A Journal of the Gesellschaft Deutscher Chemiker

Angewandte International Edition Chemie

GDCh

www.angewandte.org

Accepted Article

Title: Prebiotic methylations and carbamoylations generate non-canonical RNA nucleosides as molecular fossils of an early Earth

Authors: Christina Schneider, Sidney Becker, Hidenori Okamura, Antony Crisp, Tynchtyk Amatov, Michael Stadlmeier, and Thomas Carell

To be cited as: *Angew. Chem. Int. Ed.* 10.1002/anie.201801919
Angew. Chem. 10.1002/ange.201801919

Link to VoR: http://dx.doi.org/10.1002/anie.201801919
http://dx.doi.org/10.1002/ange.201801919

WILEY-VCH

10.1002/anie.201801919

WILEY-VCH

Prebiotic methylations and carbamoylations generate non-canonical RNA nucleosides as molecular fossils of an early Earth

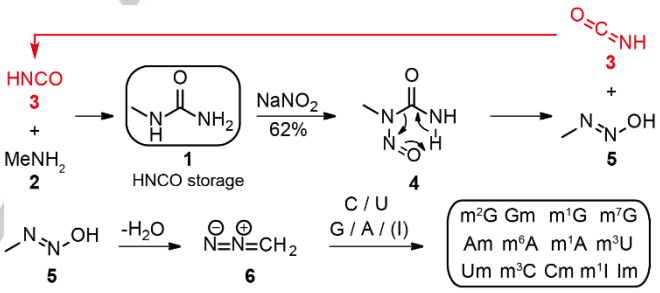
Christina Schneider, Sidney Becker, Hidenori Okamura, Antony Crisp, Tynchtyk Amatov, Michael Stadlmeier and Thomas Carell*

The RNA world hypothesis assumes that life on earth started with small RNA molecules that catalyzed their own formation. Vital to this hypothesis is the need for prebiotic routes towards RNA. Contemporary RNA, however, is not only constructed from the four canonical nucleobases (A, C, G and U), but it contains in addition many chemically modified (non-canonical) bases. A yet open question is if these non-canonical bases were formed in parallel to the canonical bases (chemical origin), or whether they were created later, when life demanded higher functional diversity (biological origin). Here we show that isocyanates in combination with sodium nitrite establish methylating and carbamoylating reactivity compatible with early Earth conditions. This chemistry leads to the formation of methylated and amino acid modified nucleosides that are still extant. Our data provide a plausible scenario for the chemical origin of certain non-canonical bases, which suggests that they are fossils of an early Earth.

More than 120 modified bases were identified in RNA, which are important for the correct folding into complex three-dimensional structures and for the fine tuning of RNA/RNA and RNA/protein interactions.[1-3] Modified nucleosides are for example found in close proximity to the anticodon stem loop in tRNA, where they are involved in translation of the genetic code.[4, 5] Methylated nucleosides such as m^6A are involved in regulating mRNA stability,[6] splicing,[7, 8] translation[9-11] and X chromosome inactivation.[12] Another methylated nucleosides, m^7G, is part of the 5'-cap structure of eukaryotic mRNA.[13]

The RNA world hypothesis postulates that life started with self-replicating RNA molecules that were amenable to process of chemical evolution involving replication, randomization and selection.[14] Since RNA is able to store genetic information and perform catalytic processes, the hypothesis further posits that an early replicating cell could proliferate and maintain a primitive metabolism in the absence of coded proteins. Non-coded polypeptides[15, 16] and simple anabolic pathways[17, 18] may have supported an early RNA based metabolism.

This hypothesis requires the presence of the key building blocks of life such as nucleosides and amino acids or of primitive anabolic processes that led to their formation.[19] This raises the question if life began with only the four canonical bases (A, C, G and U)[20, 21] or if an early preRNA was chemically more diverse,[22] containing non-canonical nucleosides.[22] Those non-canonical bases that are until today found in RNA, may be considered fossils of this early phase of chemical evolution.[23-25] Finding evidence for this idea requires simple chemistry compatible with early Earth geochemical models that generate these non-canonical bases. Here we show that the majority of methylated nucleosides, which play important roles in RNAs of all three domains of life, can be prebiotically generated upon reaction of canonical nucleosides with *N*-methyl-*N*-nitrosourea, which is readily formed from *N*-methylurea and sodium nitrite[26] (Scheme 1).



