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*

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

* Corresponding author
E-mail: thomas.carell@lmu.de

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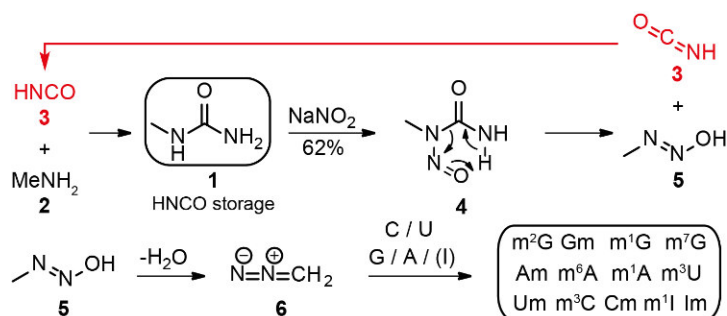
Scheme 1. Chemistry that leads to the formation of methylated derivatives of canonical nucleobases that are today found in RNA in all three domains of life. Methylurea functions as a storage for reactive isocyanic acid.

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 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_2^-(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^{1}A, Am and m^{6}A, together with the 3' and 5' methylated derivatives (marked as m^{x}A, SI). When guanosine was methylated under the same conditions, we detected m^{7}G (7%) as well as Gm, m^{1}G and m^{2}G, all of which are known non-canonical bases. In the presence of cytidine, the bases m^{3}C and Cm were generated. Furthermore, the reaction of uridine furnished the methylated compounds Um and m^{3}U. m^{3}U 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^{1}I (see SI). Importantly, nearly all of the methylated nucleosides we observed are today found in RNAs of all three domains of life.[2, 48]
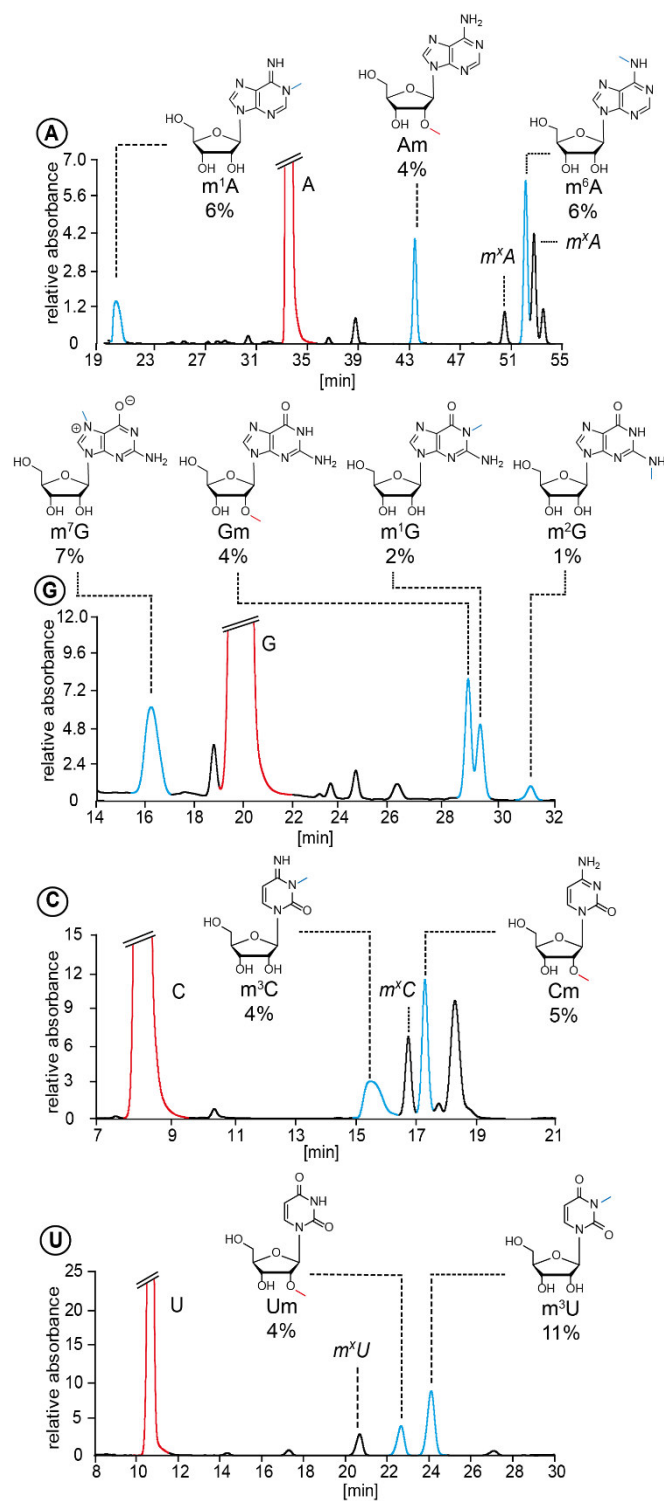
Figure 1. HPLC traces of the reaction mixtures obtained in the reaction of *N*-methyl-*N*-nitrosourea (**1**) in the presence of the canonical nucleobases A, G, C and U. The modified nucleosides are shown in blue, the canonical ones in red. Peaks labeled with "m^x" were identified as sugar modified modifications 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^{6}A and g^{6}A.
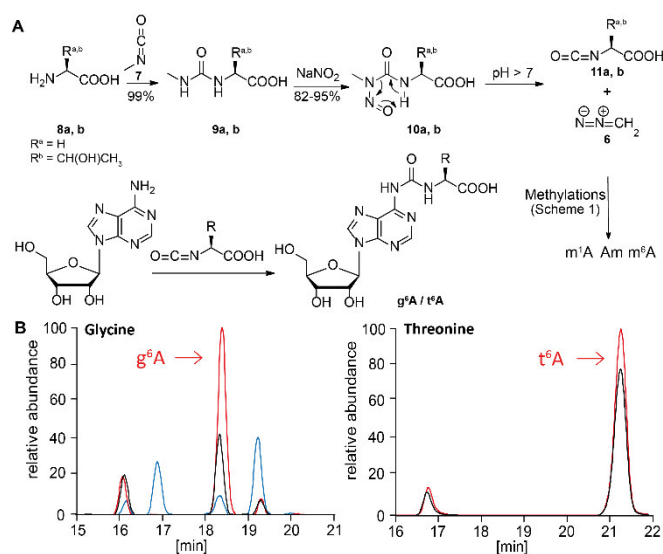
Figure 2. A: Plausible reaction scheme for the prebiotic access to t^6A and g^6A. B: MS-chromatograms of the reactions of *N*-methyl urea derivatives **9** with the canonical nucleoside A under formation of isocyanates of the corresponding amino acids. The chromatogram in blue shows the reaction without 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).
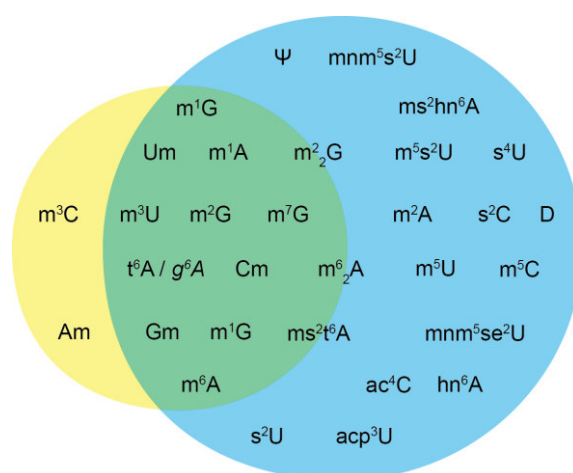

Figure 3 Non-canonical nucleosides generated in this study (yellow area) and non-canonical nucleosides that were found based on phylogenetic analysis to be likely early nucleobases bases in the biosphere (blue area).[50] Modified nucleosides that were found in both studies are shown in the green area. 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.

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.

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]

[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.

Supporting information

Materials and Methods

<u>General Information</u>

Chemicals were purchased from Sigma-Aldrich, Fluka, ABCR, Carbosynth or Acros Organics and used without further purification. Solutions were concentrated *in vacuo* on a Heidolph rotary evaporator. The solvents were of reagent grade or purified by distillation. Chromatographic purification of products was accomplished using flash column chromatography on Merck Geduran Si 60 (40-63 µm) silica gel (normal phase). Thin layer chromatography (TLC) was performed on Merck 60 (silica gel F254) plates. Visualization of the developed chromatogram was performed using fluorescence quenching or standard staining solutions. ^{1}H- and ^{13}C-NMR spectra were recorded in deuterated solvents on Varian Oxford 200, Bruker ARX 300, Varian VXR400S, Varian Inova 400, Bruker AMX 600 and Bruker AVIIIHD 400 spectrometers and calibrated to the residual solvent peak. Multiplicities are abbreviated as follows: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, br = broad. High-resolution ESI spectra were obtained on the mass spectrometers Thermo Finnigan LTQ FT-ICR. IR measurements were performed on Perkin Elmer Spectrum BX FT-IR spectrometer with a diamond-ATR (Attenuated Total Reflection) setup. Melting points were measured on a Büchi B-540. For preparative HPLC purification a Waters 1525 binary HPLC Pump in combination with a Waters 2487 Dual Absorbance Detector was used, with Macherey-Nagel VP 250/10 Nucleosil 100-7 C18 reversed phase column. The analysis of the prebiotic reactions were analyzed by LC-ESI-MS on a Thermo Finnigan LTQ Orbitrap XL and were chromatographed by a Dionex Ultimate 3000 HPLC system with a flow of 0.15 mL/min over an Interchim Uptisphere120-3HDO C18 column. The column temperature was maintained at 30 °C. Eluting buffers were buffer A (2 mM HCOONH$_4$ in H2O (pH 5.5)) and buffer B (2 mM HCOONH$_4$ in H$_2$O/MeCN 20/80 (pH 5.5)). For samples containing methylated and aminoacylated nucleosides the following gradient was used: 0 → 55 min, 0% → 8% HPLC/MS Puffer B; 55 → 70 min, 8% → 60% HPLC/MS Puffer B; 70 → 72 min, 60% → 100% HPLC/MS Puffer B; 72 → 78 min, 100% HPLC/MS Puffer B; 78 → 81 min, 100 → 0% HPLC/MS Puffer B; 81 → 90 min, 0% HPLC/MS Puffer B. The elution was monitored at 260 nm (Dionex Ultimate 3000 Diode Array Detector). The chromatographic eluent was directly injected into the ion source without prior splitting. Ions were scanned by use of a positive polarity mode over a full-scan range of m/z 120-1000 with a resolution of 30000. Parameters of the mass spectrometer were tuned with a freshly mixed solution of adenosine (5 µM). The parameters used in this section were sheath gas flow rate, 16 arb; auxiliary gas flow rate, 11 arb; sweep gas flow rate, 4 arb; spray voltage, 5.0 kV; capillary temperature, 200 °C; capillary voltage, 20 V, tube lens 65 V.

