Application of TMAH thermochemolysis to the detection of nucleobases: Application to the MOMA and SAM space experiment

Y. He, Arnaud Buch, M. Morisson, Cyril Szopa, Caroline Freissinet, A. Williams, Maeva Millan, Melissa Guzman, R. Navarro-González, Jean-Yves Bonnet, et al.

Application of TMAH thermochemolysis to the detection of nucleobases: application to the MOMA and SAM space experiment

Y. He[1], A. Buch[1*], M. Morisson[1], C. Szopa[2, 3], C. Freissinet[2], A. Williams[4], M. Millan[5,8], M. Guzman[2], R. Navarro Gonzalez[6], J. Y. Bonnet[2, 7], D. Coscia[2], J.L. Eigenbrode[8], C. A. Malespin[8], P. Mahaffy[8], D.P. Glavin[8], J.P. Dworkin[8], P. Lu[9], S. S. Johnson[5]

[1] LGPM, CentraleSupélec, University Paris-Saclay, 8-10 rue Joliot-Curie, 91190, Gif-sur-Yvette, France
[2] LATMOS/IPSL, UVSQ Université Paris-Saclay, Sorbonne Université, CNRS, 11 Bd d'Alembert, 78280 Guyancourt, France.
[3] Institut Universitaire de France, Paris, France.
[4]University of Florida Gainesville, FL, USA
[5] Georgetown University, Washington D.C., USA
[6] Instituto de Ciencias Nucleares, Universidad Nacional Autónoma de México, Circuito Exterior, Ciudad Universitaria, Apartado Postal 70-543, Ciudad de México 04510, Mexico
[7]Telespazio France, Toulouse, France
[8]NASA Goddard Space Flight Center, Greenbelt, MD, USA
[9]LGPM, CentraleSupélec, SFR Condorcet FR CNRS 3417, Université Paris-Saclay, Centre Européen de Biotechnologie et de Bioéconomie (CEBB), 3 rue des Rouges Terres 51110 Pomacle, France

Abstract: Thermochemolysis of seven nucleobases—adenine, thymine, uracil, cytosine, guanine, xanthine, and hypoxanthine—in tetramethylammonium hydroxide (TMAH) was studied individually by pyrolysis gas chromatography mass spectrometry in the frame of the Mars surface exploration. The analyses were performed under conditions relevant to the Sample Analysis at Mars (SAM) instrument of the Mars Curiosity Rover and the Mars Organic Molecule Analyzer (MOMA) instrument of the ExoMars Rover. The thermochemolysis products of each nucleobase were identified and the reaction mechanisms studied. The thermochemolysis temperature was optimized and the limit of detection and quantification of each nucleobase were also investigated. Results indicate that 600°C is the optimal thermochemolysis temperature for all seven nucleobases. The methylated products trimethyl-adenine, 1, 3-dimethyl-thymine, 1, 3-dimethyl-uracil, trimethyl-cytosine, 1, 3, 7-trimethyl-xanthine (caffeine), and dimethyl-hypoxanthine, respectively, are the most stable forms of adenine, thymine, uracil, cytosine, guanine, and xanthine, and hypoxanthine in TMAH solutions. The limits of detection for adenine, thymine, and uracil were 0.075 nmol; the limits of detection for guanine, cytosine, and hypoxanthine were higher, at 0.40, 0.55, and 0.75 nmol, respectively. These experiments

* Author to whom correspondence should be addressed: E-mail address: arnaud.buch@centralesupelec.fr.

allowed to well constrain the analytical capabilities of the thermochemolysis experiments that will be performed on Mars to detect nucleobases.

Keywords: Nucleobases, Thermochemolysis, Derivatization, Pyrolysis, GC-MS, Tetramethylammonium hydroxide (TMAH), SAM, MOMA

1. Introduction

Searching for the biochemical precursors at the Mars surface is one of the main goals of the Sample Analysis at Mars (SAM) and Mars Organic Molecule Analyzer (MOMA) experiments onboard the Curiosity and ExoMars 2020 rovers, respectively. Among the chemical compounds, organic molecules are of particular importance and key species include carboxylic acids, amino acids, and nucleobases. Seven nucleobases are of particular importance in terrestrial biology. Five — adenine (A), guanine (G), thymine (T), cytosine (C), and uracil (U) — are contained within DNA and/or RNA. These nucleobases build codons, and each codon corresponds to an amino acid. Two others — xanthine (X) and hypoxanthine (HX) — are important metabolic intermediates. Here we show that if these seven nucleobases were present in Martian samples, they could be detected *in situ* by instruments on current Mars rover missions even at low concentrations.

