A one-pot, water compatible synthesis of pyrimidine nucleobases under plausible prebiotic conditions

Hidenori Okamura,[a] Sidney Becker,[a] Niklas Tiede,[a] Stefan Wiedemann,[a] Jonas Feldmann[a] and Thomas Carell[a*]

[a] Center for Integrated Protein Science (CiPS[M]) at the Department of Chemistry, Ludwig-Maximilians-Universität München, Butenandtstr. 5-13, 81377 München, Germany,

[*] E-mail: thomas.carell@lmu.de

Published in **Chem. Commun**. 2019, **55**, 1939-1942 on 07.01.2019, DOI: 10.1039/C8CC09435G

Abstract

Herein, we report a new prebiotically plausible pathway towards a pyrimidine nucleobase in continuous manner. The route involves simultaneous methylation and carbamoylation of cyanoacetylene-derived α,β-unsaturated thioamide with *N*-methyl-*N*-nitrosourea (MNU) in aqueous media. This provides *S*-methylpyrimidinone in one-pot, which can be converted into a variety of 4-substituted pyrimidine nucleobases including cytosine and uracil.

Main Text

Elucidating the chemical process that gave birth to biologically essential molecules on the early Earth is key for understanding the origin of life. Among the major biomolecules, RNA is thought to have played a central role during chemical evolution. RNA can store genetic information and at the same time has the potential to catalyse chemical reactions.[1–3] Therefore, RNA might be able to establish self-replicating systems amenable to Darwinian evolution.[4,5] The essential components of RNA are purine (A, G) and pyrimidine (U, C) nucleobases, which form Watson-Crick H-bonds. H-bonding and π-stacking allows RNA to fold into complex shapes and modified nucleobases that are also part of contemporary RNA provide functionality far beyond what is observed with DNA.[6] This structural diversity raise the question of how the constituting RNA nucleobases could have emerged. Did the chemical reactions that occurred on the early Earth provide only the canonical nucleosides, or were a plethora of different structures created, followed by a chemical evolution process that led to the selection of the fittest molecules for information encoding? An important additional feature of the pyrimidine and purine heterocycles is that they may be capable of providing a chiral field *via* dehydrated crystals[7,8] that could have played a role in the emergence of chirality. Finally, they can stabilize aggregates of amphiphiles and as such they could have been involved in forming cell-like vesicle surrounded by amphiphilic membranes.[9] All these properties make research about the chemical origin of the pyrimidine and purine heterocycles a pressing problem.

In pioneering work, Ferris et al. demonstrated cytosine formation from urea or sodium cyanate and cyanoacetylene.[10] Miller *et al*. reported cytosine synthesis by heating a saturated solution

of urea with cyanoacetaldehyde.[11] The high reactivity of these cyano-derivatives, however, brings into question their accumulation to significant concentrations.[12,13] Alternatively, pyrimidine nucleobases can also be generated from formamide under high energy conditions (e.g. heat, spark discharges) in the presence of meteorite or mineral powder.[14,15] These conditions, however, may be too harsh for complex chemical systems that may have given rise to RNA.

