *J. Am. Chem. Soc.* **138**, 11996–12002 (2016).
10. Li, L. et al. Enhanced nonenzymatic RNA copying with 2-aminoimidazole activated nucleotides. *J. Am. Chem. Soc.* **139**, 1810–1813 (2017).
11. Fuller, W. D., Orgel, L. E. & Sanchez, R. A. Studies in prebiotic synthesis: VI. Solid-state synthesis of purine nucleosides. *J. Mol. Evol.* **1**, 249–257 (1972).
12. Becker, S. et al. A high-yielding, strictly regioselective prebiotic purine nucleoside formation pathway. *Science* **352**, 833–836 (2016).
13. Kim, H. & Benner, S. A. Prebiotic stereoselective synthesis of purine and noncanonical pyrimidine nucleotides from nucleobases and phosphorylated carbohydrates. *Proc. Natl. Acad. Sci. USA* **114**, 11315–11320 (2017).
14. Becker, S. et al. Unified prebiotically plausible synthesis of pyrimidine and purine RNA ribonucleotides. *Science* **366**, 76–82 (2019).
15. Teichert, J. S., Kruse, F. M. & Trapp, O. Direct prebiotic pathway to DNA nucleosides. *Angew. Chem. Int. Ed.* **58**, 9944–9947 (2019).
16. Reichard, P. From RNA to DNA, why so many ribonucleotide reductases? *Science* **260**, 1773–1777 (1993).
17. Leu, K., Obermayer, B., Rajamani, S., Gerland, U. & Chen, I. A. The prebiotic evolutionary advantage of transferring genetic information from RNA to DNA. *Nucleic Acids Res.* **39**, 8135–8147 (2011).
18. Sutherland, J. D. & Whitfield, J. N. Prebiotic chemistry: a bioorganic perspective. *Tetrahedron* **53**, 11493–11527 (1997).
19. Trevino, S. G., Zhang, N., Elenko, M. P., Lupták, A. & Szostak, J. W. Evolution of functional nucleic acids in the presence of nonheritable backbone heterogeneity. *Proc. Natl. Acad. Sci. USA* **108**, 13492–13497 (2011).
20. Gavette, J. V., Stoop, M., Hud, N. V. & Krishnamurthy, R. RNA–DNA chimeras in the context of an RNA world transition to an RNA/DNA world. *Angew. Chem. Int. Ed.* **55**, 13204–13209 (2016).
21. Schöffstall, A. M. Prebiotic phosphorylation of nucleosides in formamide. *Orig. Life* **7**, 399–412 (1976).
22. Lohrmann, R. & Orgel, L. E. Urea-inorganic phosphate mixtures as prebiotic phosphorylating agents. *Science* **171**, 490–494 (1971).
23. 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).
24. Ishiwata, A., Lee, Y. J. & Ito, Y. Recent advances in stereoselective glycosylation through intramolecular aglycon delivery. *Org. Biomol. Chem.* **8**, 3596–3608 (2010).
25. Springsteen, G. & Joyce, G. F. Selective derivatization and sequestration of ribose from a prebiotic mix. *J. Am. Chem. Soc.* **126**, 9578–9583 (2004).
26. 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).
27. Vorbrüggen, H. & Ruh-Pohlennz, C. *Handbook of Nucleoside Synthesis* (Wiley, 2001).
28. Holm, N. G., Oze, C., Mousis, O., Waite, J. H. & Guillet-Lepoutre, A. Serpentization and the formation of H<sub>2</sub> and CH<sub>4</sub> on celestial bodies (planets, moons, comets). *Astrobiology* **15**, 587–600 (2015).
29. Sanchez, R. A., Ferris, J. P. & Orgel, L. E. Studies in prebiotic synthesis. II: Synthesis of purine precursors and amino acids from aqueous hydrogen cyanide. *J. Mol. Biol.* **80**, 223–253 (1967).
30. Hudson, J. S. et al. A unified mechanism for abiotic adenine and purine synthesis in formamide. *Angew. Chem. Int. Ed.* **51**, 5134–5137 (2012).
31. Giner-Sorolla, A., Thom, E. & Bendich, A. Studies on the thiation of purines. *J. Org. Chem.* **29**, 3209–3212 (1964).
32. Levy, M. & Miller, S. L. The stability of the RNA bases: implications for the origin of life. *Proc. Natl. Acad. Sci. USA* **95**, 7933–7938 (1998).
33. Ritson, D. J. & Sutherland, J. D. Synthesis of aldehydic ribonucleotide and amino acid precursors by photoredox chemistry. *Angew. Chem. Int. Ed.* **52**, 5845–5847 (2013).
34. Robertson, M. P., Levy, M. & Miller, S. L. Prebiotic synthesis of diaminopyrimidine and thiocytosine. *J. Mol. Evol.* **43**, 543–550 (1996).
35. Roberts, S. J. et al. Selective prebiotic conversion of pyrimidine and purine anhydronucleosides into Watson–Crick base-pairing arabino-furanosyl nucleosides in water. *Nat. Commun.* **9**, 4073–4082 (2018).
36. Ranjan, S., Todd, Z. R., Rimmer, P. B., Sasselov, D. D. & Babbin, A. R. Nitrogen oxide concentrations in natural waters on early Earth. *Geochim. Geophys. Geosyst.* **20**, 2021–2039 (2019).
37. Xu, J. et al. Photochemical reductive homologation of hydrogen cyanide using sulfite and ferrocyanide. *Chem. Commun.* **54**, 5566–5569 (2018).
38. Marion, G. M., Kargel, J. S., Crowley, J. K. & Catling, D. C. Sulfite–sulfide–sulfate–carbonate equilibria with applications to Mars. *Icarus* **225**, 342–351 (2013).
39. Rios, A. C. & Tor, Y. On the origin of the canonical nucleobases: an assessment of selection pressures across chemical and early biological evolution. *Isr. J. Chem.* **53**, 469–483 (2013).
40. Rios, A. C., Yu, H. T. & Tor, Y. Hydrolytic fitness of N-glycosyl bonds: comparing the deglycosylation kinetics of modified, alternative, and native nucleosides. *J. Phys. Org. Chem.* **28**, 173–180 (2014).
41. Panzica, R. P., Rousseau, R. J., Robins, R. K. & Townsend, L. B. Relative stability and a quantitative approach to the reaction mechanism of the acid-catalyzed hydrolysis of certain 7-and 9- $\beta$ -D-ribofuranosylpurines. *J. Am. Chem. Soc.* **94**, 4708–4714 (1972).
42. Lindahl, T. & Nyberg, B. Rate of depurination of native deoxyribonucleic acid. *Biochemistry* **11**, 3610–3618 (1972).
43. Hättig, C. Structure optimizations for excited states with correlated second-order methods: CC2 and ADC(2). *Adv. Quantum Chem.* **50**, 37–60 (2005).
44. Drew, A. & Wormit, M. The algebraic diagrammatic construction scheme for the polarization propagator for the calculation of excited states. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **5**, 82–95 (2015).
45. Sauer, M. C., Crowell, R. A. & Shkrob, I. A. Electron photodetachment from aqueous anions. I. Quantum yields for generation of hydrated electron by 193 and 248 nm laser photoexcitation of miscellaneous inorganic anions. *J. Phys. Chem. A* **108**, 5490–5502 (2004).
46. Pascoe, D. J., Ling, K. B. & Cockroft, S. L. The origin of chalcogen-bonding interactions. *J. Am. Chem. Soc.* **139**, 15160–15167 (2017).
47. Kim, S. C., O’Flaherty, D. K., Zhou, L., Lelyveld, V. S. & Szostak, J. W. Inosine, but none of the 8-oxo-purines, is a plausible component of a primordial version of RNA. *Proc. Natl. Acad. Sci. USA* **115**, 13318–13323 (2018).
48. Karran, P. & Lindahl, T. Hypoxanthine in deoxyribonucleic acid: generation by heat-induced hydrolysis of adenine residues and release in free form by a deoxyribonucleic acid glycosylase from calf thymus. *Biochemistry* **19**, 6005–6011 (1980).
49. Shapiro, R. & Pohl, S. H. Reaction of ribonucleosides with nitrous acid. Side products and kinetics. *Biochemistry* **7**, 448–455 (1968).
50. Mariani, A. D., Russell, A., Javelle, T. & Sutherland, J. D. A light-releasable potentially prebiotic nucleotide activating agent. *J. Am. Chem. Soc.* **140**, 8657–8661 (2018).