Low concentrations of nucleobases have been identified in carbonaceous chondrite-type meteorites[1–10] . Callahan [11] analyzed the component of five different martian meteorites after formic acid extraction, purification and concentration followed by liquid chromatography-mass spectrometry. However, no nucleobases or nucleobase analogs were observed above their detection limit of 1 ppb in any of the martian meteorite samples studied [13]. These laboratory studies of carbonaceous chondrites and martian meteorites provide a solid baseline for future searches for nucleobases on the martian surface.

Though Basiuk and Douda [12] investigated the recovery rate of underivatized adenine, guanine, cytosine, and uracil after sublimation at temperatures between 400°C and 1000°C using a laboratory furnace under a N_2 or CO_2 atmosphere. Results indicated that the temperature range from 500-600°C represented the best compromise between the thermal decomposition and volatilization rate, suggesting that thermal introduction of nucleobases is a viable strategy for sample analysis. Pyrolysis-Gas Chromatography-Mass Spectrometry (Pyr-GC-MS) is one of the operational modes used to detect the organic compounds as part of the SAM and MOMA experiments. Unlike sublimation, pyrolysis does not preserve the initial organic compounds, some derivatization solvent should be used, such as N,N-tert-butyl-dimethylsilyl-trifluoroacetamide (MTBSTFA) and

dimethylformamide (DMF) and tetramethylammonium hydroxide (TMAH) will be used on MOMA and SAM to protect polar compounds released from the pyrolysis experiment. MTBSTFA and DMF have been proved as good derivatization solvent. For example, Glavin [13] detected the adenine, thymine, uracil, cytosine, and xanthine (interpreted as a product of guanine degradation) in the extract of *Escherichia coli* strains *(E. coli)* inoculated with serpentine samples. Results showed that sublimation coupled with chemical derivatization using MTBSTFA and DMF as the derivatization solvent, with GC-MS analysis, could detect nucleobases in Martian soil analogs. TMAH works as a methylation agent boosting the cleavage of macromolecules, thereby methylating the products released from sample pyrolysis, and giving rise to improved detection by GC-MS. For example, thymine can react with TMAH [14,15]. TMAH thermochemolysis can provide much more important information about organic compounds of different samples [16–20], such as amino acids [21–26], aliphatic and aromatic carboxylic acids [27], natural water, soil samples, and kerogen [28,29]. TMAH derivatization can increase the volatility of the nucleobases by decreasing the polarity of the labile molecules, which makes it easier to identify nucleic acids in the complex mixtures [14].

However, the classical process of offline thermochemolysis requires several steps, including the process of solvent derivatization, the thermochemolysis products extraction, and the supernatant drying under a flow of N_2 to remove solvent or separate derivatives prior to GC-MS analysis [30,31]. At present, these steps are too technically complex to be fully automated for *in situ* analysis in space, especially for the detection of organic compounds on Mars. Therefore, a simpler "on-line" thermochemolysis method suited to such an application has been implemented for SAM and MOMA. 500 µL of 25% TMAH in methanol and 15 µL of TMAH are stored in the SAM oven [32] and MOMA derivatization capsules [33], respectively. The TMAH thermochemolysis process needs to be tested and optimized according to the analysis conditions of the SAM and MOMA experiments, i.e. by contacting the sample and the TMAH directly in contact inside the pyrolysis furnace without any further sample processing. The results of TMAH derivatization are affected by the experimental conditions, such as TMAH concentrations, pyrolysis temperature, and volume of TMAH (which in turns affects pH) [24,34] . TMAH thermochemolysis, however, has rarely been considered for the analysis of nucleobases based on the SAM and MOMA missions, thus the experimental conditions as well as the direct TMAH thermochemolysis process in SAM and MOMA conditions have yet to be optimized.

In this study, the TMAH thermochemolysis products of seven nucleobase standards and mixtures pyrolyzed at different temperatures were analyzed individually by Pyr-GC-MS. The derivatizations of all nucleobases at the optimal pyrolysis temperature, as well as the limit of detection (LOD) and limit of quantification (LOQ) of the nucleobases were determined. In addition, the thermochemolysis mechanisms of all nucleobases in TMAH were analyzed. This data is the first of its kind, benchmarking whether it would be possible to detect all or some of the studied nucleobases with the *in situ* TMAH thermochemolysis experiments on the SAM and MOMA experiments of the Mars Science Laboratory and ExoMars 2020 space missions. However, it should be noted that on Mars the nucleobases are contained in a solid matrix from which they will have to be extracted.