Synthethis of the methylated nucleosides under prebiotic conditions:

Methylated Adenosine

Adenosine (6.00 mg, 22.5 µmol, 1 eq.) was dissolved in a suspension of borax (105 mg, 275 µmol, 12 eq) in H$_2$O (250 µL) and formamide (100 µL).1-methyl-1-nitrosourea (80.0 mg, 559 µmol, 25 eq.) was dissolved in formamide (150 µL). The suspension was stirred for 18 h at 50°C before 25 µL of the mixture were diluted into 975 µL H$_2$O and analyzed with LC-MS. The modifications were confirmed via coinjection of the prebiotic samples with synthetic standards (see Fig. S4). Yields of the prebiotic reactions were determined with calibration curves (see Fig.S6 and S7).

Methylated Guanosine

Guanosine (6.40 mg, 22.5 µmol, 1 eq.) was dissolved in a suspension of borax (105 mg, 275 µmol, 12 eq) in H$_2$O (250 µL) and formamide (100 µL).1-methyl-1-nitrosourea (80.0 mg, 559 µmol, 25 eq.) was dissolved in formamide (150 µL). The suspension was stirred for 18 h at 50°C before 25 µL of the mixture were diluted into 975 µL H$_2$O and analyzed with LC-MS. The modifications were confirmed via coinjection of the prebiotic

samples with synthetic standards (see Fig. S4). Yields of the prebiotic reactions were determined with calibration curves (see Fig.S6 and S7).

Methylated Cytidine

Cytidine (5.50 mg, 22.5 µmol, 1 eq.) was dissolved in a suspension of borax (105 mg, 275 µmol, 12 eq) in H_2O (250 µL) and formamide (100 µL).1-methyl-1-nitrosourea (80.0 mg, 559 µmol, 25 eq.) was dissolved in formamide (150 µL). The suspension was stirred for 18 h at 50°C before 25 µL of the mixture were diluted into 975 µL H_2O and analyzed with LC-MS. The modifications were confirmed via coinjection of the prebiotic samples with synthetic standards (see Fig. S4). Yields of the prebiotic reactions were determined with calibration curves (see Fig.S6 and S7).

Methylated Uridine

Uridine (5.50 mg, 22.5 µmol, 1 eq.) was dissolved in a suspension of borax (105 mg, 275 µmol, 12 eq) in H_2O (250 µL) and formamide (100 µL).1-methyl-1-nitrosourea (80.0 mg, 559 µmol, 25 eq.) was dissolved in formamide (150 µL). The suspension was stirred for 18 h at 50°C before 25 µL of the mixture were diluted into 975 µL H_2O and analyzed with LC-MS. The modifications were confirmed via coinjection of the prebiotic samples with synthetic standards (see Fig. S4). Yields of the prebiotic reactions were determined with calibration curves (see Fig.S6 and S7).

Methylated Inosine

Inosine (6.00 mg, 22.5 µmol, 1 eq.) was dissolved in a suspension of borax (105 mg, 275 µmol, 12 eq) in H_2O (250 µL) and formamide (100 µL).1-methyl-1-nitrosourea (80.0 mg, 559 µmol, 25 eq.) was dissolved in formamide (150 µL). The suspension was stirred for 18 h at 50°C before 25 µL of the mixture were diluted into 975 µL H_2O and analyzed with LC-MS. The modifications were confirmed via coinjection of the prebiotic samples with synthetic standards (see Fig. S4). Yields of the prebiotic reactions were determined with calibration curves (see Fig.S6 and S7).

Synthesis of the amino acid modified nucleosides under prebiotic conditions:

Without Nickel:

Adenosine (2.67 mg, 10.0 µmol, 1 eq.) was dissolved in phosphate buffer (370 µL, 30 mM, pH 8) and H_2O. (40 µL). The (methyl(nitroso)carbamoyl)amino acid (Glycine: 3.32 mg, 20.9 µmol, 2 eq; Threonine: 4.10 mg, 20.9 µmol, 2 eq) was suspended in water (40 µL) and was slowly added to the reaction mixture. After 17 h at 70°C 50 µL samples were removed and diluted to 1 mL. The samples were used for LC-MS analysis. The modifications were confirmed via coinjection of the prebiotic samples with synthetic standards

With Nickle:

Adenosine (2.67 mg, 10.0 µmol, 1 eq.) was dissolved in phosphate buffer (370 µL, 30 mM, pH 8). [Ni(ClO$_4$)$_2$] x 6 H_2O (91.4 mg, 250 µmol, 25 eq) was added in H_2O (40 µL). The (methyl(nitroso)carbamoyl)amino acid (Glycine: 3.32 mg, 20.9 µmol, 2 eq; Threonine: 4.10 mg, 20.9 µmol, 2 eq) was suspended in water (40 µL) and was slowly added to the reaction mixture. After 17 h at 70°C 50 µl samples were removed and diluted to 1 mL. The samples were used for LC-MS analysis. The modifications were confirmed via coinjection of the prebiotic samples with synthetic standards

Prebiotic access to Urea und Urea-Nitroso-Compounds

1-methylurea (1)
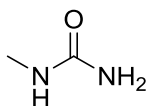
An aqueous solution of ammonia (100 µL, 25%) was added to H_2O (150 µL) and cooled to 0°C. Methylisocyanate (20.0 µL, 326 µmol, 1.0 eq) was added and the mixture was stirred at 20°C. After 3.5 h the solvent was evaporated and the product could be obtained as a white solid (23.92 mg, 322 µmol, 99%).
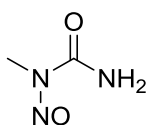
^{1}H-NMR (600 MHz, CD_3CN) δ = 5.04 (bs, 1H, NH), 4.71 (s, 2H, NH_2), 2.60 (d, 3H, NCH_3).
^{13}C-NMR (600 MHz, CD_3CN) δ = 160.32 (1C, CO), 26.94 (s, 1C, NCH_3).
HRMS (EI$^+$): calc. for [$C_2H_5N_3O_2$] 103.0383, found: 103.0374 [M].
IR (cm^{-1}): ṽ = 3412 (w), 3320 (m), 3203 (w), 1650 (m), 1625 (m), 1565 (s), 1420 (m), 1348 (m), 1167 (m), 1024 (s), 1004 (m), 657 (m).

1-methyl-1-nitrosourea (4)

The reaction was carried out slightly modified to Arndt et al.[1].

Methylurea (1.00 g, 13.5 mmol, 1 eq) was dissolved in H_2O (8 mL). $NaNO_2$ (1.02 g, 14.9 mmol, 1.1 eq) was added at 0°C and aq. HCl (1.9 mL, 18.5 mmol, 1.4 eq) was added dropwise to the reaction mixture. The precipitated solid was filtered off, washed with H_2O and dried in vacuo. The product was obtained as a white solid (0.86 g, 8.31 mmol, 62%).
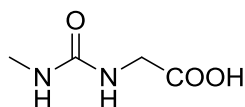
^{1}H-NMR (400 MHz, DMSO-d_6) δ = 8.15 (s, 1H, NH), 7.82 (s, 1H, NH), 3.06 (s, 3H, NCH_3). **^{13}C-NMR** (400 MHz, DMSO-d_6) δ = 150.08 (1C, CO), 48.64 (s, 1C, NCH_3).
HRMS (EI$^+$): calc. for [$C_2H_5N_3O_2$] 103.0383, found: 103.0374 [M].
mp: 108°C
IR (cm-1): ṽ = 3377 (w), 3245 (w), 1722 (m), 1604 (m), 1457 (m), 1417 (s), 1369 (m), 1206 (m), 1086 (m), 974 (s), 841 (m), 775 (m), 641 (s).

(Methylcarbamoyl)glycine (9a)

The reaction was carried out slightly modified to Machinami et al. [2].

Glycine (150 mg, 2.00 mmol) was dissolved in 10% NaOH (1mL) and methylisocyanate (150 µL, 2.42 mmol, 1.2 eq; prepared by thermal cleavage of 3-methyl-1,1-diphenylurea[3]) was added to the solution at 0°C. The reaction mixture was stirred for 3 hours at room temperature before the pH was adjusted to 2-3 with Amberlite IR-120 (H$^+$-type). The reaction mixture was freeze-dried to obtain a white powder (261 mg, 1.98 mmol, 99%). The crude product could be used without further purification.

^{1}H-NMR (400 MHz, DMSO-d_6) δ 6.10 (t, J = 5.8 Hz, 1H, NHCH$_2$), 6.02 (q, J = 4.8 Hz, 1H, NHCH$_3$), 5.75 (s, 1H), 3.65 (d, J = 5.5 Hz, 2H, CH_2), 2.54 (d, J = 4.6 Hz, 3H, NCH_3).