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© Crown 2020

## Data and materials availability

The Supplementary Information available for this Article contains all procedures, characterization data, NMR spectra, HPLC traces, X-ray data and Cambridge Crystallographic Data Centre (CCDC) numbers, plus theoretical methods and data. Any additional data are available from the corresponding author upon reasonable request.

## Code availability

All custom code used to generate the data in this study is available upon reasonable request.

**Acknowledgements** The authors thank all JDS group members for discussions. This research was supported by the Medical Research Council (MC\_UP\_A024\_1009), the Simons Foundation

(290362 to J.D.S., 494188 to R.S.), and a grant from the National Science Centre Poland (2016/23/B/ST4/01048 to R.W.G.). M.J.J. acknowledges the support of the ‘Diamond Grant’ (0144/DIA/2017/46) from the Polish Ministry of Science and Higher Education and a computational grant from Wrocław Centre of Networking and Supercomputing (WCSS). R.S. thanks the Foundation for Polish Science for support from the START Fellowship.

**Author contributions** Experimental contributions by J.X., V.C., N.J.G., D.A.R. and A.D.B. Theoretical contributions by M.J.J., R.W.G. and R.S. Crystallography by A.D.B. This work was supervised by J.D.S. All authors co-wrote the manuscript.

**Competing interests** The authors declare no competing interests.

### Additional information

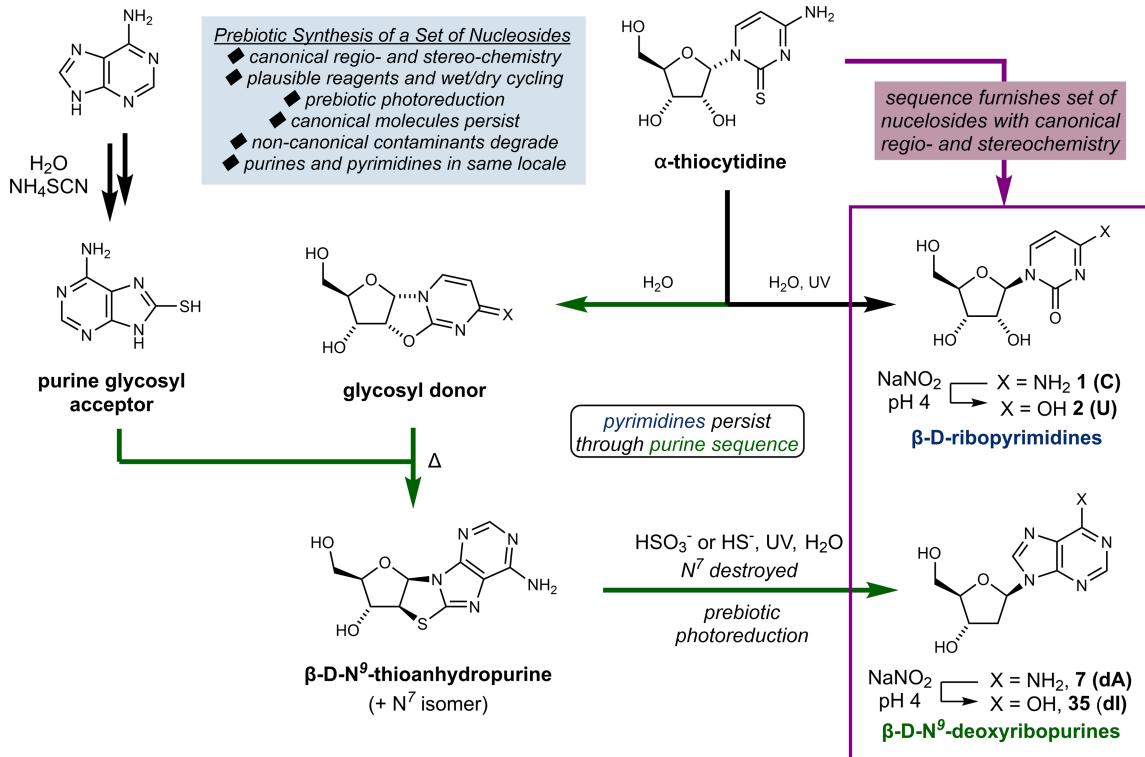
**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-020-2330-9>.

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

**Peer review information** *Nature* thanks Hannes Mutschler and Yitzhak Tor for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.

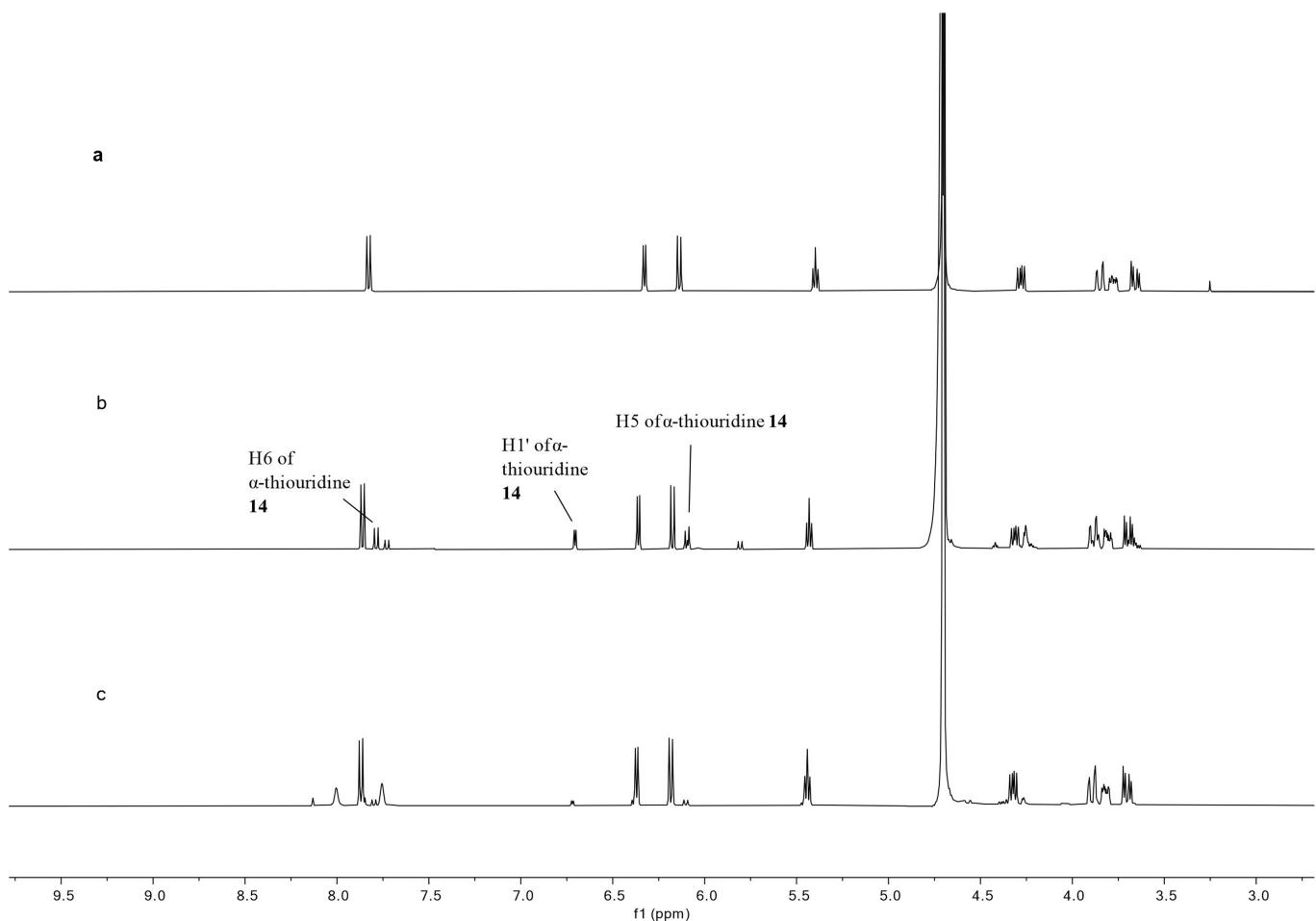
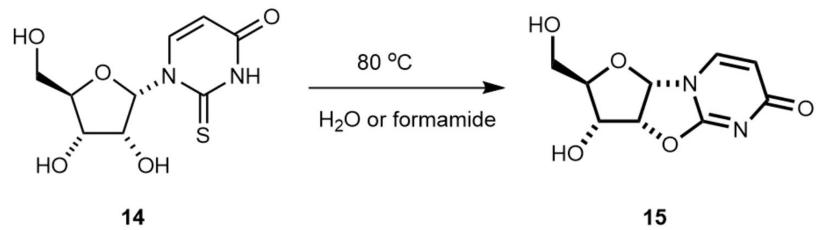
# Article