Scheme 1. Chemistry that leads to the formation of methylated derivatives of canonical nucleobases that are today found in RNA in all three domains of life. Methylurea functions as a storage for reactive isocyanic acid.

NO_2^- was potentially available on the early Earth from NO and NO_2,[27] which are formed during lightning in an N_2 atmosphere.[28] Alternatively, NO can form by reaction of N_2 with CO_2 in hot impact plumes.[29]

Next to the methylated RNA bases, we also find amino acid modified nucleosides among the many contemporary non-canonical RNA bases.[30, 31] They are directly involved in decoding the genetic information.[32, 33] We show that our NO_2^--based chemistry provides these modified bases as well, which suggests an early intimate contact between nucleobases and amino acids that might have been the basis for the co-evolution of RNA and proteins and the establishment of primitive proto-metabolic pathways.

The chemistry starts with methyl urea **1**, which is one of the molecules that was likely present on the early Earth[34]. Methyl urea (**1**) is for example formed by reaction of ammonia with methyl isocyanate which was detected on comet 67P/Churyumov-Gerasimenko.[35, 36] Methyl urea was also

[a] Christina Schneider, Sidney Becker, Dr. Hidenori Okamura, Antony Crisp, Dr. Tynchtyk Amatov, Michael Stadlmeier and Prof. Dr. T. Carell
Center for Integrated Protein Science (CiPS^M) at the Department of Chemistry, LMU München
Butenandtstr. 5-13, 81377 München
Fax: (+) 49 2180 77756
E-mail: Thomas.Carell@lmu.de
Homepage: www.carellgroup.de

Supporting information for this article is given via a link at the end of the document.

shown to form directly in the Urey-Miller experiment[37] and it is available with high yields by the reaction of methylamine **2** with HNCO (**3**) in water (90%).[38] HNCO in turn was detected in interstellar gases,[39] and likewise on comet 67P/Churyumov-Gerasimenko.[35, 40] Urea itself is also known to decompose into ammonium isocyanate.[41, 42] Despite the potential formation of **3**, however, it is difficult to conceive the accumulation of HNCO (**3**) because of its high reactivity. If, however, a little methyl urea (**1**) is present, it can readily react with NO^+ (Scheme 1). Methyl urea (**3**) is easily nitrosylated, which gives *N*-methyl-*N*-nitrosourea (**4**)[26] with a yield of 62%. This compound can physically separate as a foam from the aqueous phase, which allows **4** to potentially accumulate, so that it may have been locally available at high concentrations. Under slightly basic conditions, for example in the presence of borax (reported to be important for ribose forming reactions),[43] **4** quickly decomposes to furnish 1-hydroxy-2-methyldiazene (**5**) under liberation of HNCO (**3**). As such, only small amounts of HNCO are required to help converting $MeNH_2$ and $NaNO_2$ into 1-hydroxy-2-methyldiazene (**5**). 1-Hydroxy-2-methyldiazene (**5**) in turn eliminates water and decomposes to diazomethane (**6**), which is a common methylating agent.[44] Since all starting materials are likely components of the organic matter on the early Earth, it is therefore plausible that diazomethane was an accessible component. The controlled release of **6** from the stable precursor methyl urea (**1**) could have made it available for chemical transformations despite its high reactivity and consequently short half-life time on the early Earth.

When we performed this base-catalyzed diazomethane (**6**) formation in the presence of the canonical nucleobases, we obtained a large set of methylated compounds (Fig. 1). For the experiment we dissolved the nucleosides in a 1:1 mixture of borate buffer and formamide. Formamide is accessible under early Earth conditions through the reaction of HCN and H_2O.[45] *N*-methyl-*N*-nitrosourea (**4**) was then added to the nucleoside mixture in one portion. After one hour at 70°C, samples were taken and analysed by LC-MS and tandem mass spectrometry.