2. Experiments

2.1 Materials and methods

In this study, individual solutions of adenine (Sigma, purity > 99%), thymine (Fluka, purity > 99%), uracil (Alfa Aesar, purity > 99%), cytosine (Fluka, purity ~ 97%), guanine (Fluka, purity > 99%), xanthine (Alfa Aesar, purity > 99%) and hypoxanthine (Sigma, purity > 99%) were prepared by diluting the pure solid nucleobases in TMAH solutions (25% by weight in methanol, Sigma-Aldrich). The concentration of each nucleobase and their mixtures used in this study are shown in Table 1. TMAH was always in excess compared to the nucleobase, with a minimum of six molecules of TMAH for three methylation sites in the case of cytosine. Naphthalene-d8 (Sigma-Aldrich, isotopic purity: 99 atom % D) was used as an internal standard. All of the samples were stored in the refrigerator at 3°C.

2.2 Pyrolysis-GC-MS and methods

During the pyrolysis experiments, several amount of each nucleobase solution was pyrolyzed at 400°C, 500°C, and 600 °C, and 0.1 µl of the mixture of seven nucleobases in solution were pyrolyzed at 600°C. The pyrolysis experiments were performed with an EGA/PY-3030D micro-oven pyrolyzer (Frontier Lab), installed on the Split/SplitLess (SSL) injector of a gas chromatograph (Trace GC Ultra, Thermo Scientific) coupled to a quadrupole mass spectrometer (ISQ LT, Thermo Scientific). An Optic 4 injector (GL Sciences) was used for the low temperature analyses (injection at 250°C and 300°C). It was connected to a GC (Trace GC Ultra, Thermo Scientific) coupled to a quadrupole mass spectrometer (DSQ II, Thermo Scientific). The GC was equipped with a Zebron ZB-5HT

Inferno column (30 m × 0.25 mm i.d. × 0.25 μm film thickness) with a 5 m integrated guard column. The temperature programming of the column started at 50°C, was held for 5 min, then heated at a rate of 6°C.min^{-1} up to 240°C, then raised to 300°C at a rate of 10°C.min^{-1} and held for2 min. Helium was used as the carrier gas, with the helium flow rate in the column at 1.5 mL min^{-1}. The split flow was 50 mL min^{-1}. The temperature of the SSL injector was 280°C. The masses were scanned between m/z 40 and m/z 500. The ionization energy was 70 eV. These conditions are similar to those on SAM and MOMA since it can be used to analyze whether or not it will be able to detect nucleobases on Mars.

2.3 Quantification of all nucleobases

A solution of naphthalene-d8 in dichloromethane was prepared at a concentration of 0.05 mol·L^{-1}, and 0.4 μl of naphthalene-d8 internal standard solution was added in each solution of the nucleobase. A standard solution of all seven nucleobases (A, G, T, U, C, X, and HX) at a concentration of 0.01 mol·L^{-1} was prepared with the mix of TMAH (25% by weight in methanol) and H_2O in the same volume. For the calibration curve of guanine, different amounts of guanine (0, 15, 20, 40, 80, 120 nmol) were injected; for the other six nucleobases, 0, 5, 10, 15, 20, 30, 50 nmol were injected. The main peaks of each nucleobase (*m/z* = 177, 154, 140, 153, 194, 194, 164) were chosen respectively as target products for A, T, U, C, G, X, and HX quantification.

In order to analyze the mixtures of all nucleobases, a new GC temperature program was used. The initial column temperature was set at 40°C and held for 2 min, followed by a ramp of 3°C min^{-1} up to 200°C. The temperature raised to 300°C at a rate of 6°C min^{-1} and maintained at this temperature for 2 min. The split ratio was set at 20:1. All of the nucleobase solutions were sampled in a clean capsule (pre-cleaned at 1000°C for 30 sec) and dried completely with a stream of nitrogen. Then 3.0 μl of TMAH was added. All points were repeated more than three times. The quantification curves of all nucleobases were used to quantify the corresponding nucleobase.