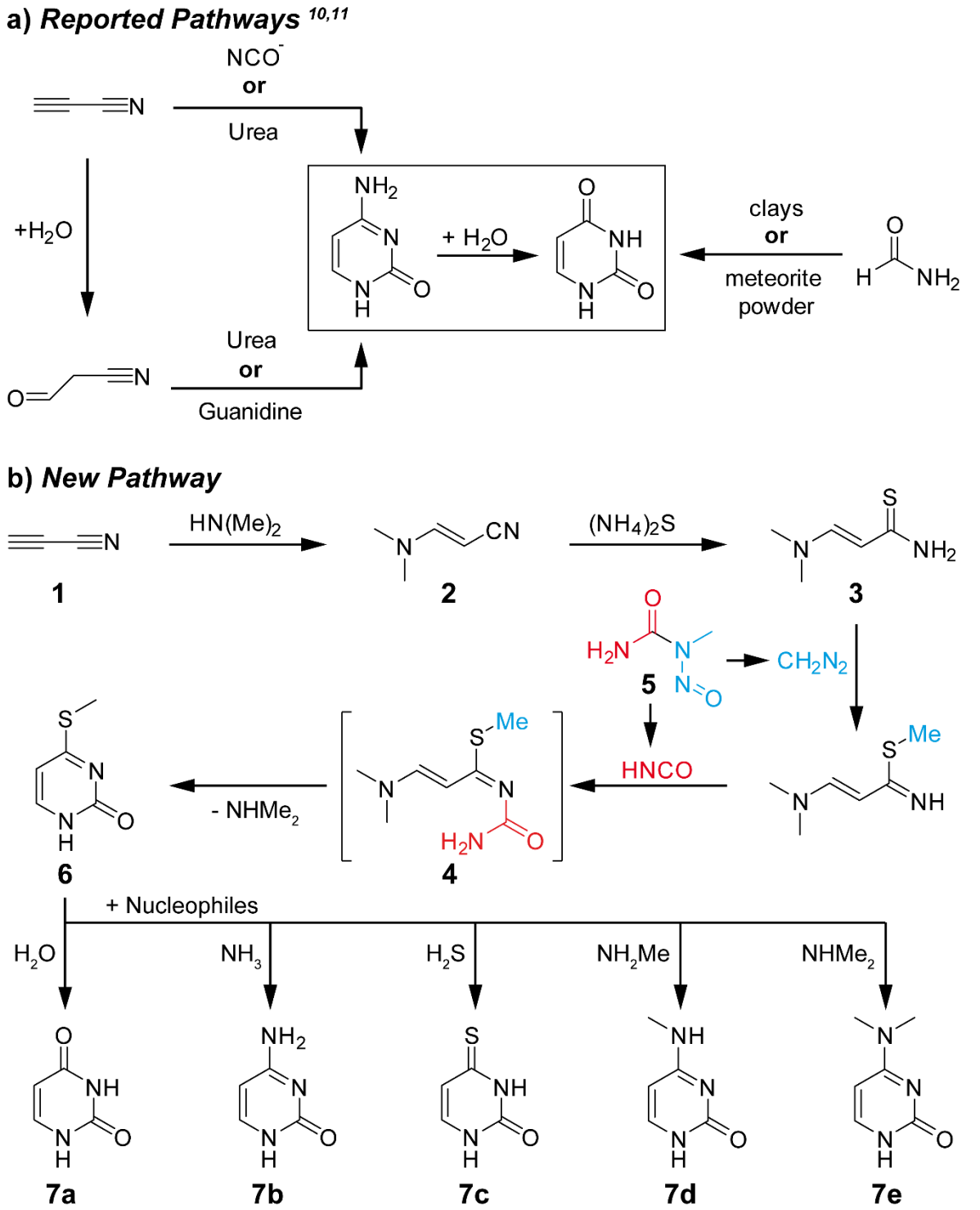
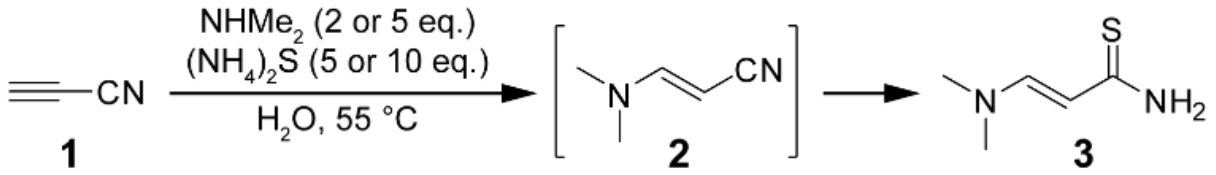


Fig. 1. (a) Previously reported pathway for pyrimidine nucleobase synthesis10,11 and (b) proposed pathway in this study.