## Extended Data Fig. 1 | A summary of the main findings of the work.

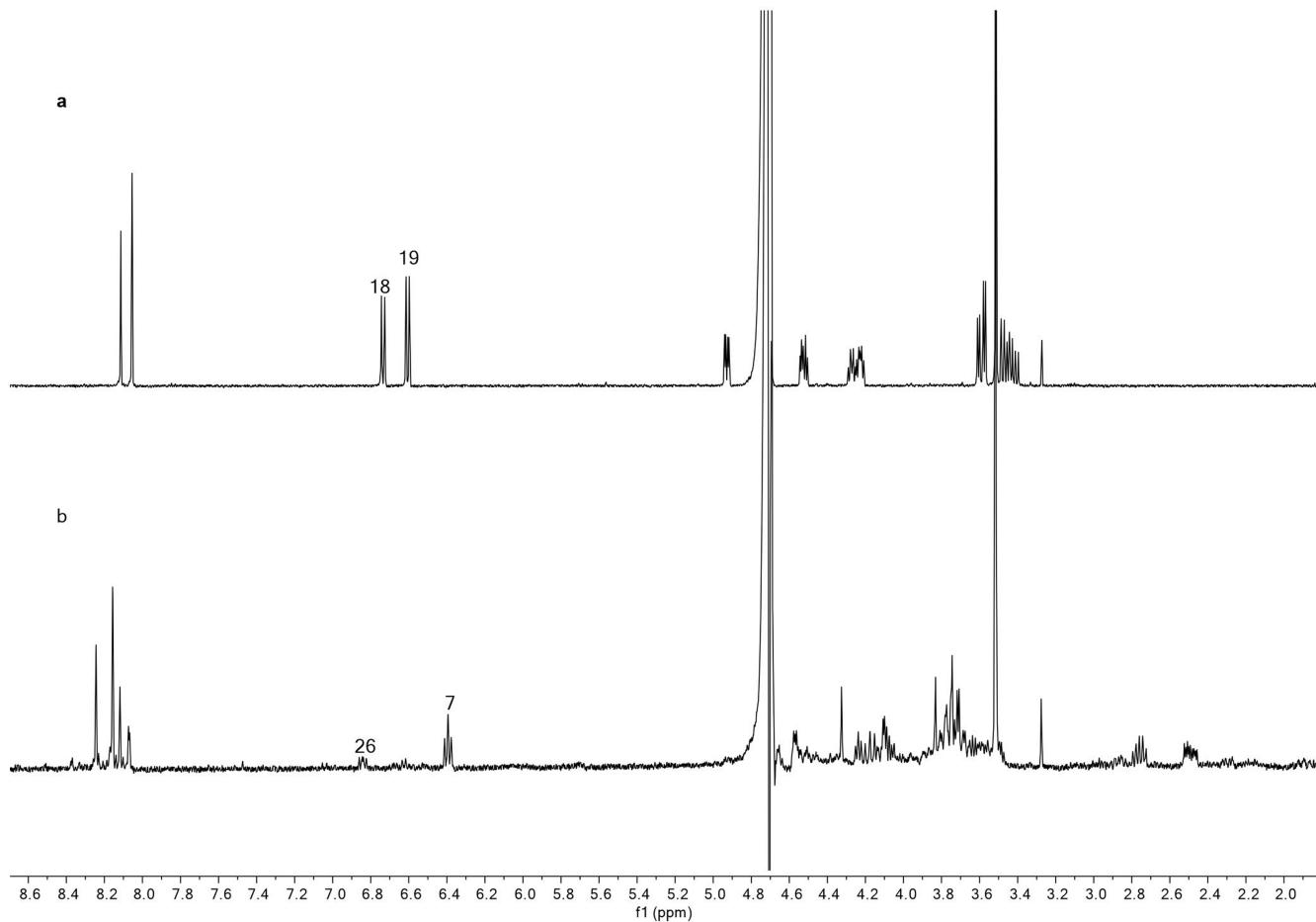
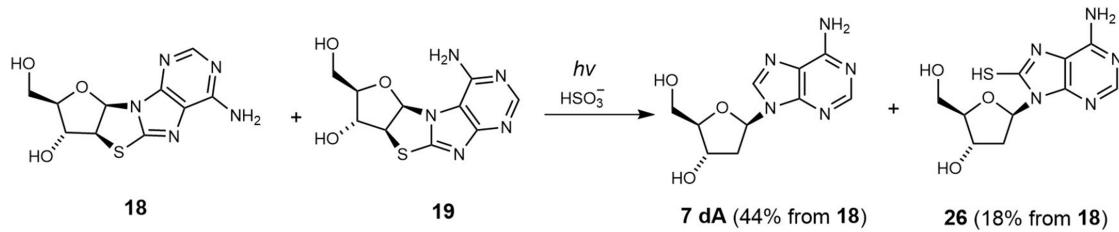
Previously, a prebiotically plausible synthesis of  $\beta$ -ribopyrimidines C and U has been identified using  $\alpha$ -thiocytidine. Herein, we demonstrate that the same intermediate can undergo a distinct prebiotically plausible process that could have happened in a similar—or the same—environment. This process furnishes  $\beta$ -D- $N^9$ -deoxyribopurine nucleosides dA and dI alongside the pyrimidines.