The observed results are depicted in Fig. 1. In order to correctly assign the resulting methylated nucleosides, co-injections with synthetic reference compounds were performed (SI). The products were further elucidated by analysis of the fragmentation patterns in LC-MS^2 experiments. When we performed the reaction in the presence of adenosine, we obtained m^1A, Am and m^6A, together with the 3' and 5' methylated derivatives (marked as m^xA, SI). When guanosine was methylated under the same conditions, we detected m^7G (7%) as well as Gm, m^1G and m^2G, all of which are known non-canonical bases. In the presence of cytidine, the bases m^3C and Cm were generated. Furthermore, the reaction of uridine furnished the methylated compounds Um and m^3U. m^3U is formed in high yield of 11%. We also investigated the methylation reaction with inosine (I) as the hydrolysis product of A.[46, 47] When I was subjected to the same conditions, we detected formation of Im and m^1I (see SI). Importantly, nearly all of the methylated nucleosides we observed are today found in RNAs of all three domains of life.[2, 48]

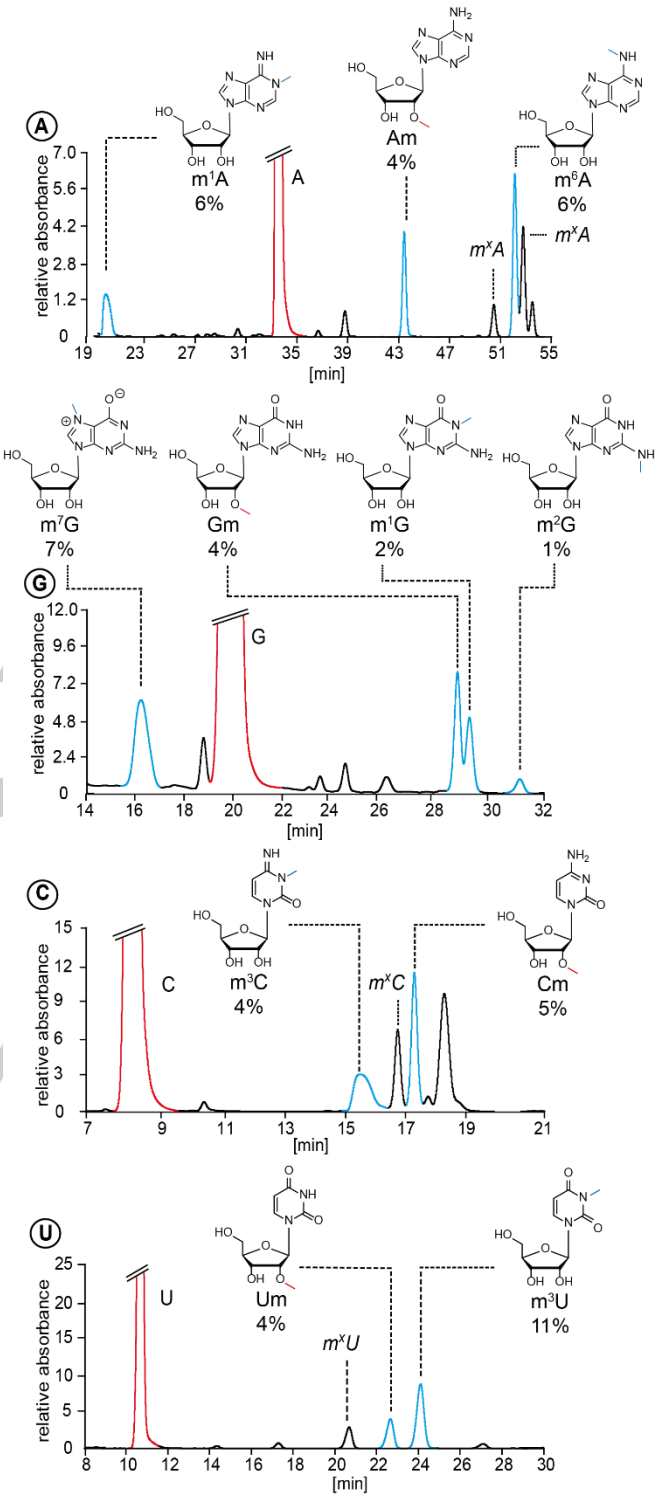


Figure 1. HPLC traces of the reaction mixtures obtained in the reaction of *N*-methyl-*N*-nitrosourea (**1**) in the presence of the canonical nucleobases A, G, C and U. The modified nucleosides are shown in blue, the canonical ones in red. Peaks labeled with "m^x" were identified as sugar modified modifications based on data from fragmentation studies (SI).