3. Results and Discussion

3.1 Optimization of the thermochemolysis for nucleobases

To determine the optimal temperature of TMAH thermochemolysis by direct Pyr-GC-MS analysis, each of the seven nucleobases diluted in TMAH was pyrolyzed at 400°C, 500°C, and 600°C.

3.1.1 Adenine

2.5 nmol of adenine dissolved in TMAH was injected successively at 400°C, 500°C, and 600°C, and all derivatives of adenine were analyzed. The chromatograms obtained are given in Figure 1A. At 500°C, 3-methyl-adenine was not detected. The five peaks 5, 6, 7, 8 and 9 were methylated derivatives of adenine which cannot be detected at 400°C, however, their exact tautomeric forms cannot be determined. The major peak is N, N, 9-trimethyl-adenine (or N, N, 9-trimethyl-9H-purin-6-amine, peak 2), and its three labile hydrogen sites are fully methylated at every tested temperature. At this amount of adenine (2.5 nmol), these two compounds were detected from 400°C to 600°C. However, it should be noted that the intensity of all peaks increased strongly with the increase of pyrolysis temperature, and the optimal detection temperature of adenine in TMAH thermochemolysis was 600°C. According to our previous work, when the temperature is higher than 600°C, TMAH can be degraded. This gives rise to more byproducts [35], which are degradational to the thermochemolysis of nucleobases in TMAH; hence 600 °C is considered the optimal temperature for TMAH thermochemolysis experiments. Some siloxane peaks were detected in the thermochemolysis products of seven nucleobases at different temperatures, there are also siloxane in the blank experiments. Thus, these siloxanes may originate either from the thermal degradation of the septum of the GC injector or from the degradation of the stationary phase of the column caused by the corrosive effect of TMAH, or both.

Figure 1B shows the chromatograms of different amounts of adenine thermochemolyzed at 600°C. For the 0.125 nmol adenine solution (Figure 1B(a)), three different methylated derivatives of adenine were identified, with low intensity peaks: a dimethylated form (peak 2: N, 9-dimethyl-adenine) and two kinds of trimethylated tautomeric forms (peak 3: N, N, 9-trimethyl-adenine, and peak 4: N, N, 3-trimethyl-adenine). No additional peaks were observed for an injection of 0.25 nmol (Figure 1B (b)). There were six adenine related peaks in addition to the three mentioned previously when 2.5 nmol of adenine was analyzed (Figure 1B (c)). Peak 1 is 7-methyl-adenine, which corresponds to the methylation of adenine. The other five products could not be unambiguously attributed to specific derivatives of adenine. However, if the peaks with the highest *m/z* values represent the molecular ion (and not the fragment of a heavier molecule), the chromatographic data of five peaks correspond to a dimethyl derivative of adenine (peak 5: *m/z* 163) and four trimethylated tautomeric forms (peaks 6, 7, 8 and 9: *m/z* 177). The mass spectra corresponding to peaks 5 to 9 are given in Figure S1(see the supplementary material) . All of the corresponding retention times, and the masses of the

molecular ions and main fragments are given in Table 2. It should be noted that there are some other peaks on the chromatograms that are not related to adenine.

3.1.2 Thymine and Uracil

Once thermochemolysis occurs, there are similar derivatized pyrimidine bases between thymine and uracil since they differ only by the presence of an additional methyl group at carbon position 5 in the thymine (i.e. 5-methyl-uracil, see Figure 2). Thymine and uracil solutions at a concentration of 2.6×10^{-3} mol·L^{-1} in TMAH were analyzed by Pyr-GC-MS.

Since the compounds respond well to that concentration (high signal intensity), there was no need to inject a larger amount of thymine and uracil. The chromatograms of thymine and uracil obtained at 400°C, 500°C, and 600°C are shown in Figure 2, respectively. 1, 3-dimethyl-thymine (1, 3, 5-trimethyl-2, 4 (1H, 3H) -pyrimidinedione) was the only compound derived from thymine detected at a retention time of 21.5 min. The only derivative of uracil was 1, 3-dimethyl-uracil (1, 3-dimethyl-2, 4 (1H, 3H) -pyrimidinedione) at a retention time of 20.4 min. The molecular ion masses of 1, 3-dimethyl-thymine and 1, 3-dimethyl-uracil and their major fragments are shown in Table 2, respectively.