Remarkable selectivity enforced by UV irradiation and hydrolysis operates throughout the reported ribosylpyrimidine synthesis and the discovered deoxyribosylpurine synthesis, resulting in a set of nucleosides with only the canonical regio- and stereochemistry. The coexistence in one location of a set of nucleosides similar to this is thought to be a precondition for the spontaneous emergence of life on Earth<sup>6,47</sup>.



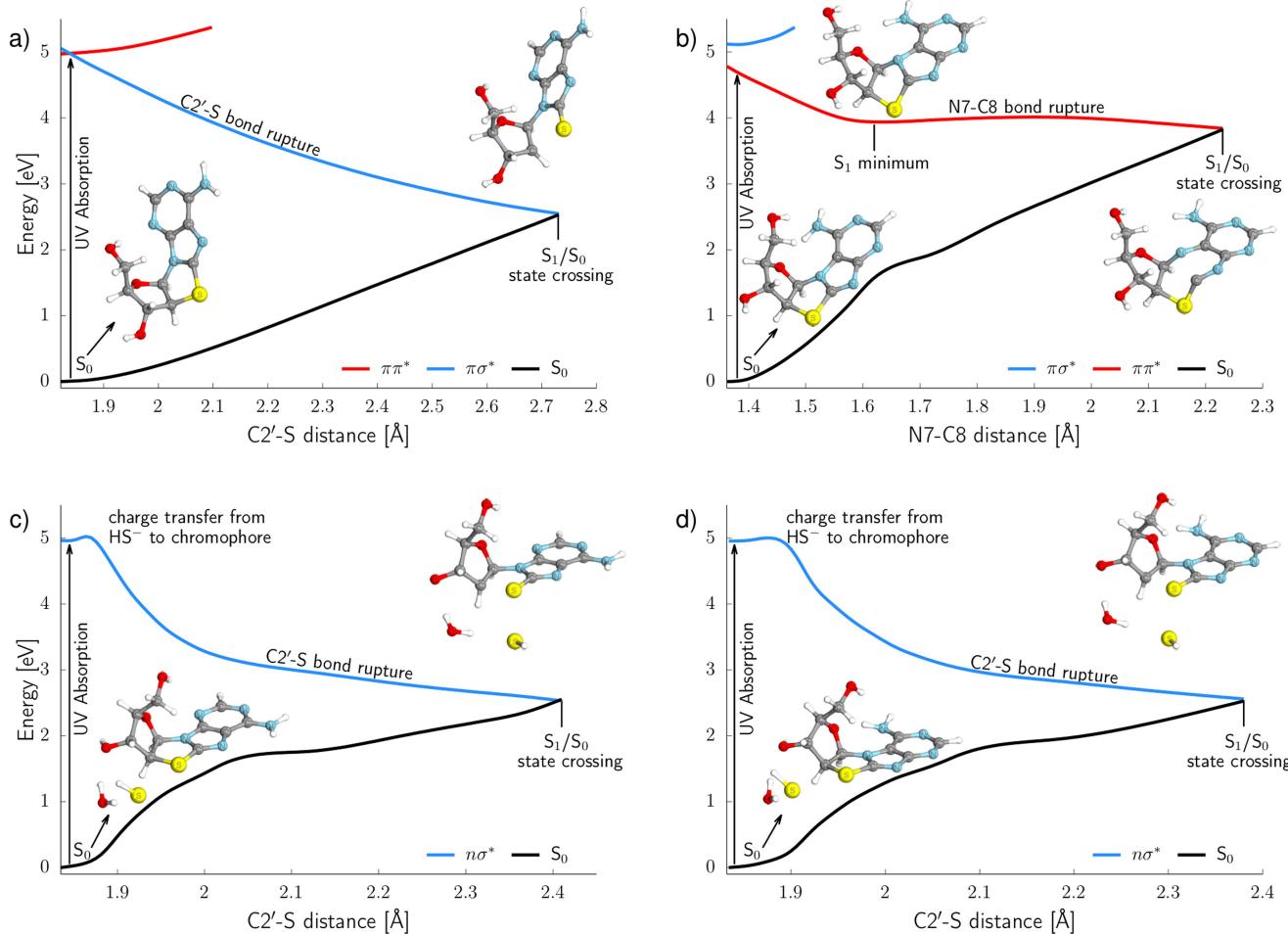
**Extended Data Fig. 2 |  $^1\text{H}$  NMR spectra of conversion of  $\alpha$ -anhydrouridine (**15**) from  $\alpha$ -thiouridine (**14**). a.**  $^1\text{H}$  NMR spectrum of **15**. **b,**  $^1\text{H}$  NMR spectrum of the reaction mixture after heating **14** in  $\text{H}_2\text{O}$ . **c,**  $^1\text{H}$  NMR spectrum of the reaction mixture after heating **14** in formamide. f1, chemical shift ( $\delta$ ).