^{13}**C NMR** (101 MHz, DMSO) δ 173.10 (1C, COOH), 158.99 (1C, CON$_2$), 42.14 (1C, CH$_2$), 26.77 (1C, NCH$_3$).

HRMS (ESI$^-$): calc. for [C$_4$H$_7$N$_2$O$_3$]$^-$: 131.0462, found: 131.0462 [M-H]$^-$.

IR (cm^{-1}): ṽ = 3354 (w), 2976 (w), 2936 (w), 2359 (br., w), 1708 (m), 1623 (s), 1552 (s), 1413 (m), 1223 (m), 1185 (s), 1185 (m), 1122 (s) 1077 (w), 1014 (w) 949 (w), 904 (w), 862 (w), 761 (w).

3-hydroxy-2-(3-methylureido)butanoic acid (9b)
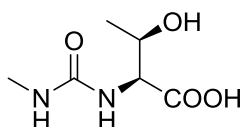
The reaction was carried out slightly modified to Machinami *et al.* [2].

Threonine (238 mg, 2.00 mmol) was dissolved in 10% NaOH (1 mL) and methylisocyanate (150 μl, 2.42 mmol, 1.2 eq; prepared by thermal cleavage of 3-methyl-1,1-diphenylurea[3]) was added to the solution at 0°C. The reaction mixture was stirred for 3 hours at room temperature before the pH was adjusted to 2-3 with Amberlite IR-120 (H$^+$-type). The reaction mixture was freeze-dried to obtain a white powder (348 mg, 1.98 mmol, 99%). The crude product could be used without further purification.

^{1}**H-NMR** (400 MHz, D$_2$O) δ = 4.32 (m, 1H, CHOH), 4.23 (m, 1H, CHNHCOOH), 2.69 (s, 3H, NCH$_3$), 1.19 (d, *J*=6.4 Hz, 3H, CHCH$_3$).

^{13}**C-NMR** (400 MHz, DMSO-d$_6$) δ = 173.58 (s, 1C, COOH), 158.66 (s, 1C, N$_2$CO), 65.31 (s, 1C, CHCOOH), 56.67 (s, 1C, CHOH), 24.01 (s, 1C, NCH$_3$), 16.58 (s, 1C, CHCH$_3$).

HRMS (ESI$^-$): calc. for [C$_6$H$_{11}$N$_2$O$_4$]$^-$ 175.0724, found: 175.0723 [M-H]$^-$.

IR (cm^{-1}): ṽ = 3368 (w), 2976 (w), 2938 (w), 1701 (s), 1528 (m), 1483 (m), 1405 (m), 1224 (m), 1163 (s), 1127 (m), 1079 (m) 981 (s), 923 (m), 841 (w), 781 (s), 663 (m).

(Methyl(nitroso)carbamoyl)glycine (10a)
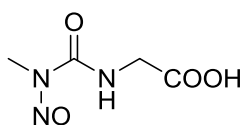
The reaction was carried out slightly modified to Machinami *et al.* [2].

(Methylcarbamoyl)glycine (754 mg, 5.71 mmol, 1 eq) was dissolved in water (10 mL) and acetic acid (1.0 mL). NaNO$_2$ (393 mg, 5.71 mmol, 1 eq) was added at 0°C. The reaction mixture was stirred for 4 h before the pH was adjusted to 3 using Dowex 50WX8-100. The reaction mixture was freeze-dried to obtain a light-yellow solid (735 mg, 4.68 mmol, 82%). The crude product could be used without further purification.

^{1}**H-NMR** (400 MHz, DMSO-d$_6$) δ = 9.00 (t, *J* = 6.0 Hz, 1H, NH), 3.94 (d, *J* = 6.0 Hz, 2H, CH$_2$), 3.11 (s, 2H, NCH$_3$).
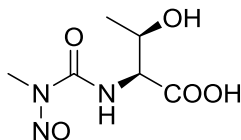
^{13}**C-NMR** (101 MHz, DMSO-d$_6$) δ = 170.85 (s, 1C, COOH), 153.47 (s, 1C, CONHCH$_3$), 42.01 (s, 1C, CH$_2$), 27.04 (s, 1C, NCH$_3$).

HRMS (ESI$^-$): calc. for [C$_4$H$_6$N$_3$O$_4$]$^-$ 160.0364, found: 160.0362 [M-H]$^-$.

mp: 115°C

IR (cm⁻¹): ṽ = 3310 (w), 2943 (w), 2578 (w), 1723 (m), 1689 (s), 1537 (m), 1489 (s), 1434 (m), 1404 (m), 1348 (w), 1258 (m), 1174 (s) 1109 (w), 1067 (m), 1025 (m), 922 (s), 988 (m), 849 (m) 786 (m), 766 (m), 703 (m).

3-hydroxy-2-(3-methyl-3nitroureido)butanoic acid (10b)

The reaction was carried out slightly modified to Machinami *et al.* [2].

3-hydroxy-2-(3-methylureido)butanoic acid (99.0 mg, 0.56 mmol, 1 eq) was dissolved in water (5 mL) and acetic acid (0.5 mL). NaNO$_2$ (58.2 mg, 0.84 mmol, 1.5 eq) was added at 0°C. The reaction mixture was stirred for 4 h before the pH was adjusted to 3 using Dowex 50WX8-100. The reaction mixture was freeze-dried to obtain a yellow oil (110 mg, 0.54 mmol, 95%). The crude product could be used without further purification.

¹H-NMR (400 MHz, DMSO-d$_6$) δ = 8.05 (d, *J*=1.20, 1H, NH), 4.36 (dd, *J* = 8.7, 3.1 Hz, 1H, CHOH), 4.25 (qd, *J* = 6.4, 3.1 Hz, 1H CHNHCOOH), 3.31 (s, 3H, NCH$_3$), 1.18 (d, *J*=6.4, 3H, CHCH$_3$).
**¹³C-NMR** (400 MHz, DMSO-d$_6$) δ = 172.08 (s, 1C, COOH), 153.72 (s, 1C, N$_2$CO), 66.69 (s, 1C, CHCOOH), 59.90 (s, 1C, CHOH), 27.51 (s, 1C, NCH$_3$), 21.04 (s, 1C, CHCH$_3$).
HRMS (ESI⁻): calc. for [C$_6$H$_{10}$N$_3$O$_5$]⁻ 204.0626, found: 204.0624 [M-H]⁻.
IR (cm⁻¹): ṽ = 3394 (w), 2979 (w), 2938 (w), 1701 (m), 1528 (s), 1483 (m), 1405 (s), 1350 (m), 1291 (m), 1224 (w), 1163 (m), 1127 (s) 1079 (w), 981 (m), 923 (m), 874 (s), 841 (m), 781 (m) 753 (m), 663.

Synthetic access to amino acid-urea-compounds

N-methyl-1H-imidazole-1-carboxamide (12)
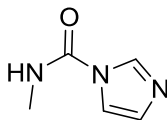
The reaction was carried out slightly modified to Duspara *et al.* [4].

CDI (15.0 g, 92.5 mmol, 1.1 eq) was dissolved dry DMF (14 mL) and dry ACN (41 mL). MeNH$_2$Cl (5.68 mg, 84.1 mmol, 1.5 eq) was added. After stirring the mixture at rt for 3 h the solvent was removed *in vacuo*. The crude product was purified by column chromatography (DCM : MeOH, 19:1 → 17:3) to afford the product as a white solid (9.87 g, 78.9 mmol, yield: 94%).

¹H-NMR (400 MHz, CDCl$_3$) δ = 8.20 (bs, 1H), 7.73 (d, *J* = 3.6 Hz, 1H, NH), 7.48 (t, *J* = 1.5 Hz, 1H), 7.01 (dd, *J* = 1.5, 0.9 Hz, 1H), 2.99 (d, *J* = 4.6 Hz, 3H, CH$_3$).
¹³C NMR (101 MHz, CDCl$_3$) δ 149.81 (s, 1C, CO), 135.87 (s, 1H, C$_{Ar}$), 129.74 (s, 1H, C$_{Ar}$), 116.73 (s, 1H, C$_{Ar}$), 27.49 (s, 1C, NCH$_3$).
HRMS (ESI⁺): calc. for [C$_5$H$_8$N$_3$O]⁺: 126.0662, found: 126.0664 [M+H]⁺.

Ethyl (methylcarbamoyl)glycinate (13)
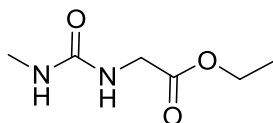
N-methyl-1H-imidazole-1-carboxamide **12** (3.00 g, 23.9 mmol, 2.0 eq) was dissolved in DCM (60 mL)). Ethyl glycinate hydrochloride (1.67 g, 11.9 mmol, 1.5 eq) and NEt$_3$ (3.32 mL, 23.9 mmol, 2.0 eq) were added. After stirring the mixture at rt for 18 h the solvent was removed *in vacuo*. The crude product was purified by column chromatography (DCM : MeOH, 9:1, 0.1 % NEt$_3$) to afford the product as a white solid (9.87 g, 78.9 mmol, yield: 94%).