For thymine and uracil, the thermochemolysis temperature had no effect on the number of derivatives detected, since only the canonical form was observed. However, the signal intensities of thymine and uracil were stronger at a higher pyrolysis temperature (i.e. 600°C), suggesting a high reaction rate. The optimal temperature for analysis of thymine and uracil pyrolysis products is 600°C, same as the case of adenine. As for adenine, siloxane peaks, coming from the column bleeding or septum decomposition, were detected.

3.1.3 Cytosine

Cytosine solution at a concentration of 0.3 mol·L^{-1} in TMAH was analyzed by Pyr-GC-MS. Because injection of 0.3 and 2.5 nmol of cytosine in TMAH gave no chromatographic response, 15 nmol of cytosine in TMAH was used to prevent a possible decrease in the response of the cytosine derivatization at a lower temperature. Compared with the nucleobases studied above, the cytosine solution was much more concentrated. The difficulty of detecting cytosine is probably due to the high polar primary amino group. The analysis of amines is particularly problematic with Pyr-GC-MS not only because of their adsorption on the transfer line and the liner of the pyrolyzer, but also because of their

interaction with the stationary phase of the chromatographic column (5% diphenyl/95% dimethyl polysiloxane), which results in significant spreading peaks.

The chromatograms of 15 nmol cytosine pyrolyzed at 400°C, 500°C, and 600°C are shown in Figure 3. Whatever the pyrolysis temperature, the same molecules were detected, including N, N, N-trimethyl-cytosine corresponding to the canonical form of the three methylated cytosine, and 2-O-methyl-cytosine corresponding to an enol form of the cytosine methylated on the oxygen atom site. Besides, peak 2 (m/z = 139), peaks 3 and 5 (m/z = 153) correspond to a dimethyl and trimethyl forms of cytosine, respectively. All of the cytosine derivatives in TMAH are summarized in Table 2. The mass spectra corresponding to peak 2, 3, and 5 are given in Figure S2 (see the supplementary material), respectively. The spectrum of peak 5 spreads over several minutes, and may overlap with other coeluted tautomers; however, according to the mass spectra, the main compound is trimethyl-cytosine. In summary, the major peaks correspond to two tautomeric forms of trimethyl-cytosine: N, N, N'-trimethyl-cytosine, which is the canonical form of the fully methylated cytosine, and another trimethylated tautomeric form.

As in the case of thymine, the thermochemolysis temperature of cytosine had no effect on the number of functionalized derivatives detected; six methylated derivatives of cytosine were observed at 400°C, 500°C, and 600°C, respectively. However, from a quantitative point of view, the intensity of cytosine peaks is higher at 600°C than that at other temperatures. Hence, 600°C is the optimal temperature.

3.1.4 Guanine

11 nmol of guanine in TMAH solution was injected to study the effect of thermochemolysis temperature. The chromatograms of guanine thermochemolyzed at 400°C, 500°C, and 600°C are shown in Figure 4. 1, 3, 7-trimethylxanthine (caffeine) was identified using the NIST database when guanine is thermochemolyzed at 500°C and 600°C. The other peaks were identified from the *m/z* ratios of their supposed molecular ions, including dimethyl-guanine (peak 2, m/z 179), three tautomers of trimethyl-guanine (peaks 3, 5 and 9, *m/z* 193) and four tautomers of tetramethyl-guanine (peaks 4, 6, 7 and 8, *m/z* 207). The mass spectra corresponding to peaks 2 to 9 are given in Figure S3 (see the supplementary material). Some other compounds are also detected with lower intensity, such as glycocyamine (RT=20.0 min) and mesalamine (RT=21.3 min).

The intensity of peaks generally increased with the increase in temperature. However, Peak 9 (trimethyl-guanine) was observed only at 400°C while peaks 1, 4, and 6

were not detected at 400 °C. Peak 1 corresponds to the methylated xanthine from the degradation of guanine as mentioned before, which probably occurs at temperatures above 400°C. Peaks 4 and 6 correspond (according to the *m/z* ratios of the molecular ion) to two tautomers of tetramethylated guanine. Trimethyl-guanine (peak 9) certainly coeluted with peak 7 at a higher temperature than 400°C. All detected compounds of guanine can be seen in Table 2.