# Article



**Extended Data Fig. 3 |  $^1\text{H}$  NMR spectra of photoreduction of  $\text{N}^7$ -8,2'-anhydro-thioadenosine (**18**) and  $\text{N}^9$ -8,2'-anhydro-thioadenosine (**19**) mixture with bisulfite.** **a**,  $^1\text{H}$  NMR spectrum of the crude mixture before

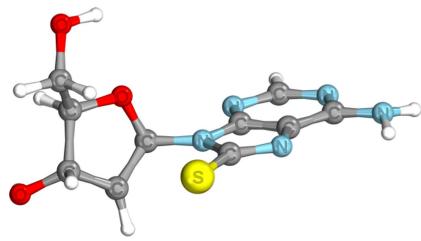
irradiation; the ratio of  $\text{N}^7:\text{N}^9$  isomer was 4:5. **b**,  $^1\text{H}$  NMR spectrum of the mixture after irradiation for 7 h; the  $\text{N}^9$  isomers **dA** (**7**) and **26** are the only detectable products. f1, chemical shift ( $\delta$ ).



**Extended Data Fig. 4 | Potential energy surfaces and S<sub>1</sub>/S<sub>0</sub> state crossings of the key photochemical steps in deoxyadenosine synthesis calculated using ADC(2) and the ma-def2-TZVP basis set.** See Supplementary Information for details. **a**, Potential energy profile of UV-induced C–S bond scission of **18**. C–S bond opening may spontaneously occur in **18**, leading to a peaked S<sub>1</sub>/S<sub>0</sub> state crossing; however, a reducing agent is necessary to maintain that geometry after reaching the S<sub>0</sub> state. **b**, Potential energy profile of

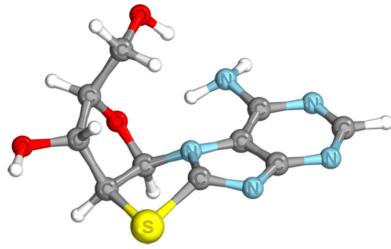
UV-induced N7–C8 bond scission of **19**. N7–C8 bond rupture is the lowest-energy photochemical process in **19** and results in destruction of the purine ring. **c, d**, Potential energy profiles of the UV-induced C–S bond scission of encounter complexes **18** (**c**) and **19** (**d**) with HS<sup>−</sup>. Photochemical C–S bond rupture induced by charge transfer from HS<sup>−</sup> to a chromophore and is a barrierless process.