^{1}H-NMR (400 MHz, DMSO-d$_6$) δ = 6.21 (t, *J* = 6.0 Hz, 1H, NH), 6.01 (q, *J* = 4.8 Hz, 1H, NHCH$_3$), 4.07 (q, *J* = 7.1 Hz, 2H, CH$_2$), 3.73 (d, *J* = 6.0 Hz, 2H, CH$_2$NH), 2.54 (d, *J* = 4.6 Hz, 3H, NHCH$_3$), 1.18 (t, *J* = 7.1 Hz, 3H, CH$_3$).
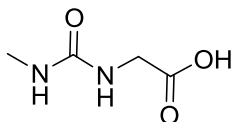**^{13}C-NMR** (101 MHz, DMSO-d$_6$) δ = 171.31 (s, 1C, COOEt), 158.53 (s, 1C, N$_2$CO), 60.17 (s, 1C, OCH$_2$), 41.56 (s, 1C, NHCH$_2$), 26.36 (s, 1C, NCH$_3$), 14.16 (s, 1C, CH$_3$).
HRMS (ESI$^+$): calc. for [C$_6$H$_{13}$N$_2$O$_3$]$^+$: 161.0924, found: 161.0921 [M+H]$^+$.
mp: 83 °C
IR (cm^{-1}): ṽ = 3341 (m), 2936 (w), 1754 (m), 1635 (m), 1575 (s), 1479 (m), 1456 (w), 1424 (w), 1407 (m), 1370 (m), 1283 (m), 1183 (s), 1115 (m) 1065 (m), 1019 (s), 910 (w), 870 (w), 779 (w).

(Methylcarbamoyl)glycine (9a)

Ethyl (methylcarbamoyl)glycinate **13** (0.98 g, 6.12 mmol, 1.0 eq) was dissolved in 10% KOH (5.0 mL). After stirring the mixture at rt 1 h it was treated with Amberlite IR 120 (H$^+$ type) and the solvent was removed in vacuo. The product could be afforded as a white solid (0.53 g, 3.99 mmol, yield: 65%).

^{1}H-NMR (400 MHz, DMSO-d$_6$) δ = 6.12 (t, *J* = 5.9 Hz, 1H, NHCH$_2$), 6.01 (q, *J* = 4.9 Hz, 1H, NHCH$_3$), 3.67 (q, *J* = 5.9 Hz, 2H, CH$_2$), 2.54 (d, *J* = 4.6 Hz, 3H, NCH$_3$).
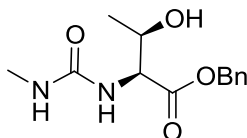**^{13}C-NMR** (101 MHz, DMSO-d$_6$) δ = 172.73 (s, 1C, COOH), 158.64 (s, 1C, N$_2$CO), 41.60 (s, 1C, CH$_2$), 26.39 (s, 1C, NCH$_3$).
HRMS (ESI$^-$): calc. for [C$_4$H$_7$N$_2$O$_3$]$^-$: 131.0462, found: 131.0462 [M-H]$^-$.
mp: 144°C
IR (cm^{-1}): ṽ = 3390 (m), 3367 (m), 2363 (br., w), 1959 (br., w), 1707 (m), 1600 (s), 1550 (s), 1496 (m), 1442 (s), 1414 (m), 1354 (m), 1234 (s) 1087 (m), 1040 (m), 1005 (m), 912 (m), 772 (w), 740 (w).

Benzyl (methylcarbamoyl)-*L*-threoninate (15)

N-methyl-1H-imidazole-1-carboxamide **12** (1.50 g, 11.9 mmol, 2.0 eq) was dissolved in DCM (30 mL). Threonine benzyl ester hydrochloride (1.48 g, 5.98 mmol, 1 eq) and NEt$_3$ (1.80 mL, 11.9 mmol, 2.0 eq) were added. After stirring the mixture at rt for 18 h the solvent was removed *in vacuo*. The crude product was purified by column chromatography (MeOH : Hex : EtOAc, 2:4.5:4.5) to afford the product as a white solid (0.78 mg, 4.45 mmol, yield: 74%).

^{1}H-NMR (400 MHz, DMSO-d$_6$) δ = 7.41 – 7.30 (m, 5H, H$_{Ar}$), 6.23 (q, *J* = 4.6 Hz, 1H, NHCH$_3$), 6.08 (d, *J* = 9.0 Hz, 1H, NH), 5.12 (s, 2H, CH$_2$C$_{Ph}$), 5.12 (d, *J* = 4.6 Hz, 1H, OH), 4.13-4.19 (m, 2H, CHOH, CHNH), 2.56 (d, *J* = 4.6 Hz, 3H, NCH$_3$), 1.07 (d, *J* = 6.1 Hz, 3H, CH$_3$).
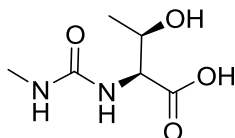**^{13}C-NMR** (101 MHz, DMSO-d$_6$) δ = 172.02 (s, 1C, COOH), 158.70 (s, 1C, N$_2$CO), 136.20 (s, 1C, C$_{Ph}$), 128.41 (s, 2C, C$_{Ph}$), 127.93 (s, 1C, C$_{Ph}$), 127.65 (s, 2C, C$_{Ph}$), 66.57 (s, 1C, CHCOOH), 65.68 (s, 1C, CHOH), 26.21 (s, 1C, NCH$_3$), 20.48 (s, 1C, CHCH$_3$).
HRMS (ESI$^+$): calc. for [C$_{13}$H$_{19}$N$_2$O$_4$]$^+$: 267.1339, found: 267.1340 [M+H]$^+$.
mp: 136°C
IR (cm^{-1}): ṽ = 3467 (w), 3349 (w), 1707 (m), 1628 (s), 1588 (m), 1536 (w), 1500 (w), 1448 (w), 1377 (w), 1290 (s), 1127 (w), 1082 (m), 1021 (s) 870 (w), 727 (s), 692 (m).

(Methylcarbamoyl)-*L*-threonine (9b)

Benzyl (methylcarbamoyl)-L-threoninate **15** (0.10 g, 0.38 mmol, 1.0 eq) was dissolved in MeOH (5.0 mL). Pd/C (10.0 mg, 10 wt%) was added and the mixture was stirred in an H$_2$-atmosphere. After 1 h the catalysator was filtered off and the solvent was removed *in vacuo* at 25°C. The product could be afforded as a colorless oil (58.0 mg, 0.33 mmol, yield: 88%).

^{1}H-NMR (400 MHz, D$_2$O) δ = 4.32 (m, 1H, CHOH), 4.23 (m, 1H, CHNHCOOH), 2.69 (s, 3H, NCH$_3$), 1.19 (d, *J* = 6.4 Hz, 3H, CHCH$_3$).
**^{13}C-NMR** (400 MHz, D$_2$O) δ = 173.58 (s, 1C, COOH), 158.66 (s, 1C, N$_2$CO), 65.31 (s, 1C, CHCOOH), 56.67 (s, 1C, CHOH), 24.01 (s, 1C, NCH$_3$), 16.58 (s, 1C, CHCH$_3$).
HRMS (ESI$^-$): calc. for [C$_6$H$_{11}$N$_2$O$_4$]$^-$ 175.0724, found: 175.0723 [M-H]$^-$.
IR (cm^{-1}): ṽ = 3354 (w), 2976 (w), 2936 (w), 1709 (m), 1622 (s), 1500 (m), 1550 (m), 1414 (m), 1277 (m), 1223 (m), 1185 (s), 1122 (m) 1077 (s), 1014 (m) 949 (w), 904 (w), 862 (m), 761 (w).

Synthesis of synthetic nucleoside standards

9-((3aR,4R,6R,6aR)-6-(hydroxymethyl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)-2-(((p-tolylthio)methyl)amino)-1,9-dihydro-6H-purin-6-one (16)
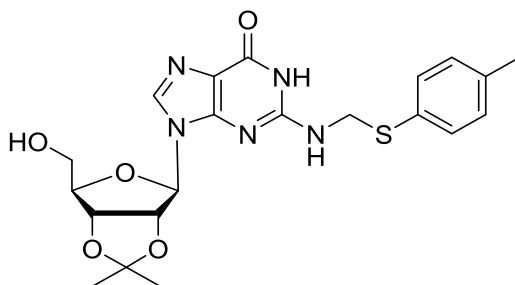
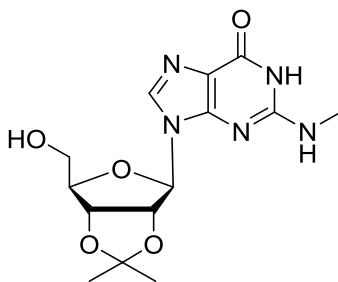The reaction was carried out slightly modified to Bridson *et al.*.[5]

2′,3′-O-isopropylideneguanosine (2.50 g, 7.73 mmol, 1 eq.) and p-thiocresole (1.25 g, 10.1 mmol, 1.3 eq.) were suspended in EtOH (60.0 mL). Acetic acid (1.9 mL) and formaldehyde (2.0 mL) were added to the suspension and the mixture was heated to 90°C. After 14 h the solvent was removed *in vacuo* and the crude product was washed with EtOH to afford the product as a white solid (2.75 g, 5.88 mmol, yield: 76%).

^{1}H NMR (400 MHz, DMSO-d_6) δ = 10.79 (s, 1H, $N_{Ar}H$), 7.95 (s, 1H, HC8), 7.41 – 7.31 (m, 2H, J = 8.0 Hz, HC_{Ar}), 7.17 (d, J = 8.0 Hz, 2H, HC_{Ar}), 6.01 (d, J = 2.6 Hz, 1H, HC1′), 5.36 (dd, J = 6.3, 2.7 Hz, 1H, HC2′), 5.02 (t, J = 5.5 Hz, 1H, OH), 4.95 (dd, J = 6.3, 3.0 Hz, 1H, HC3′), 4.84 (h, J = 6.7, 6.3 Hz, 2H, CH_2C_{Ar}), 4.13 (td, J = 5.5, 3.0 Hz, 1H, HC4′), 3.54 (hept, J = 5.7 Hz, 2H, HC5′), 2.28 (s, 3H, $C_{Ar}CH_3$), 1.54 (s, 3H, CH_3), 1.35 (s, 3H, CH_3).