Compared with adenine and cytosine, guanine is the third nucleobase with a primary amine moiety that can decrease the effectiveness of GC-MS analysis dramatically. At low temperatures of 250°C and 300°C, no thermochemolysis products of guanine in TMAH solutions were detected. Thus, we conclude that thermochemolysis reactions barely occur at temperatures lower than 400°C to 600°C. However, it was difficult to analyze guanine because of the numerous derivatives and the spread and often coeluted peaks.

3.1.5 Xanthine

2.2×10^{-3} mol·L^{-1} of xanthine in TMAH was used. Like thymine and uracil, xanthine responded well at the lower concentration. The chromatograms of xanthine obtained at 400°C, 500°C, and 600°C are shown in Figure 5. The only derivative of xanthine detected was 1, 3, 7-trimethylxanthine (caffeine; TR = 28.90 min at 600°C), which is the trimethylated form of xanthine, with a mass of the molecular ion of 194 g.mol^{-1}. As shown in Figure 5, the signal intensity increased with the increase of thermochemolysis temperature, and the optimal analysis temperature was 600°C.

3.1.6 Hypoxanthine

The chromatograms of hypoxanthine in TMAH solutions obtained at the pyrolyzed temperature of 400 °C, 500 °C and 600 °C are given in Figure 5. There are two derivatives of hypoxanthine, 1, 7-dimethyl-hypoxanthine (the main peak, at a retention time of 29.60 min at 600 °C) with relatively higher intensity and dimethyl-hypoxanthine tautomer (at a retention time of 31.52 min at 600 °C) with lower intensity. There were virtually no compounds released at 400 °C; the derivative compounds were only detectable above 500 °C. Some nucleobases—such as thymine, uracil, xanthine, and hypoxanthine, which were completely methylated on their intracyclic nitrogen (trimethyl substitutions of xanthine, dimethyl substitutions of hypoxanthine)—were detectable at low concentration.

3.2 Thermochemolysis patterns

All derivatized compounds detected are listed in Table 2. It can be seen that most of the derivatized compounds are the methylated products of the seven nucleobases. However, the nucleobases exist in different tautomeric forms because of the presence of π bonds and non-binding doublet. Tautomerism of the nucleobases is manifested by the delocalization of π electrons and the displacement of a hydrogen atom, thus the protons move from one site to another and transform from a functional group into another within the same molecule. There are two kinds of prototropic tautomerisms in nucleobases, one is the keto-enol tautomerism, where the ketone form is in equilibrium with the enol form (C=C with OH functional group in the vinyl position) and imine-enamine tautomerism, where the enamine forms (ethylenic alpha amine) and is in equilibrium with the ketimine form. Imine-enamine tautomerism is involved in all the above-mentioned nucleic bases. In all cases, different tautomeric forms of the same molecule coexist in an equilibrium that can be displaced towards one or the other, depending on pH, temperature, and solvent (with the pH value of dependent on the solvent). As a result, the position of the methylation sites varies with labile protons moving within the molecule. The proton is substituted by the methyl group and acquires the stable form. Hence, there are different methylation forms of nucleobases besides the canonical form.

For adenine, there are three possible methylation sites: one is at the N9 position, and the two others are at the N10 position, as shown in Figure 6. During the TMAH thermochemolysis process, dimethyl and trimethyl-adenine were detected. As shown in Figure 1A, the peak intensity of N, N, 9-trimethyl-adenine, which is the methylated products of the N9H form of adenine, was the highest. This is because the N9H tautomeric form is the major form of adenine and the N7H and N3H tautomeric forms are minor forms [36,37], however, the formation and observation of the 7(H) and 3(H) tautomers may be facilitated in the solution phase [38]. The N, 9-dimethyl-adenine form was the second highest peak at 600°C, which is also the methylated derivative of the N9H form of adenine. According to Fonseca Guerra [37], the natural N9H form of adenine was the most stable tautomer of the twelve systems studied and the stability of the N3H form of adenine was preceded only by the N9H form. Hence, the methylated derivatives of the N7H form of adenine were also detected during the TMAH thermochemistry process. 7-methyl-adenine and N, N, 3-trimethyl-adenine are the monomethyl substitution and trimethyl substitution, respectively. The content of trimethyl substitution products was higher than that of

monomethyl substitution or dimethyl substitution, which demonstrates that the trimethyl substitution products are more stable than other substitutions.