Here we report a new prebiotically plausible pathway to pyrimidines, which takes place in water. We discovered that cyanoacetylene **1** (Fig. 1) reacts quickly with dimethylamine. Both compounds were likely present on the early Earth. The reaction provides the stable acrylonitrile derivative **2**, in which the reactivity of cyanoacetylene is captured. **2** is subsequently converted into a stable α,β-unsaturated thioamide **3** upon sulfurization. Simultaneous methylation and carbamoylation of **3** gives carbamoylimidothioate **4**, with spontaneous cyclization to the *S*-methylpyrimidinones (SMePy, **6**) under basic conditions. Simultaneous methylation and carbamoylation under prebiotically plausible conditions is achieved with *N*-methyl-*N*-nitrosourea (MNU, **5**).[16]



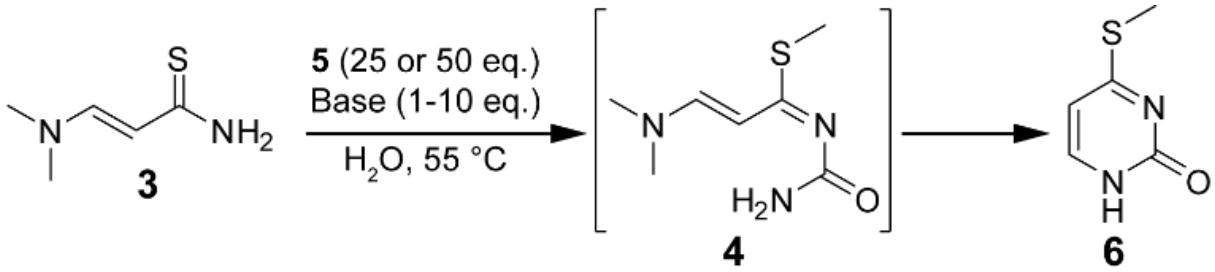
1 **2** **3**

Entry	$NHMe_2$ (eq.)	$(NH_4)_2S$ (eq.)	Yield of **3** (%) [a]
1	2	5	86
2	2	10	73
3	5	5	97
4	5	10	92

Table 1. Yields obtained under different reaction conditions for the conversion of **1** to **3**. [a] Yields were determined by HPLC.

We first investigated the synthesis of α,β-unsaturated thioamide **3** from cyanoacetylene **1** under prebiotically plausible conditions. Due to the electron-withdrawing nature of the nitrile group, the alkyne moiety of cyanoacetylene **1** is highly susceptible to a nucleophilic addition at its alkyne terminus. Indeed, upon addition to aqueous dimethylamine solution, **1** reacted immediately and quantitatively to form 3-(dimethylamino)acrylonitrile **2** (Scheme S1). Compound **2** was subsequently subjected to sulfurization. In initial attempts, we tested solutions of sodium hydrogen sulphide in phosphate, borate or carbonate buffer with different pH as a model solution of hydrogen sulphide.[17] These conditions, however, gave only trace amount of the desired α,β-unsaturated thioamide **3** (data not shown). We next turned our attention to ammonium sulphide as a prebiotically relevant sulfurization reagent. Ammonium sulphide is a salt of ammonia and hydrogen sulphide, both plausible prebiotic molecules. Ammonium sulphide is for example found in the atmosphere of Titan and Jupiter.[18,19] When the acrylonitrile derivative **2** was reacted with different concentrations of ammonium sulphide in water at 55°C, formation of the desired α,β-unsaturated thioamide compound **3** was observed. Notably, it turned out that thioamide **3** (63%) crystallized out from the reaction mixture upon cooling to room temperature, which can be considered to be a naturally occuring enrichment step by temperature change.

We next examined if **3** can be formed in one-pot, despite the presence of competing nucleophiles. Cyanoacetylene **1** was slowly added to the solution containing different equivalents of dimethylamine and ammonium sulphide, and the reaction was gently heated to 55°C and analysed by reverse phase HPLC. The reaction provided thioamide **3** in high yields (up to 97%, Table 1) showing that **1** reacts selectively with dimethylamine followed by capturing of the intermediate with $(NH_4)_2S$.