^{13}C NMR (101 MHz, DMSO-d_6) δ = 154.14 (C6), 151.76 (C2), 149.98 (C4), 137.49 (C8), 136.99 ($C_{Ar}CH_3$), 136.31 ($C_{Ar}S$), 131.43 ($C_{Ar}S$), 130.24 (2C, C_{Ar}), 118.02 (C5), 113.49 ($C(CH_3)_2$), 89.17 (C1′), 87.10 (C4′), 83.75 (C2′), 81.74 (C3′), 61.99 (C5′), 46.71 (CH_2C_{Ar}), 27.41 (CH_3), 25.52(CH_3), 21.04 ($C_{Ar}CH_3$).

HRMS (ESI$^+$): calc. for $[C_{21}H_{26}N_5O_5S]^+$ 460.1649, found: 460.1650 $[M+H]^+$.

9-((3aR,4R,6R,6aR)-6-(hydroxymethyl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)-2-(methylamino)-1,9-dihydro-6H-purin-6-one (17)

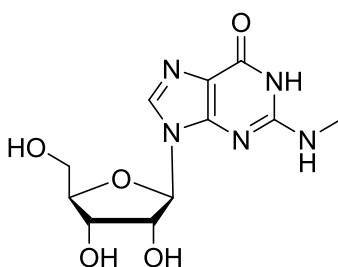The reaction was carried out slightly modified to Bridson *et al.*.[5]

Compound **16** (0.65 g, 1.41 mmol, 1 eq.) was dissolved in DMSO (9.0 mL). NaBH$_4$ (0.11 g, 2.81 mmol, 2 eq.) was added and the mixture was heated to 100°C. After cooling to rt the mixture was purified by column chromatography (DCM : MeOH, 9:1) to afford the product as a white solid (0.41 g, 1.22 mmol, yield: 86%).

^{1}H NMR (400 MHz, DMSO-d_6) δ = 10.80 (s, 1H, $N_{Ar}H$), 7.90 (s, 1H, HC8), 6.46 (m, 1H, $NHCH_3$), 6.00 (d, J = 2.6 Hz, 1H, HC1′), 5.31 (dd, J = 6.2, 2.6 Hz, 1H, HC2′), 5.00 (t, J = 5.5 Hz, 1H, OH), 4.95 (dd, J = 6.2, 3.0 Hz, 1H, HC3′), 4.12 (td, J = 5.5, 3.0 Hz, 1H, HC4'), 3.53 (m, 2H, HC5′), 2.82 (d, J = 4.6 Hz, 3H, NCH_3), 1.53 (s, 3H, CH_3), 1.32 (s, 3H, CH_3).

^{13}C NMR (101 MHz, DMF-d_7) δ = 157.49 (C2), 154.08 (C6), 150.98 (C4), 136.86 (C8), 117.84 (C5), 113.59 ($C(CH_3)_2$), 89.79 (C1′), 87.52 (C4′), 84.22 (C2′), 82.17 (C3′), 62.50 (C5′), 27.88 (NCH_3), 27.03 (CH_3), 25.06 (CH_3).

HRMS (ESI$^+$): calc. for $[C_{14}H_{20}N_5O_5]^+$ 338.1457, found: 338.1459 $[M+H]^+$.

9-((2R,3R,4S,5R)-3,4-dihydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2-(methyl-amino)-1,9-dihydro-6H-purin-6-one (m^{2}G, 18)

The reaction was carried out slightly modified to Ubiali *et al.*.[6]

The protected nucleoside **17** (0.28 g, 0.83 mmol, 1 eq.) was dissolved in H_2O (25.0 mL) and TFA (25.0 mL). After stirring for 1.5 h the solvent was removed *in vacuo*. The residue was suspended in acetone and H_2O. The solvent was removed *in vacuo* and the crude product was recrystallized in EtOH/ H_2O to yield the product as a white solid (0.16 mg, 0.54 mmol, yield: 65%).

1H-NMR (400 MHz, CD$_3$OD) δ = 10.77 (s, 1H, N$_{Ar}$H), 7.95 (s, 1H, HC8), 6.28 (d, ^3J = 5.0 Hz, 1H, NH), 5.73 (d, ^3J = 6.0 Hz, 1H, HC1'), 4.52 (t, J = 5.5 Hz, 1H, HC2'), 4.12 (dd, J = 5.1, 3.4 Hz, 1H, HC3'), 3.88 (d, J = 3.9 Hz, 1H, HC4'), 3.75 – 3.45 (m, 2H, HC5'), 2.81 (d, ^3J = 4.7 Hz, 3H, NCH$_3$).

13C-NMR (101 MHz, CD$_3$OD) δ = 156.77 (C6), 153.29 (C2), 150.97 (C4), 136.15 (C8), 116.65 (C5), 86.79 (C1'), 85.31 (C4'), 73.39 (C2'), 70.55 (C3'), 61.60 (C5'), 27.74 (NCH$_3$).

HRMS (ESI$^+$): calc. for [C$_{11}$H$_{16}$N$_5$O$_4$]$^+$ 298.1146, found: 298.1142 [M+H]$^+$.

Ethyl 9-(2',3',5'-tri-*O*-acetyl-*β*-D-ribofuranosyl)-9*H*-purine-6-carbamate (19)[7]
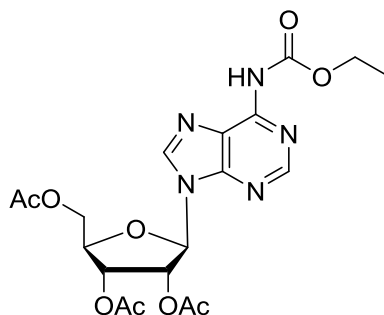

Acetyl protected adenosine (2.00mg, 5.08 mmol, 1 eq) was dissolved in absolute pyridine (37 mL) and cooled to 0°C. Ethyl chloroformate (33.8 µL, 0.35 mmol, 3 eq) was then added dropwise with stirring during 30 min. The mixture was allowed to warm to rt and stirred overnight. The solvent was removed *in vacuo* by co-evaporation with toluene and the crude product was purified by column chromatography (DCM : MeOH, 100:1) to afford the product as a colorless foam (1.05 g, 2.26 mmol, yield: 44%).

1H NMR: (400 MHz, DMSO-d$_6$) δ: 10.56 (s, 1H, NH), 8.66 (s, 1H, HC8), 8.65 (s, 1H, HC2), 6.31 (d, J=5.3, HC1'), 6.06 (t, 1H,J = 5.6 Hz, HC2'), 5.65 (t, 1H,J = 5.6 Hz, HC3'), 4.46-4.33 (m, 2H, HC4', HC5' ^{a}), 4.32-4.22 (m, 1H, HC5'^{b}), 4.18 (q, 1H, J = 7.1 Hz, CH$_2$), 2.12 (s, 3H, OAc), 2.04 (s, 3H, OAc), 2.00 (s, 3H, OAc), 1.26 (7, 3H, J=7.1, CH$_3$).

13C NMR: (101 MHz, DMSO-d$_6$) δ: 170.03 (COCH$_3$), 169.45 (COCH$_3$), 169.28 (COCH$_3$), 152.21 (C(NH)(OCH$_2$CH$_3$)) 151.92 (C6), 151.41(C2), 150.11 (C4), 143.16 (C8), 123.93 (C5), 85.76 (C1'), 79.53 (C4'), 71.90 (C2'), 69.99 (C3'), 62.72 (C5'), 61.02 (CH$_2$), 20.49 (COCH$_3$), 20.38 (COCH$_3$), 20.21 (COCH$_3$), 14.37 (CH$_3$).

HRMS (ESI$^+$): calc. for [C$_{19}$H$_{24}$N$_5$O$_9$]$^+$ 466.1569, found: 466.1569 [M+H]$^+$.

2-(3-(9-((2R,3R,4R,5R)-3,4-diacetoxy-5-(acetoxymethyl)tetrahydrofuran-2-yl)-9H-purin-6-yl)ureido)-3-hydroxybutanoic acid (20)
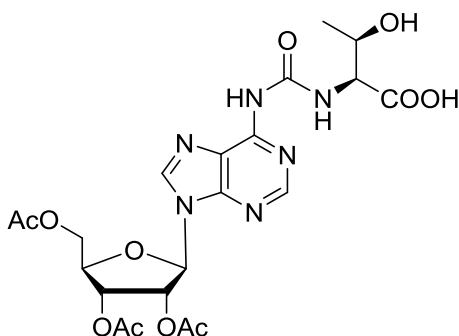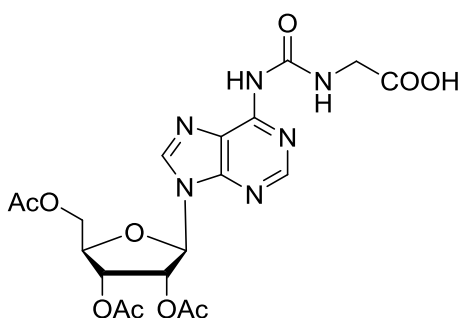
Carbamate **19** (0.30 g, 0.64 mmol, 1 eq) was dissolved in absolute pyridine (30 mL) and L-threonine (0.15 g, 1.29 mmol, 2 eq) was added. The mixture was stirred under reflux for 6 h and cooled down to room temperature. Excess of L-threonine was removed by filtration and the. The solvent was removed *in vacuo* by co-evaporation with toluene and the crude product was recrystallized in EtOH to afford the product as a white solid (0.29 g, 0.54 mmol, yield: 84%).

^{1}H NMR: (400 MHz, DMSO-d$_6$) δ: 9.92 (s, 1H, COOH), 9.68 (d, J = 8.4 Hz, 1H, NH), 8.66 (s, 1H, HC8), 8.58 (s, 1H, HC2), 6.31 (d, J = 5.6 Hz, 1H, HC1'), 6.04 (t, J = 5.7 Hz, 1H HC2'), 5.62 (dd, J = 5.9, 4.4 Hz, 1H, HC3'), 4.48 − 4.36 (m, 2H, CHOH, CHNH), 4.36 − 4.13 (m, 3H, HC5',HC4'), 2.14 (s, 3H, OAc), 2.04 (s, 3H, OAc), 2.03 (s, 3H, OAc), 1.15 (d, J = 6.2 Hz, 3H, CH$_3$).