For thymine and uracil, only the methylated canonical forms were detected during the TMAH thermochemolysis. Although the keto form, enol form, and keto-enol tautomerism of thymine and uracil have been reported before [39–43], the canonical tautomer is thermodynamically more stable than all enol and dienol forms. Therefore, the canonical tautomer is the main thermochemolysis product of thymine and uracil in TMAH solution as shown in Figure 2. There are two active sites (secondary amine) on the thymine and uracil at N1 and N3 position, and they are replaced by a methyl group during TMAH thermochemolysis, as shown in Figure 6. The absence of methylation of the enolactic tautomeric forms of thymine and uracil is explained by the keto-enolic thermodynamic equilibrium shifted very strongly in favor of the ketone form. Therefore, 1, 3-dimethyl-thymine and 1, 3-dimethyl-uracil are detected, which are the most stable methylated products of thymine and uracil, respectively.

For cytosine, various tautomeric forms have been detected in different studies [44]. These are mainly amino-oxo, imino-oxo and amino-hydroxy forms of cytosine. The computational investigations have obtained different results via different methods. The amino-hydroxy tautomer was found to be the most stable in the gas phase, and the canonical amino-oxo form seemed to be the most stable one under aqueous solvation [45–47]. In TMAH solutions, methylated cytosine in canonical forms, such as trimethyl-cytosine and N, N, N'-trimethyl-cytosine, were the main products, which means that the canonical amino-oxo form of cytosine was also the most stable form in TMAH solutions. In addition, there are relatively weak peaks of the methylated compounds of cytosine in enol form, such as 4-amino-2-methoxy-pyrimidine (Figure 3, peak 1), 2-(dimethylamino)-1-phenyl-1-propanone (Figure 3, RT=21.36 min), 2-amino-4,6-dimethyl-pyrimidine(Figure 3, RT=24.40 min), and 3,4,5-trimethyoxybenzylamine(Figure 3, RT=28.95 min).

According to the guanine thermochemolysis chromatograph, there were more tautomeric transformations of guanine compared with cytosine. The keto (amino-oxo) and enol (amino-hydroxy) forms were the most stable forms according to the value of the relative energies of the tautomers of guanine and both exist in approximately equal concentrations[48]. The 7-, 9- or 1-methylations of guanine and dimethylated guanine were identified by some researchers using Infrared spectroscopy (IR) [49–51]. In our study, the dimethyl, trimethyl, and tetramethyl-guanine were detected and are probably the methylated products of the 9NH or 7NH form of guanine. These are the products of

methylation and oxidative deamination of guanine, as shown in Figure 6. The deamination and oxidation of guanine to xanthine, which was methylated by TMAH on its three labile hydrogen atoms, 1, 3, 7-trimethylxanthine (caffeine) was consequently detected.

Hypoxanthine and guanine are structurally very similar, and they play an important role during the DNA replication process and can form a base pair with cytosine. From this point of view, therefore, hypoxanthine also has keto-enol and prototropic tautomeric phenomena [49,52,53]. The keto-N7H and keto-N9H forms are the most stable forms of hypoxanthine, and there are two possible substituted sites at N1 and N7 position. According to the chromatograph of the hypoxanthine, dimethyl products were detected in excessive TMAH solutions, which must be the most stable methylated forms of hypoxanthine. Xanthine also has a similar structure with hypoxanthine, there are three possible substituted sites, N1, N3, and N7 position. Therefore, the trimethyl form must be the most stable methylated product of xanthine.

TMAH can improve pyrolysis by mild thermal decomposition and a selective cleavage of ester and ether bonds as opposed to random and uncontrolled thermal decomposition and fragmentation of the organic material via standard pyrolysis [32]. During the thermochemolysis of nucleobases, tautomerism transformations of all nucleobases occur. TMAH also plays an important role by promoting the cracking of N-H bonds at 600°C; thus the fragments of nucleobases in stable tautomerism forms are methylated and form the stable derivatives. It is not surprising that the methylated compounds of all of the nucleobases detected by GC-MS were their most stable forms in a or aqueous phase. Their characteristic methylated compounds are summarized in Table 3. Thymine, uracil, and xanthine each have only one unique characteristic peak individually; however, the case is different for the other four kinds of nucleobases, as at least two characteristic peaks correspond to their identification. Hence, the characteristic thermochemolysis products of each nucleobases (list in Table 3) could be used to identify these nucleobases on MOMA or SAM.