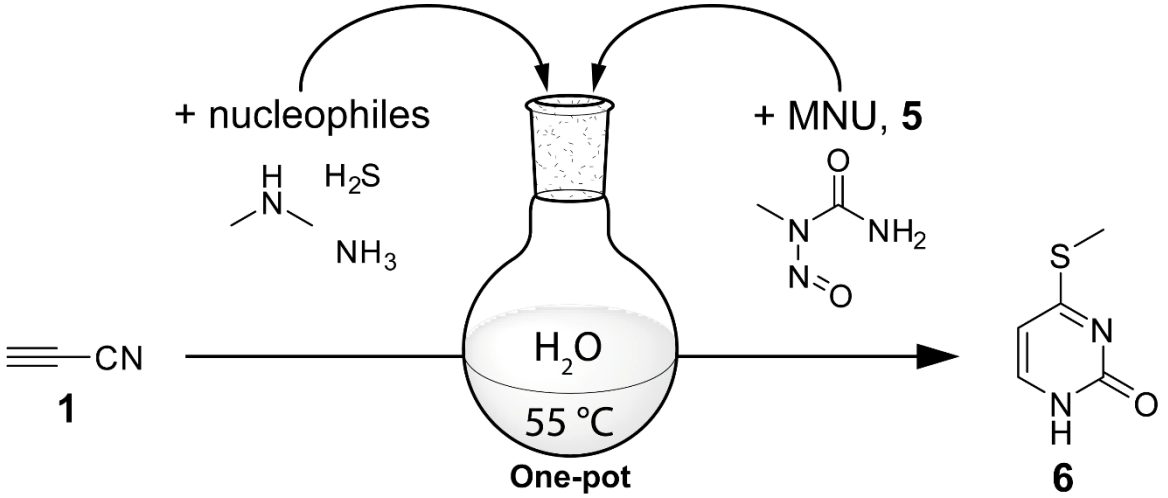
Entry	Base (eq.)	**5** (eq.)	Yield of **6** (%)[a]
1	$NaHCO_3$ (10)	25	15
2	$NaHCO_3$ (10)	50	22
3	Borax (1)	25	8
4	Borax (1)	50	18
5	$(NH_4)_2S$ (3)	25	11
6	$(NH_4)_2S$ (3)	50	17
7	Na_2HPO_4 (10)	25	13
8	Na_2HPO_4 (10)	50	18
9	$CaCO_3$ (5)	25	9
10	$CaCO_3$ (5)	50	13

Table 2. Yields obtained under different reaction conditions for the conversion of **3** to **6**. [a] Yields were determined by HPLC.

We further investigated the reaction of the so formed thioamide **3** with MNU **5** for the formation of SMePy **6**. MNU **5** can be synthesized under plausible prebiotic conditions by reacting *N*-methylurea with sodium nitrite in a slightly acidic environment. MNU **5** separates spontaneously from the water solution, allowing it to be locally available in high quantities. It is known that MNU produces diazomethane and isocyanic acid under basic conditions or by thermal degradation.[20] Accordingly, we assumed that MNU would operate as a bifunctional reagent. We expected that the carbamoylimidothioate **4**, which is formed by the methylation of thiocarbonyl group of thioamide **3** with diazomethane would undergo carbamoylation with isocyanic acid, followed by a base-catalysed cyclization to build up the pyrimidine skeleton. For this, MNU **5** was added to the solution containing thioamide **3** and sodium bicarbonate under mild heating at 55°C. The reaction was monitored by HPLC (Fig. S2a). The newly present peak was isolated and analysed by ^{1}H-NMR, confirming formation of the desired SMePy **6**. We did

not observe the carbamoylimidothioate **4** in the HPLC chromatogram, presumably because it cyclizes immediately to SMePy **6** after the carbamoylation reaction. The yield of SMePy **6** was determined to be 22% by HPLC-MS based on the calibration curve that was prepared using a chemically synthesized reference compound.[21]

The reaction was further investigated regarding its sensitivity to the presence of different prebiotically relevant bases and minerals. The data are summarized in Table 2 and Fig. S2. All tested conditions gave similar yields of SMePy **6**, showing that the reaction is robust, which is one of our main criteria for calling a reaction prebiotically plausible. More importantly, formation of SMePy **6** was observed in the presence of the ammonium sulphide under conditions also utilized for the formation of thioamide **3**. This observation led to the idea to conduct a one-pot synthesis of the SMePy **6**, a precursor of canonical and non-canonical pyrimidine nucleobases.