We next asked the question of whether the simple chemistry can be used to enable the attachment of larger chemical moieties

such as amino acids to the canonical nucleobases to give RNA modifications such as t^6A and g^6A.

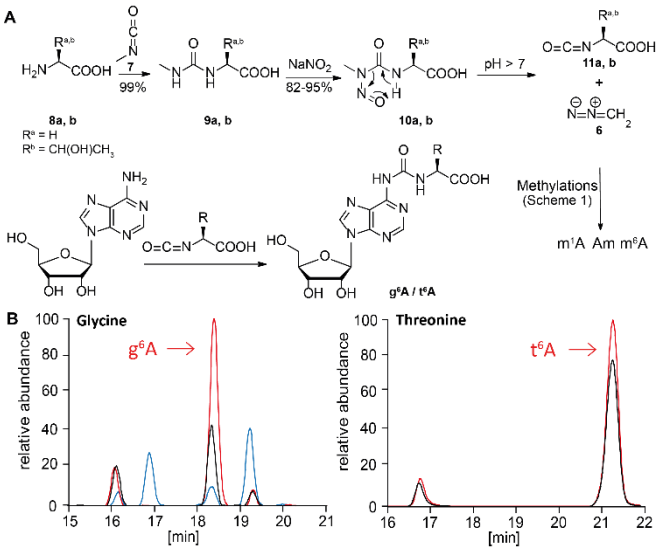


Figure 2. A: Plausible reaction scheme for the prebiotic access to t^6A and g^6A. B: MS-chromatograms of the reactions of *N*-methyl urea derivatives **9** with the canonical nucleoside A under formation of isocyanates of the corresponding amino acids. The chromatogram in blue shows the reaction without the addition of Ni^{2+}-salts and the ones in black represent the reaction in the presence of Ni^{2+}-salts. Co-injections with synthetic standards are shown in red. The additional peaks arise from the reaction of the sugars with the amino acid isocyanate. Selectivity can be increased by the addition of $[Ni(ClO_4)_2]$.

This was indeed possible (Fig. 2), when we replaced HNCO by methyl isocyanate CH_3NCO (**7**). **7** can be generated under prebiotic conditions *via* UV-irradiation of CH_4 and HNCO. In an aqueous environment, we observed that **7** reacts rapidly with amino acids such as glycine (**8a**) and threonine (**8b**) to give the corresponding methyl urea derivatives **9** (Fig. 2) in nearly quantitative yields. The compounds **9a** and **b** can be nitrosylated[49] under the same conditions as methyl urea **1** to form the nitroso compounds **10a** and **10b** in high yields of 82-95%. A pH switch to slightly basic conditions with either phosphate or borate buffer converts the intermediate nitroso compounds **10a**,**b** into the isocyanates of the corresponding amino acids **11a**,**b**. Upon treatment with adenosine, these intermediates react to give the corresponding N^6-derivatives g^6A and t^6A. Since the reaction takes place under basic conditions, not only N^6 but also the 2'-, 3'-, and 5'-hydroxyl groups can react with the isocyanate-derivative of the amino acids (see Fig. 2). Interestingly, the selectivity of the reaction can be controlled to favour the N^6 position by the addition of Ni^{2+}, which is generated during prebiotic nucleoside formation.[22] At the same time CH_2N_2 (**6**) is formed which can facilitate subsequent methylations (see SI).

Interestingly, the amino acid modified nucleosides that are formed as described here, are present today in all three domains of life.[2, 48] Recently, a comparative phylogenetic analysis[50] has suggested that non-canonical bases were likely already present in the ancient parent of all life on Earth, known conventionally as LUCA, the last universal common ancestor. An overlay of the nucleosides accessed in this study with those derived from the genetic analysis shows surprising consensus (Fig. 3). Most of the simple modifications that were present in LUCA could also be formed by the chemistry presented here.