## Article



**31 radical anion**

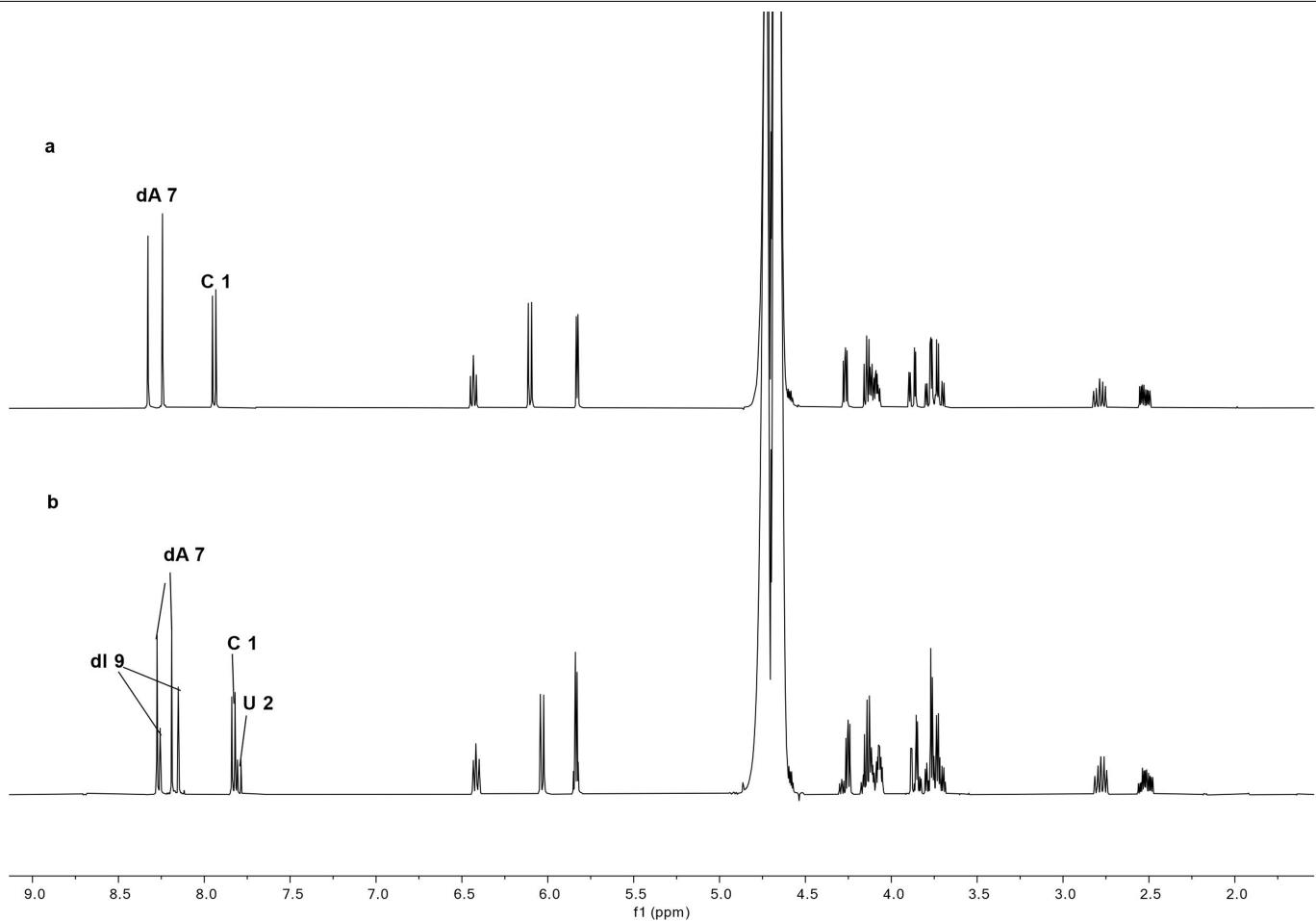
Adiabatic electron affinity: 2.50 eV



**32 radical anion**

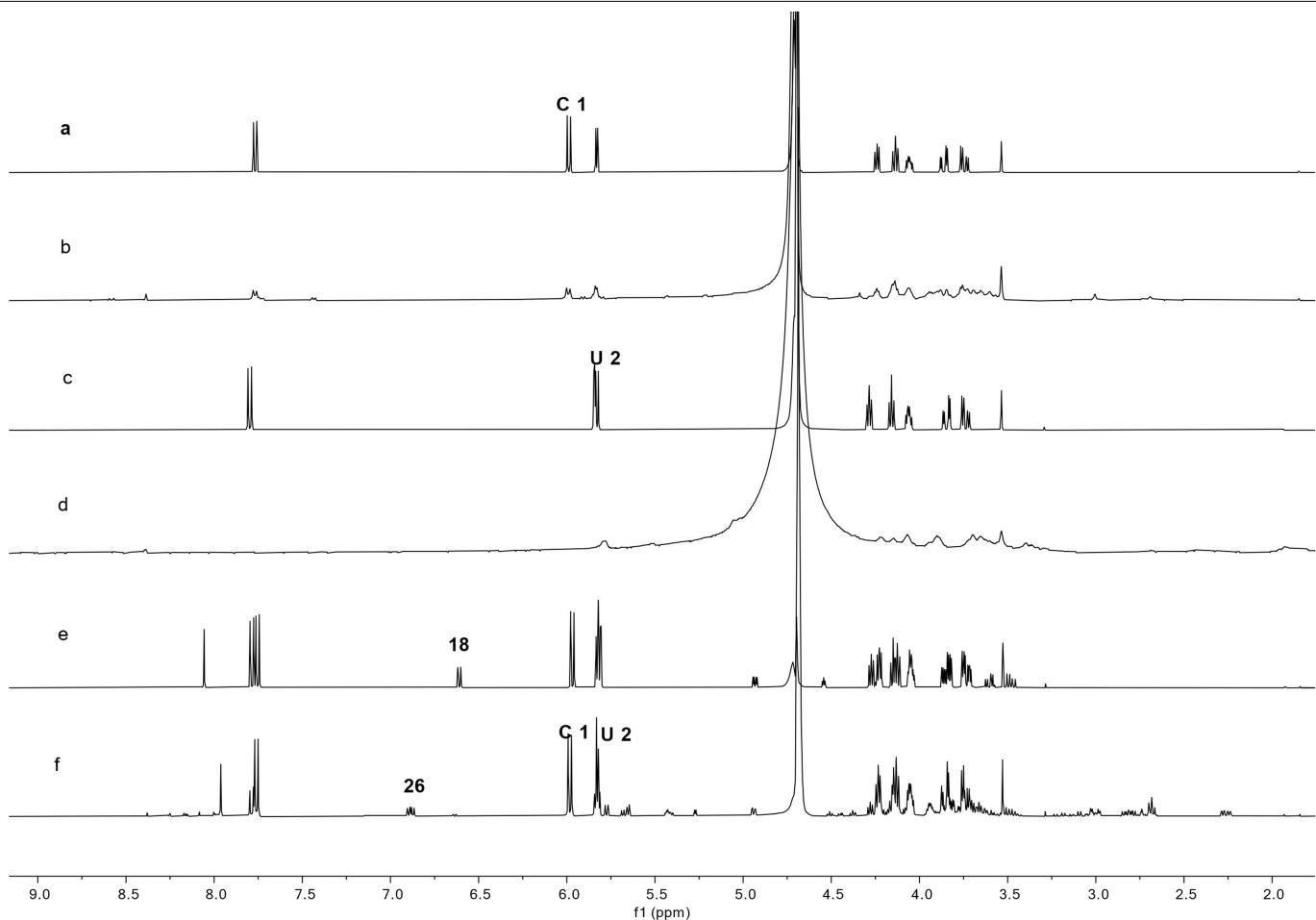
Adiabatic electron affinity: 1.83 eV

**Extended Data Fig. 5 | Equilibrium geometries of C2, S8 radical anion (31) and C8, N9 radical anion (32).** Radical anions may be formed after accepting a hydrated electron from the environment. The adiabatic electron affinities are calculated using ωB97X-D/IEFPCM and the ma-def2-TZVP basis set.