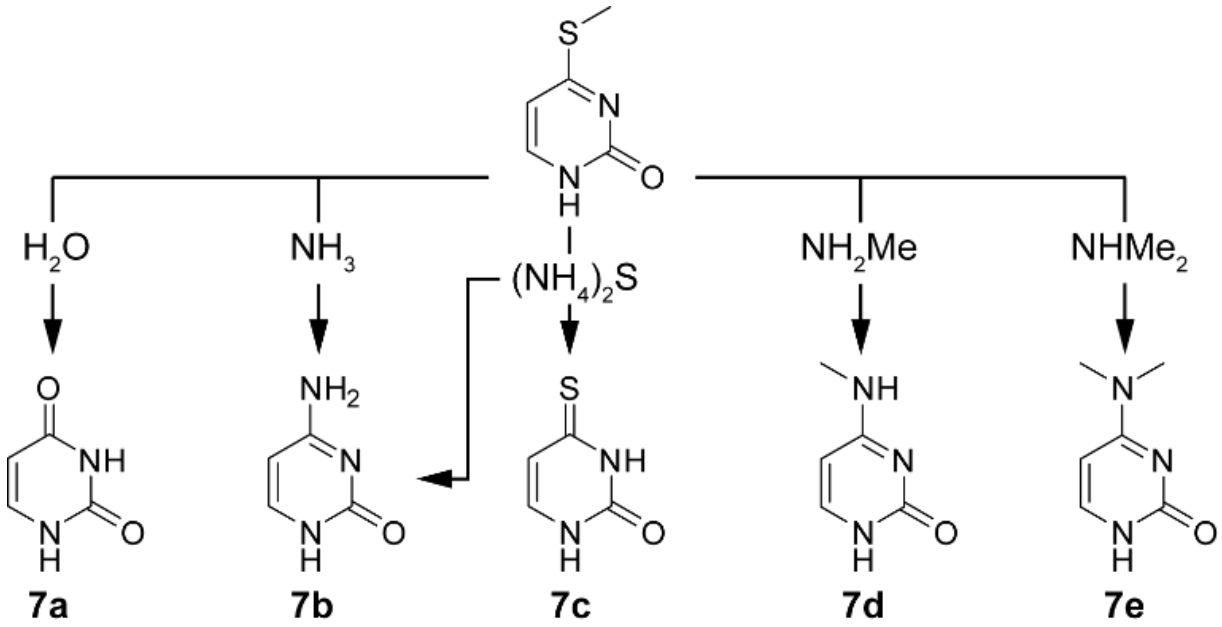
Entry	$NHMe_2$ (eq.)	$(NH_4)_2S$ (eq.)	**5** (eq.)	Yield of **6** (%) [a]
1	2	5	50	17
2	2	10	50	14
3	5	5	50	16
4	5	10	50	15

Figure 2. One-pot synthesis of SMePy **6** in the presence of different equivalents of the reagents. [a] Yields were determined by HPLC.

To achieve a one-pot synthesis of **6** (Fig. 2), cyanoacetylene **1** was first reacted with a dimethylamine solution containing ammonium sulphide in different amounts. After warming the reaction to 55°C for 24 hrs, the reaction mixture was diluted with water and MNU **5** was added. The reaction mixture was again warmed to 55°C for another 24 hrs and finally analysed by reverse phase HPLC. The desired SMePy **6** was obtained in 14-17% yield over three steps under these one-pot conditions.