^{13}C-NMR (101 MHz, DMSO-d$_6$) δ = 172.85 (COOH), 170.53 (COCH$_3$), 169.94 (COCH$_3$), 169.74 (COCH$_3$), 154.12 (CON$_2$), 151.61 (C2), 150.94 (C6), 150.56 (C4), 143.09 (C8), 121.00 (C5), 86.07 (C1'), 80.11 (COH), 72.42 (C4'), 70.50 (C2'), 66.65 (C3'), 63.26 (CHCOOH), 59.09 (C5'), 21.26 (COCH$_3$), 20.99 (COCH$_3$), 20.86 (COCH$_3$), 20.67 (CH$_3$) ppm.

HRMS (ESI$^+$): calc. for [C$_{21}$H$_{27}$N$_6$O$_{11}$]$^+$ 539.1732, found: 539.1736 [M+H]$^+$.

((9-((2R,3R,4R,5R)-3,4-diacetoxy-5-(acetoxymethyl)tetrahydrofuran-2-yl)-9H-purin-6-yl)carbamoyl)glycine (21)


Carbamate **19** (0.75 g, 1.61 mmol, 1 eq) was dissolved in absolute pyridine (60 mL) and L-glycine (0.24 g, 3.22 mmol, 2 eq) was added. The mixture was stirred under reflux for 7 h and cooled down to room temperature. Excess of glycine was removed by filtration and the. The solvent was removed *in vacuo* by co-evaporation with toluene and the crude product was recrystallized in EtOH to afford the product as a white solid (0.24 g, 0.50 mmol, yield: 31%).

^{1}H NMR: (400 MHz, **DMSO-d$_6$**) δ: 12.69 (s, 1H, COOH), 9.99 (s, 1H, NH), 9.61 (bs, 1H, NH), 8.65 (s, 1H, HC8), 8.59 (s, 1H, HC2), 6.30 (d, J = 5.3 Hz, 1H, HC1'), 6.03 (t, J = 5.6 Hz, 1H HC2'), 5.63 (t, J = 5.3, 1H, HC3'), 4.45 − 4.38 (m, 2H, HC5'), 4.31 − 4.20 (m, 1H, HC4'), 3.99 (d, J = 5.3 Hz, 1H, CH$_2$), 2.12 (s, 3H, COCH$_3$), 2.04 (s, 3H, COCH), 2.01 (s, 3H, COCH).

HRMS (ESI+): calc. for $[C_{19}H_{23}N_6O_{10}]^+$ 495.1470, found: 495.1466 $[M+H]^+$.
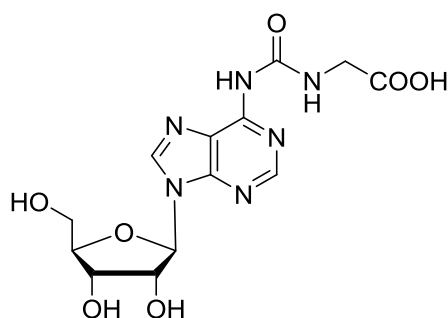

t⁶A (22)
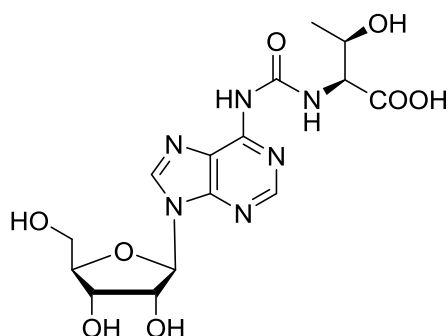
Protected nucleoside **20** (0.17 g, 0.32 mmol, 1 eq) was dissolved in 7 N Ammonia in MeOH (8 mL). The mixture was heated to 40°C for 1 h and then stirred overnight at rt. The solvent was removed *in vacuo* and the crude product was recrystallized in EtOH to afford the product as a white solid (0.97 g, 0.24 mmol, yield: 75%).

¹H NMR: (400 MHz, **DMSO-d₆**) δ: 9.61 (d, *J* = 6.4 Hz, 1H, NH), 8.64 (s, 1H, HC8), 8.52 (s, 1H, HC2), 5.98 (d, *J* = 5.6 Hz, 1H, HC1'), 4.59 (t, *J* = 5.3 Hz, 1H HC2'), 4.18 (t, *J* = 4.1 Hz, 1H, HC3'), 4.06 – 4.01 (m, 2H, CHOH, CHNH), 3.96 (q, *J* = 3.8 Hz, 1H, HC4'), 3.63 (ddd, *J* = 48.7, 12.0, 4.0 Hz, HC5'), 1.00 (d, *J* = 6.1 Hz, 3H, CH₃).
¹³C-NMR (101 MHz, DMSO-d₆) δ = 172.61 (COOH), 153.21 (CON₂), 151.42 (C2), 150.83 (C6), 150.62 (C4), 142.49 (C8), 120.80 (C5), 88.13 (C1'), 86.08 (COH), 74.22 (C4'), 70.69 (C2'), 66.48 (C3'), 61.70 (C5'), 58.86 (CHCOOH), 19.89 (CH₃) ppm.
HRMS (ESI⁺): calc. for $[C_{15}H_{21}N_6O_8]^+$ 413.1415, found: 413.1415 $[M+H]^+$.

g⁶A (21)

Protected nucleoside **21** (202 mg, 0.41 mmol, 1 eq) was dissolved in 7 N Ammonia in MeOH (10 mL). The mixture was heated to 40°C for 1 h and then stirred overnight at rt. The solvent was removed *in vacuo* and the crude product was recrystallized in EtOH to afford the product as a white solid (74.0 mg, 0.20 mmol, yield: 49%).

¹H NMR: (400 MHz, **DMSO-d₆**) δ: 9.55 (d, *J* = 6.4 Hz, 1H, NH), 8.66 (s, 1H, HC8), 8.54 (s, 1H, HC2), 5.99 (d, *J* = 5.6 Hz, 1H, HC1'), 4.18 (dd, *J* = 4.9, 3.7 Hz, 1H HC2'), 3.97 (t, *J* = 4.1 Hz, 1H, HC3'), 3.78 (d, *J* = 4.9 Hz, 1H, HC4'), 3.64 (ddd, *J* = 48.7, 12.0, 4.0 Hz, HC5').

¹³C-NMR (101 MHz, DMSO-d₆) δ = 171.67 (COOH), 153.51 (C2), 151.34 (C6), 150.74 (C4), 142.56 (C8), 120.73 (C5), 88.14 (C1'), 86.14 (C4'), 74.24 (C2'), 70.73 (C3'), 61.73 (C5'), 43.85 (CH₂) ppm.
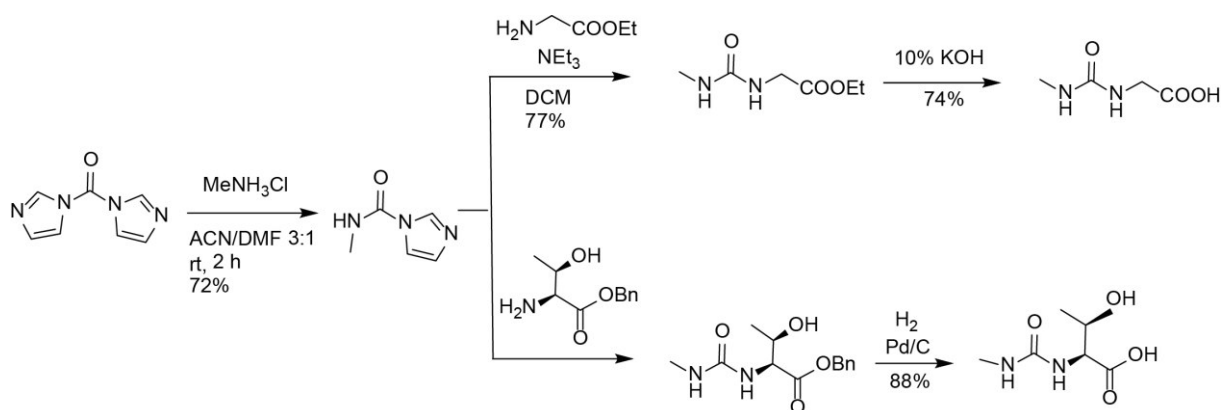**HRMS** (ESI⁺): calc. for $[C_{13}H_{17}N_6O_7]^+$ 369.1153, found: 369.1152 $[M+H]^+$.

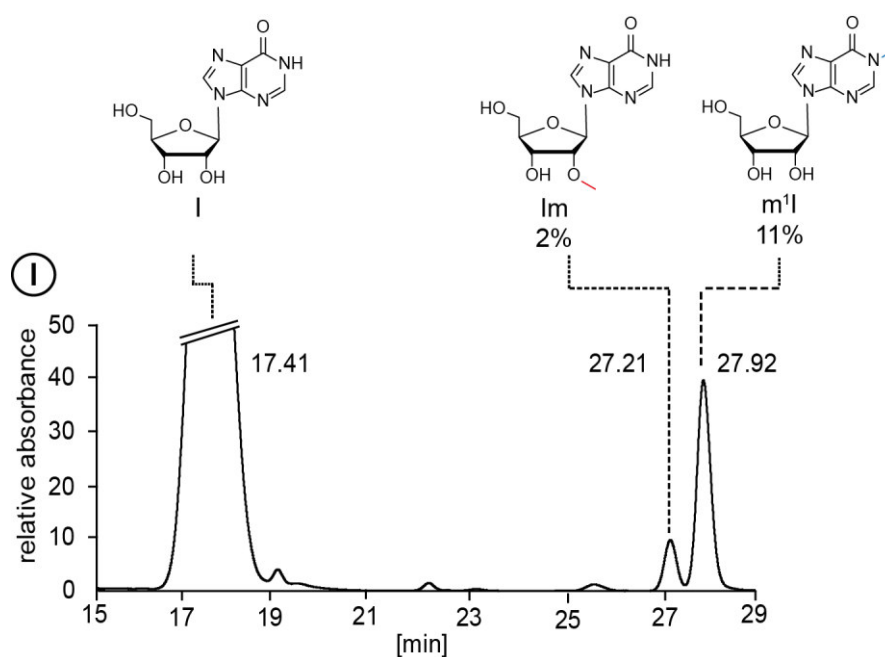**Figure S 1:** Reaction scheme of the organic synthesis of the amino acid urea compounds.