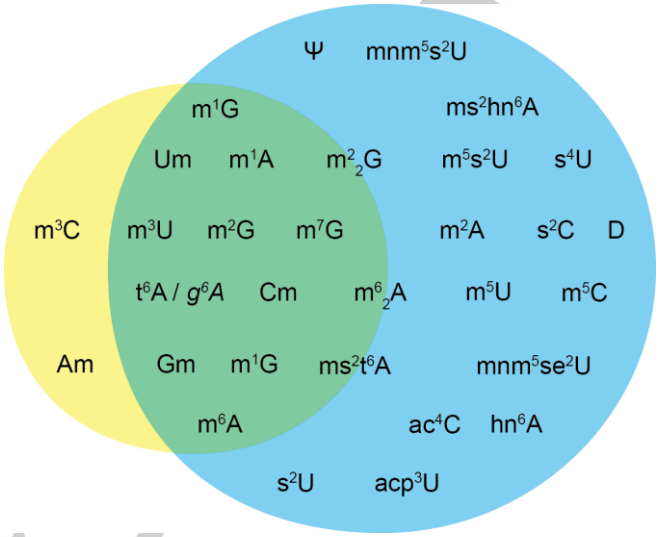


Figure 3 Non-canonical nucleosides generated in this study (yellow area) and non-canonical nucleosides that were found based on phylogenetic analysis to be likely early nucleobases bases in the biosphere (blue area).[50] Modified nucleosides that were found in both studies are shown in the green area. ms^2t^6A, $m^2{}_2G$ and $m^6{}_2A$ were deposited at the border since they could be generated as well using the here described chemistry starting from ms^2A, m^2G and m^6A.

In summary, we report a simple cascade reaction that starts with isocyanic acid, methylisocyanate, methylamine, ammonia and sodium nitrite. In this cascade the unstable molecule isocyanic acid (**3**) is captured by methylamine and stored in form of methyl urea. It can be released under basic conditions from *N*-methyl-*N*-nitrosourea (**4**) which is produced by nitrosylation of methylurea (**1**). The chemistry allows us to convert the canonical pyrimidine and purine bases, for which prebiotically plausible formation processes were recently described[20, 21, 51, 52] into non-canonical nucleosides. As such, the reported results provide chemical evidence that the canonical and many non-canonical ribonucleosides can form spontaneously under plausible prebiotic conditions. The here described chemistry can be linked to the nitrosylation chemistry that was recently reported to enable the parallel formation of canonical and non-canonical bases.[22] The non-canonical bases, particularly the amino acid modified purines, potentially increase the chemical diversity of RNA in order to broaden its folding and catalytic capabilities. This complements ideas that non-canonical base pairs might have existed in pre-RNA.[22, 53]

Acknowledgements

We thank the Deutsche Forschungsgemeinschaft for financial support via SFB1032 (TP-A5), SPP1784, CA275/11-1 and the Excellence Cluster CiPS[M]. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme

(grant agreement n° EPiR 741912). We thank Dr. Markus Müller for comments on this manuscript.

Keywords: prebiotic chemistry • origin of life • modifications • methylated nucleosides • t^6A • m^6A