Finally, we investigated the conversion of SMePy **6** to canonical nucleobases under plausible prebiotic conditions (Table 3). The conversion of SMePy **6** into uracil **7a** can be catalysed by Fe^{2+}. Heating SMePy **6** in aqueous solution containing $FeSO_4$, provided **7a** quantitatively. $FeSO_4$ is naturally available from minerals such as Rozenite and Melanterite.[22] Cytosine **7b** was in contrast obtained by heating SMePy **6** in aqueous ammonia solution in the presence of $FeSO_4$.

This demonstrates an efficient "mineral guided" formation of the canonical pyrimidine bases. Furthermore, reaction of **6** with ammonium sulphide furnished 4-thiouracil **7c**, another well-known modification present in the tRNA of bacteria and archaea.[23] Interestingly, the concentration of ammonium sulphide determines the outcome of this reaction, while either cytosine or 4-thiouracil is formed (Table 3). Since cytosine hydrolysed slowly to uracil, the ammonium sulphide conditions provide cytosine, uracil and 4-thiouracil simultaneously. Finally, we investigated whether also non-canonical bases can form from SMePy **6**. Such bases are present in contemporary RNA as constituents to modify structural and functional features.[6] If they would form under identical conditions, this would strengthen our hypothesis that canonical and non-canonical bases emerged simultaneously as competitor and companion molecules that found their roles in a subsequent chemical selection process.[16,17,24–26] In order to investigate this possibility we reacted SMePy **6** with different amines in aqueous solutions. Upon treatment with methylamine and dimethylamine, 4-mono- and 4-dimethylcytosine (**7d** and **7e**) were obtained respectively, which are indeed found as modified bases in contemporary RNA species.[27]



Entry	Nucleophile (%)	$FeSO_4$ (eq.)	Product (yield in %)
1	H_2O	1	**7a** (quant.)[a]
2	NH_3 (5)	1	**7b** (77)[a]
3	$(NH_4)_2S$ (1)	-	**7b** (5)/ **7c** (80)[b]
4	$(NH_4)_2S$ (5)	-	**7b** (21)/ **7c** (0)[b]
5	$(NH_4)_2S$ (10)	-	**7b** (43)/**7c** (0)[b]
6	NH_2Me (1)	-	**7d** (70)[b]
7	$NHMe_2$ (1)	-	**7e** (95)[b]

Table 3. Conversion yields of 6 to 4-substituted pyrimidines with different nucleophiles. [a] Isolated yield. [b] Yields were determined by HPLC.

In summary, we discovered a new plausible prebiotic pathway, which takes place in continuous manner in water. The process provides canonical as well as C4-modified pyrimidine nucleobases. Although cyanoacetylene has been widely investigated as the building block of nucleobases and nucleosides, its high-electrophilic character has always questioned its involvement. Here we show that one can start the pyrimidine synthesis with a trapped version of cyanoacetylene, namely the stable acrylonitrile derivative. We show that the cascade reaction proceeds under one-pot conditions in a continuous manner to provide SMePy **6**. Importantly the key intermediate SMePy **6** gives rise not only to canonical but also to non-canonical bases arguing for the simultaneous prebiotic formation of a diverse set of pyrimidines under prebiotically plausible conditions.

Conflicts of interest

The authors declare no conflicts of interest.

Acknowledgement

This study was supported by the Deutsche Forschungsgemeinschaft (DFG) for financial support via the programs SFB1032 (TP-A5), SFB749 (TP-A4), SPP-1784, CA275/11-1. Additional funding from the Excellence Cluster EXC117 CiPSM is acknowledged. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Project ID: 741912). H.O. received a Marie Sklodowska-Curie Individual Fellowship (Project ID: 752420) from the European Commission.