**Extended Data Fig. 6 | <sup>1</sup>H NMR spectra for the reactions of deoxyadenosine (dA, 7) and cytidine (C, 1) with nitrous acid. a.** <sup>1</sup>H NMR spectrum of the mixture of dA (7) and C (1). **b.** <sup>1</sup>H NMR spectrum of the reaction mixture after

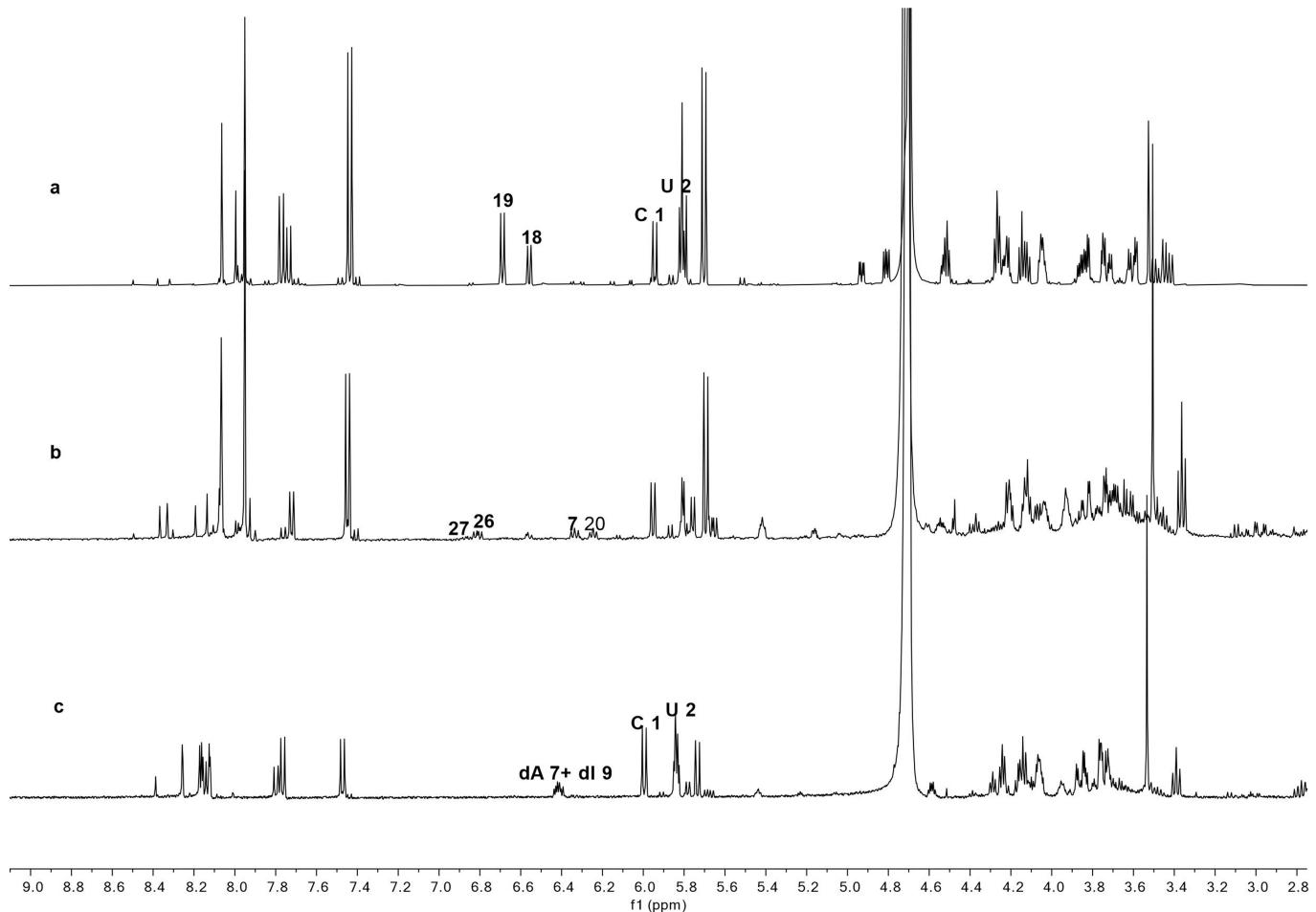
4 d, showing that the ratio of the four (deoxy)nucleosides dA (7), deoxyinosine (dl, 9), C (1), and uridine (U, 2) is 30:17:42:11. f1, chemical shift ( $\delta$ ).



**Extended Data Fig. 7 | <sup>1</sup>H NMR spectra for stability study of cytidine (C1)**

**and uridine (U2) at 254 nm irradiation with bisulfite.** **a,** <sup>1</sup>H NMR spectrum of the mixture of C (1), bisulfite and K<sub>4</sub>Fe(CN)<sub>6</sub> in the dark. **b,** As in **a**, after 10 h of irradiation. **c,** <sup>1</sup>H NMR spectrum of the mixture of U (2), bisulfite and K<sub>4</sub>Fe(CN)<sub>6</sub>

in the dark. **d,** As in **c**, after 10 h of irradiation. **e,** <sup>1</sup>H NMR spectrum of the mixture of C (1), U (2), *N*<sup>3</sup>-thioanhydroadenosine (18), bisulfite and K<sub>4</sub>Fe(CN)<sub>6</sub> in the dark. **f,** As in **e**, after 10 h of irradiation. f1, chemical shift ( $\delta$ ).



**Extended Data Fig. 8 | <sup>1</sup>H NMR spectra for sequential reactions with the mixture of  $\alpha$ -anhydrouridine (15), C (1) and U (2). a.** <sup>1</sup>H NMR spectrum of the mixture after heating with 8-mercaptopoadenine (16) and magnesium chloride at

150 °C for 1.5 d. **b.** As in a, after irradiation with hydrogen sulfide at 254 nm. **c.** As in a, after reacting with nitrous acid for 2 d; dA (7):dl (9):C (1):U (2) = 14:14:44:28. **f1**, chemical shift ( $\delta$ ).