Figure S 2: UV-chromatogram (at 260 nm) of the reaction of Inosine with *N*-methly-*N*-nitroso-urea under prebiotic conditions.
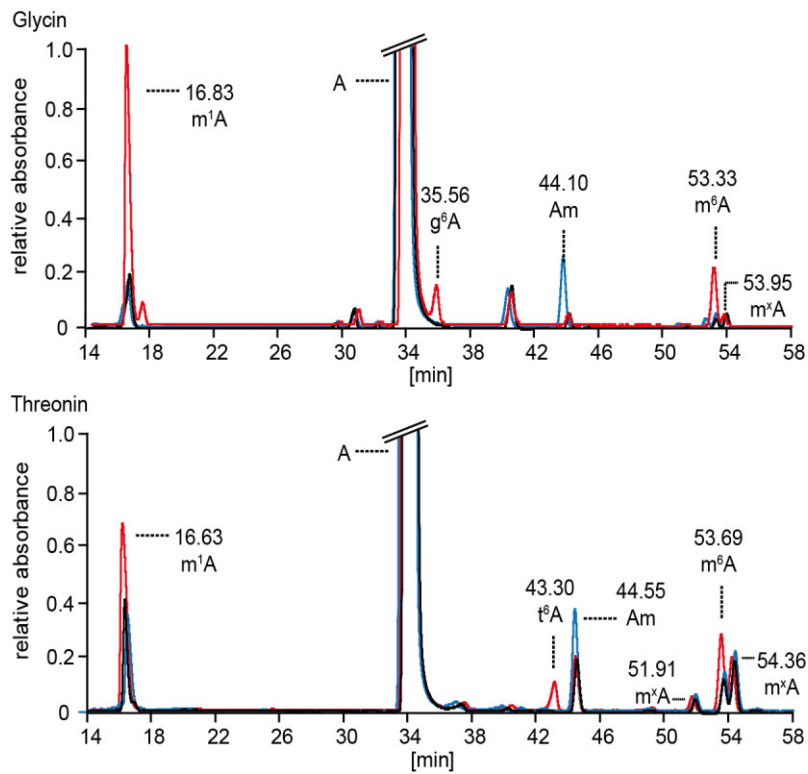
Figure S 3: UV-chromatogram (at 260 nm) of the reaction of Adenosine with amino acid- nitroso compounds under prebiotic conditions. Chromatogram of the reaction itself is shown in black, coinjections of the different modifications are shown in blue and red.
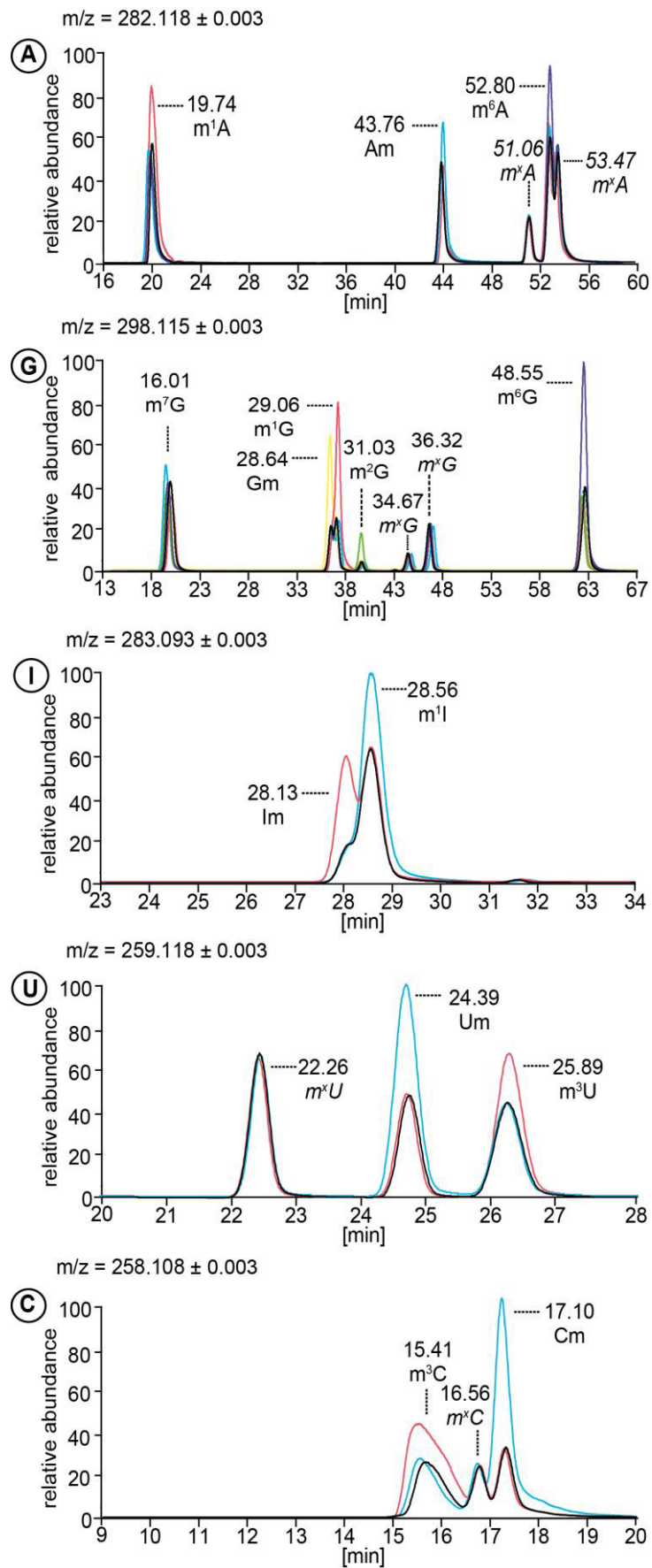
Figure S 4: MS-chromatogram of the reaction of the different canonical nucleosides and Inosine with *N*-methly-*N*-nitroso-urea (black) and the coinjections with synthetic standards (red, blue, yellow).
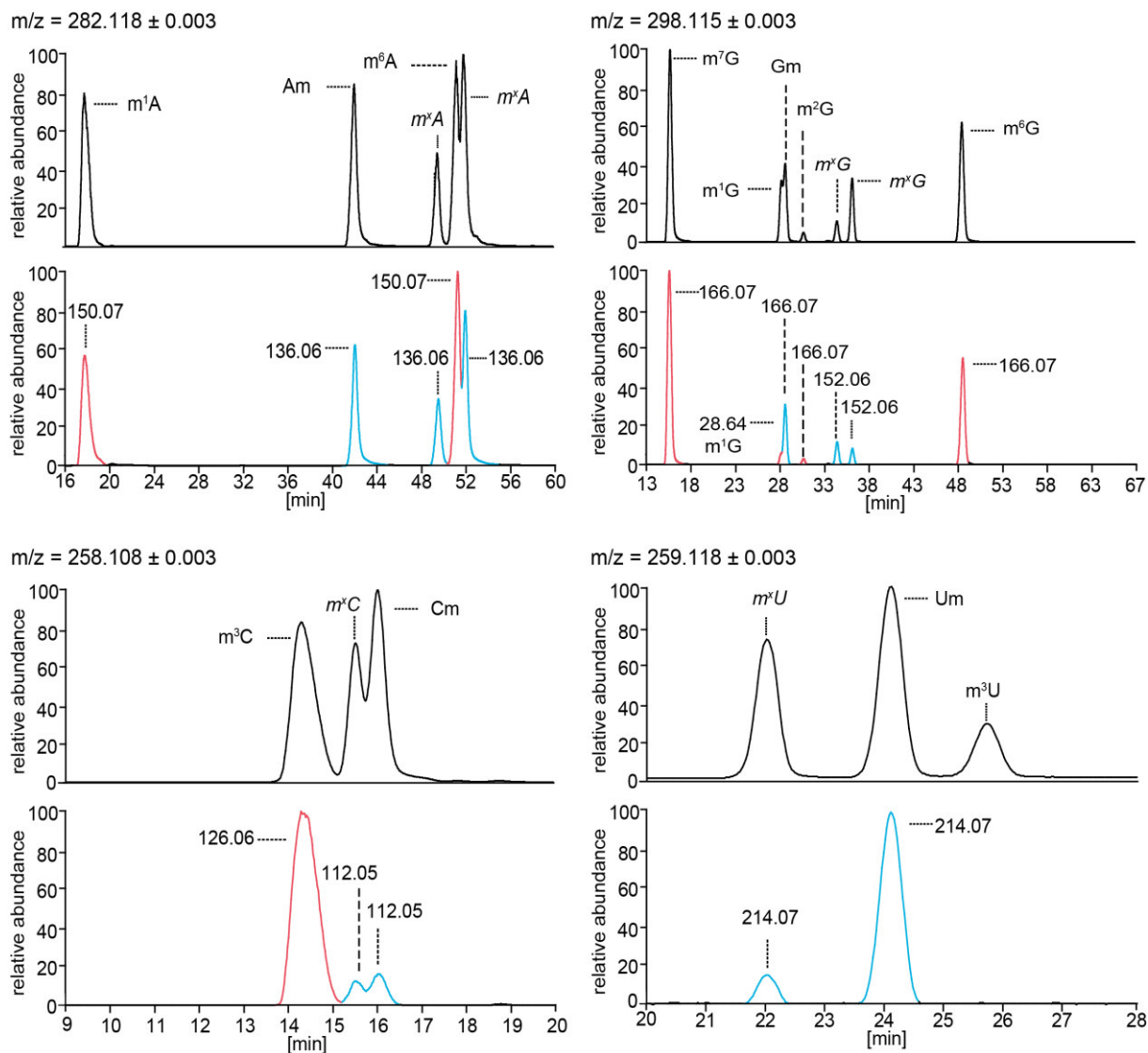
Figure S 5: Fragmentation studies of the methylation products of the prebiotic reaction of *N*-methly-*N*-nitroso-urea with the canonical nucleosides. Sugar modified nucleosides are shown in blue, and the nucleosides with a methylation at the base in red. The numbers represent the base peak of the modification.