Notes and references

1 K. Kruger, P. J. Grabowski, A. J. Zaug, J. Sands, D. E. Gottschling and T. R. Cech, *Cell*, 1982, **31**, 147–157.
2 C. Guerrier-Takada, K. Gardiner, T. Marsh, N. Pace and S. Altman, *Cell*, 1983, **35**, 849–857.
3 W. Gilbert, *Nature*, 1986, **319**, 618–618.
4 T. A. Lincoln and G. F. Joyce, *Science*, 2009, **323**, 1229–1232.
5 J. Attwater, A. Raguram, A. S. Morgunov, E. Gianni and P. Holliger, *Elife*, 2018, **7**, e35255.
6 T. Carell, C. Brandmayr, A. Hienzsch, M. Müller, D. Pearson, V. Reiter, I. Thoma, P. Thumbs and M. Wagner, *Angew. Chem. Int. Ed.*, 2012, **51**, 7110–7131.
7 S. J. Sowerby and W. M. Heckl, *Orig. Life Evol. Biosph.*, 1998, **28**, 283–310.
8 T. Kawasaki, Y. Hakoda, H. Mineki, K. Suzuki and K. Soai, *J. Am. Chem. Soc.*, 2010, **132**, 2874–2875.
9 R. A. Black, M. C. Blosser, B. L. Stottrup, R. Tavakley, D. W. Deamer and S. L. Keller, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 13272–6.
10 J. P. Ferris, R. A. Sanchez and L. E. Orgel, *J. Mol. Biol.*, 1968, **33**, 693–704.
11 M. P. Robertson and S. L. Miller, *Nature*, 1995, **375**, 772–774.
12 L. E. Orgel, *Orig. Life Evol. Biosph.*, 2002, **32**, 279–281.
13 C. Menor-Salván, D. M. Ruiz-Bermejo, M. I. Guzmán, S. Osuna-Esteban and S. Veintemillas-Verdaguer, *Chem. Eur. J.*, 2009, **15**, 4411–4418.
14 D. Niether, D. Afanasenkau, J. K. G. Dhont and S. Wiegand, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, 4272–7.
15 M. Ferus, F. Pietrucci, A. M. Saitta, A. Knížek, P. Kubelík, O. Ivanek, V. Shestivska and S. Civiš, *Proc. Natl. Acad. Sci. U. S. A.*, 2017, **114**, 4306–4311.

16 C. Schneider, S. Becker, H. Okamura, A. Crisp, T. Amatov, M. Stadlmeier and T. Carell, *Angew. Chem. Int. Ed.*, 2018, **57**, 5943–5946.
17 S. Stairs, A. Nikmal, D.-K. Bučar, S.-L. Zheng, J. W. Szostak and M. W. Powner, *Nat. Commun.*, 2017, **8**, 15270.
18 E. Karkoschka, *Icarus*, 1998, **133**, 134–146.
19 M. J. Loeffler, R. L. Hudson, N. J. Chanover and A. A. Simon, *Icarus*, 2016, **271**, 265–268.
20 B. T. Golding, C. Bleasdale, J. McGinnis, S. Müller, H. T. Rees, N. H. Rees, P. B. Farmer and W. P. Watson, *Tetrahedron*, 1997, **53**, 4063–4082.
21 T. J. Delia, M. J. Olsen and G. B. Brown, *J. Org. Chem.*, 1965, **30**, 2766–2768.
22 J. L. Jambor, D. K. Nordstrom and C. N. Alpers, *Rev. Mineral. Geochem.*, 2000, **40**, 303–350.
23 Y. Liu, X. Zhu, A. Nakamura, R. Orlando, D. Söll and W. B. Whitman, *J. Biol. Chem.*, 2012, **287**, 36683–92.
24 M. Levy and S. L. Miller, *J. Mol. Evol.*, 1999, **48**, 631–637.
25 S. Becker, I. Thoma, A. Deutsch, T. Gehrke, P. Mayer, H. Zipse and T. Carell, *Science*, 2016, **352**, 833–6.
26 S. Becker, C. Schneider, H. Okamura, A. Crisp, T. Amatov, M. Dejmek and T. Carell, *Nat. Commun.*, 2018, **9**, 163.
27 P. A. Limbach, P. F. Crain and J. A. McCloskey, *Nucleic Acids Res.*, 1994, **22**, 2183–2196.