[1] T. Carell, C. Brandmayr, A. Hienzsch, M. Müller, D. Pearson, V. Reiter, I. Thoma, P. Thumbs, M. Wagner, *Angew. Chem. Int. Ed.* **2012**, *51*, 7110-7131.
[2] M. A. Machnicka, K. Milanowska, O. Osman Oglou, E. Purta, M. Kurkowska, A. Olchowik, W. Januszewski, S. Kalinowski, S. Dunin-Horkawicz, K. M. Rother, *Nucleic Acids Res.* **2012**, *41*, D262-D267.
[3] A. Czerwoniec, S. Dunin-Horkawicz, E. Purta, K. H. Kaminska, J. M. Kasprzak, J. M. Bujnicki, H. Grosjean, K. Rother, *Nucleic Acids Res.* **2008**, *37*, D118-D121.
[4] P. C. Thiaville, B. El Yacoubi, C. Köhrer, J. J. Thiaville, C. Deutsch, D. Iwata‐Reuyl, J. M. Bacusmo, J. Armengaud, Y. Bessho, C. Wetzel, *Mol. Microbiol.* **2015**, *98*, 1199-1221.
[5] J. A. Kowalak, J. J. Dalluge, J. A. McCloskey, K. O. Stetter, *Biochem.* **1994**, *33*, 7869-7876.
[6] X. Wang, Z. Lu, A. Gomez, G. C. Hon, Y. Yue, D. Han, Y. Fu, M. Parisien, Q. Dai, G. Jia, *Nature* **2014**, *505*, 117.
[7] X. Zhao, Y. Yang, B.-F. Sun, Y. Shi, X. Yang, W. Xiao, Y.-J. Hao, X.-L. Ping, Y.-S. Chen, W.-J. Wang, *Cell Res.* **2014**, *24*, 1403.
[8] W. Xiao, S. Adhikari, U. Dahal, Y.-S. Chen, Y.-J. Hao, B.-F. Sun, H.-Y. Sun, A. Li, X.-L. Ping, W.-Y. Lai, *Mol. Cell* **2016**, *61*, 507-519.
[9] K. D. Meyer, D. P. Patil, J. Zhou, A. Zinoviev, M. A. Skabkin, O. Elemento, T. V. Pestova, S.-B. Qian, S. R. Jaffrey, *Cell* **2015**, *163*, 999-1010.
[10] J. Zhou, J. Wan, X. Gao, X. Zhang, S. R. Jaffrey, S.-B. Qian, *Nature* **2015**, *526*, 591.
[11] X. Wang, B. S. Zhao, I. A. Roundtree, Z. Lu, D. Han, H. Ma, X. Weng, K. Chen, H. Shi, C. He, *Cell* **2015**, *161*, 1388-1399.
[12] D. P. Patil, C.-K. Chen, B. F. Pickering, A. Chow, C. Jackson, M. Guttman, S. R. Jaffrey, *Nature* **2016**, *537*, 369.
[13] N. Sonenberg, M. A. Morgan, W. C. Merrick, A. J. Shatkin, *Proc. Natl. Acad. Sci. USA* **1978**, *75*, 4843-4847.
[14] A. C. Ferretti, G. F. Joyce, *Biochem.* **2013**, *52*, 1227-1235.
[15] S. Tagami, J. Attwater, P. Holliger, *Nat. Chem.* **2017**, *9*, 325.
[16] J. G. Forsythe, S. S. Yu, I. Mamajanov, M. A. Grover, R. Krishnamurthy, F. M. Fernández, N. V. Hud, *Angew. Chem. Int. Ed.* **2015**, *54*, 9871-9875.
[17] K. B. Muchowska, S. J. Varma, E. Chevallot-Beroux, L. Lethuillier-Karl, G. Li, J. Moran, *Nat. Ecol. Evol.* **2017**, *1*, 1716.
[18] A. J. Coggins, M. W. Powner, *Nat. Chem.* **2017**, *9*, 310-317.
[19] O. Leslie E, *Crit. Rev. Biochem. Mol.* **2004**, *39*, 99-123.
[20] S. Becker, I. Thoma, A. Deutsch, T. Gehrke, P. Mayer, H. Zipse, T. Carell, *Science* **2016**, *352*, 833-836.
[21] M. W. Powner, B. Gerland, J. D. Sutherland, *Nature* **2009**, *459*, 239-242.
[22] S. Becker, C. Schneider, H. Okamura, A. Crisp, T. Amatov, M. Dejmek, T. Carell, *Nat. Commun.* **2018**, *9*, 163.
[23] M. P. Robertson, G. F. Joyce, *Cold Spring Harb Perspect Biol* **2012**, *4*, a003608.
[24] B. J. Cafferty, D. M. Fialho, J. Khanam, R. Krishnamurthy, N. V. Hud, *Nat. Commun.* **2016**, *7*, 11328.
[25] A. C. Rios, Y. Tor, *Isr. J. Chem.* **2013**, *53*, 469-483.
[26] A. Tsubura, Y.-C. Lai, H. Miki, T. Sasaki, N. Uehara, T. Yuri, K. Yoshizawa, *in vivo* **2011**, *25*, 11-22.