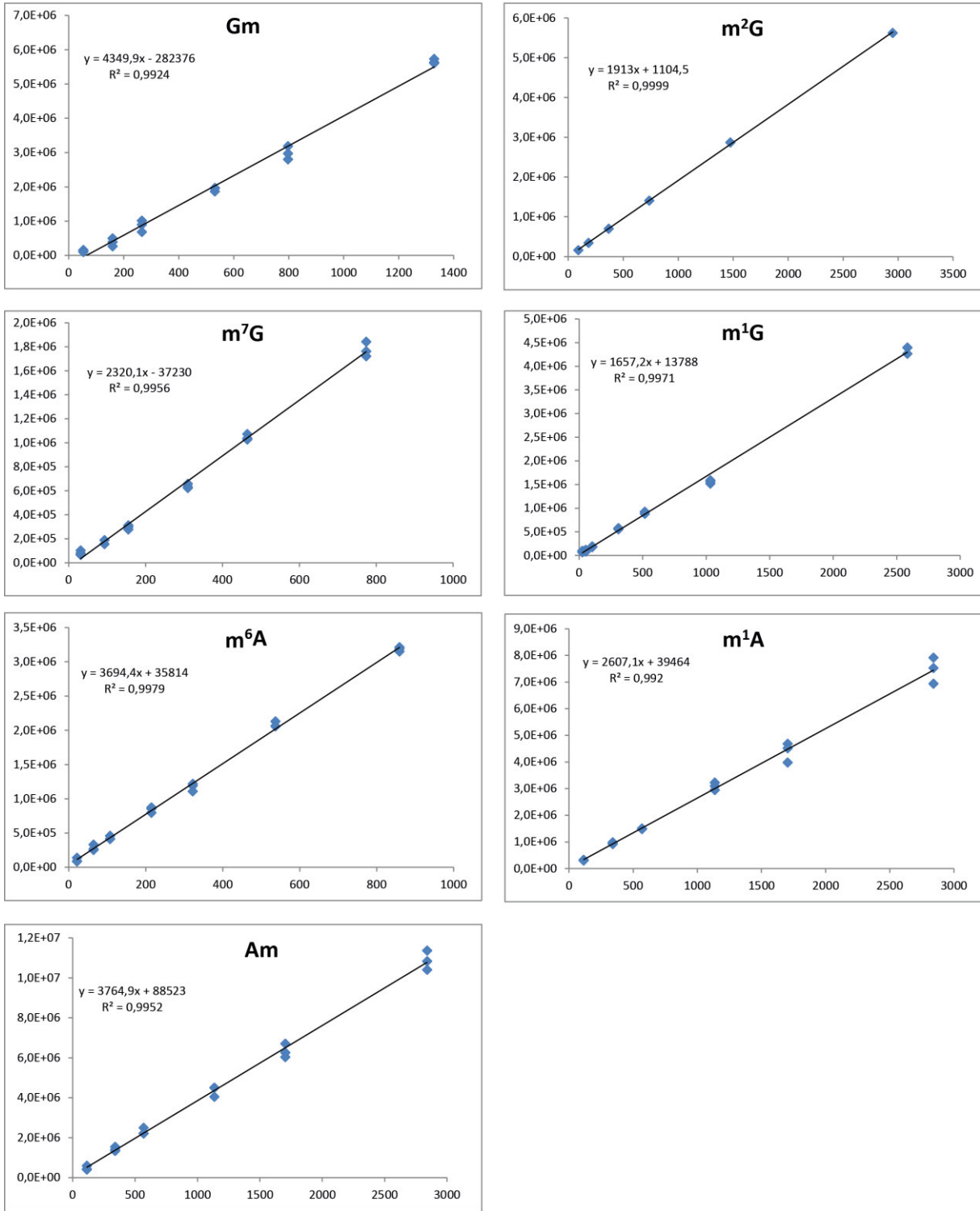
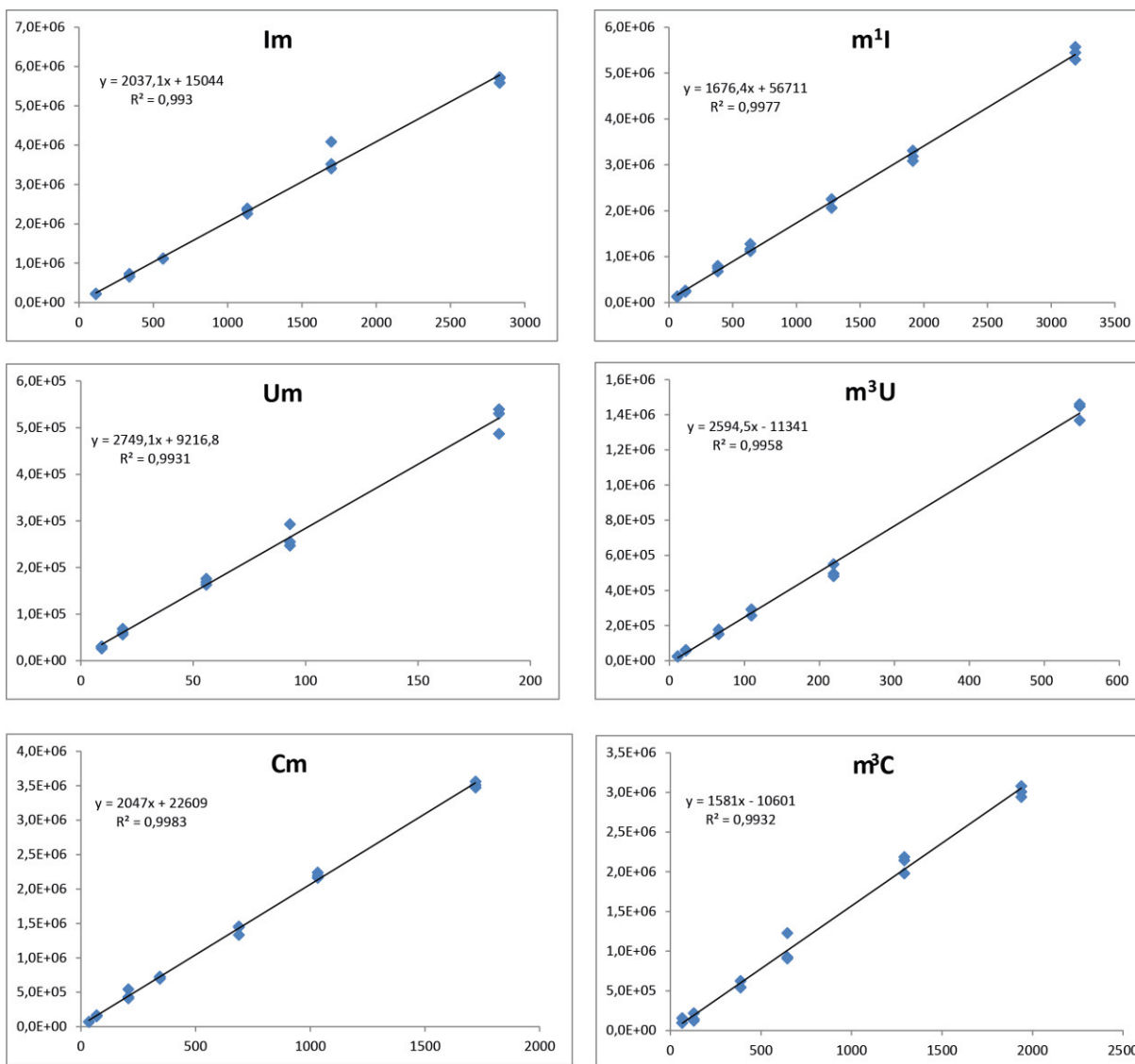**Figure S 6**: Calibration curves at 260 nm for quantification of modified A and G nucleosides from UV-signals.

Figure S 7: Calibration curves at 260 nm for quantification of modified C, U and I nucleosides from UV-signals.

Literature:

[1]   F. Arndt, *Org. Synth.* **1943**, 3-3.

[2]   T. Machinami, T. Suami, *Bull. Chem. Soc. Jpn.* **1975**, *48*, 1333-1334.

[3]   F. U. Richter, *Chem. Eur. J.* **2009**, *15*, 5200-5202.

[4]   P. A. Duspara, M. S. Islam, A. J. Lough, R. A. Batey, *J. Org. Chem.* **2012**, *77*, 10362-10368.

[5]   P. K. Bridson, C. B. Reese, *Bioorg. Chem.* **1979**, *8*, 339-349.

[6]   D. Ubiali, C. D. Serra, I. Serra, C. F. Morelli, M. Terreni, A. M. Albertini, P. Manitto, G. Speranza, *Adv. Synth. Catal.* **2012**, *354*, 96-104.

[7]   M. Matuszewski, E. Sochacka, *Bioorganic Med. Chem. Lett.* **2014**, *24*, 2703-2706.