[27] Y. L. Yung, M. B. Mcelroy, *Science* **1979**, *203*, 1002-1004.
[28] W. L. Chameides, J. C. G. Walker, *Origins of life* **1981**, *11*, 291-302.
[29] J. F. Kasting, *Orig. Life Evol. Biospheres* **1990**, *20*, 199-231.
[30] G. B. Chheda, *Life Sci.* **1969**, *8*, 979-987.
[31] C. Deutsch, B. El Yacoubi, V. de Crécy-Lagard, D. Iwata-Reuyl, *J. Biol. Chem.* **2012**, *287*, 13666-13673.
[32] S. E. Kolitz, J. R. Lorsch, *FEBS Lett.* **2010**, *584*, 396-404.
[33] F. V. Murphy IV, V. Ramakrishnan, A. Malkiewicz, P. F. Agris, *Nat. Struct. Mol. Biol.* **2004**, *11*, 1186.
[34] T. G. Waddell, L. L. Eilders, B. P. Patel, M. Sims, *Orig. Life Evol. Biosph.* **2000**, *30*, 539-548.
[35] F. Goesmann, H. Rosenbauer, J. H. Bredehöft, M. Cabane, P. Ehrenfreund, T. Gautier, C. Giri, H. Krüger, L. Le Roy, A. J. MacDermott, *Science* **2015**, *349*, aab0689.
[36] N. F. W. Ligterink, A. Coutens, V. Kofman, H. S. P. Müller, R. T. Garrod, H. Calcutt, S. F. Wampfler, J. K. Jørgensen, H. Linnartz, E. F. van Dishoeck, *Mon. Notices Royal Astron. Soc.* **2017**, *469*, 2219-2229.
[37] S. L. Miller, H. C. Urey, *Science* **1959**, *130*, 245-251.
[38] B. T. Golding, C. Bleasdale, J. McGinnis, S. Müller, H. T. Rees, N. H. Rees, P. B. Farmer, W. P. Watson, *Tetrahedron* **1997**, *53*, 4063-4082.
[39] K. Altwegg, H. Balsiger, A. Bar-Nun, J.-J. Berthelier, A. Bieler, P. Bochsler, C. Briois, U. Calmonte, M. R. Combi, H. Cottin, *Sci. Adv.* **2016**, *2*, e1600285.
[40] P. Ehrenfreund, S. B. Charnley, *Annu. Rev. Astron. Astrophys.* **2000**, *38*, 427-483.
[41] W. H. Shaw, J. J. Bordeaux, *J. Am. Chem. Soc.* **1955**, *77*, 4729-4733.
[42] B. P. Callahan, Y. Yuan, R. Wolfenden, *J. Am. Chem. Soc.* **2005**, *127*, 10828-10829.
[43] H.-J. Kim, A. Ricardo, H. I. Illangkoon, M. J. Kim, M. A. Carrigan, F. Frye, S. A. Benner, *J. Am. Chem. Soc.* **2011**, *133*, 9457-9468.
[44] K. K. Park, J. S. Wishnok, M. C. Archer, *Chemico-Biological Interactions* **1977**, *18*, 349-354.
[45] R. Saladino, C. Crestini, S. Pino, G. Costanzo, E. Di Mauro, *Phys. Life Rev.* **2012**, *9*, 84-104.
[46] R. A. Sanchez, L. E. Orgel, *J. Mol. Biol.* **1970**, *47*, 531-543.
[47] W. D. Fuller, R. A. Sanchez, L. E. Orgel, *J. Mol. Biol.* **1972**, *67*, 25-33.
[48] M. A. Machnicka, K. Milanowska, O. Osman Oglou, E. Purta, M. Kurkowska, A. Olchowik, W. Januszewski, S. Kalinowski, S. Dunin-Horkawicz, K. M. Rother, M. Helm, J. M. Bujnicki, H. Grosjean, *Nucleic Acids Res.* **2013**, *41*, D262-D267.
[49] T. Machinami, T. Suami, *Bull. Chem. Soc. Jpn.* **1975**, *48*, 1333-1334.
[50] M. C. Weiss, F. L. Sousa, N. Mrnjavac, S. Neukirchen, M. Roettger, S. Nelson-Sathi, W. F. Martin, *Nat. Microbiol.* **2016**, *1*, 16116.
[51] H.-J. Kim, S. A. Benner, *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 11315-11320.
[52] R. Saladino, E. Carota, G. Botta, M. Kapralov, G. N. Timoshenko, A. Y. Rozanov, E. Krasavin, E. Di Mauro, *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E2746-E2755.
[53] C. B. Winiger, M.-J. Kim, S. Hoshika, R. W. Shaw, J. D. Moses, M. F. Matsuura, D. L. Gerloff, S. A. Benner, *Biochem.* **2016**, *55*, 3847-3850.