3.3 Analysis of the mixture of nucleobases

The chromatogram corresponding to the thermochemolysis of 0.1 µl of the nucleobases mixture detected at 600°C is shown in Figure 7, and the concentration and the amount of each nucleobase are shown in Table 1. The methylated derivatives of the nucleobases detected as well as their retention times and the associated parent molecules are listed in Table 4. The characteristic peak of each nucleobase appears at different

retention times but in the same order as shown in Table 3 because the temperature programming of the column was changed (starting at 50°C held for 2 min, then at a heating rate of 3°C min^{-1} up to 170°C then raised to 300°C at a rate of 10°C min^{-1} and maintained for 3 min).

Each of the seven nucleobases was identified when more samples were used, and the derivatives of thymine, uracil, xanthine, and hypoxanthine were uniquely identified. All compounds were the methylated canonical form of all their labile hydrogens, except for dimethyl-hypoxanthine, where a second tautomer was detected. Compared with five different methylated derivatives during the injection of 15 nmol of cytosine, only the trimethylated canonic form, one of the weakest peaks, was detected.

Compared with the nine derivatives obtained by the pyrolysis of 2.5 nmol of adenine, two derivatives were observed in the 0.1uL mixture. The predominant peak of adenine corresponded to the trimethylated canonical structure whereas the second peak of low intensity was a trimethyl tautomer form of adenine. A single derivative of guanine with low intensity was also detected. Guanine was the only nucleobase whose methylation was uncompleted: only three out of four labile hydrogens were replaced by methyl groups. A small amount of guanine was detected due to the oxidative deamination of guanine which degraded to xanthine. This is also the reason for the high intensity of the 1, 3, 7-trimethyl xanthine peak, which is preponderant in the chromatogram. In addition, there were no interactions among seven nucleobases, because there were nearly no obvious cracking reactions of nucleobases at low temperature.

We conclude that among the seven nucleobases, four peaks with high intensity should be easily detected by thermochemolysis if there were in the nanomole range or higher amounts in natural samples, including xanthine, adenine, thymine, and uracil. Due to the oxidative deamination of guanine to xanthine during thermochemolysis, detection of xanthine may also be a strong indicator of the presence of guanine.

3.4 Quantification study of nucleobases

Limit of detection (LOD) and limit of quantitation (LOQ) can be determined by signal-to-noise (S/N) and relative standard deviation (RSD). A value of S/N=3 or RSD≈17% can be used for LOD, S/N=10 or RSD≈5% can be used for LOQ[54–58]. Therefore, S/N=3 and S/N=10 were used to determine the LOD and LOQ of the nucleobases in this study. The calibration curves of nucleobases were obtained by plotting peak area ratio values (the

peak area of nucleobases/internal standard peak area) against the concentration ratio of nucleobases (nucleobases amount/internal standard amount).

Naphthalene-d8 was used as the internal standard, and the stability was tested five times at 600°C. The average peak area of Naphthalene-d8 was about 3.2×10^8 $cps.s^{-1}.min$; the % RSD was about 6.96%, which is calculated by the average divided by the standard deviation (about 2.26×10^7 $cps.s^{-1}.min$). Moreover, we observed no degradation of the Naphthalene-d8 after a pyrolysis at 600°C compared at the same GCMS injection at 270°C in the GC injector. Therefore, because of its good stability at 600°C, naphthalene-d8 was used as our internal standard.

To confirm the value of LOD and LOQ for each sample, all experiments were run in triplicate. The LOD and LOQ value and the linearity of standard curves for each nucleobase are summarized in Table 5. We conclude that the relativity of standards curves is not a sufficient metric, as shown in Figure S4 (see the supplementary material), possibly because of the interaction of TMAH with the column.

4. Conclusion

We conclude that TMAH thermochemolysis allows for the methylation and detection of the seven nucleobases investigated in this study. The optimal temperature for nucleobase thermochemolysis under SAM and MOMA conditions was 600°C. Trimethyl-adenine, 1,3-dimethyl-thymine, 1,3-dimethyl-uracil, trimethyl-cytosine, 1, 3, 7-trimethyl-xanthine (caffeine) and dimethyl-hypoxanthine were the main methylated products of adenine, thymine, uracil, cytosine, guanine, and xanthine, and hypoxanthine, respectively, because they are the most stable forms of each nucleobase in a gas phase or in aqueous form. In addition, there was no interaction among these nucleobases when pyrolyzed in a mixture, which makes it possible to detect different nucleobases in natural samples. However, oxidative deamination of guanine to xanthine may also lead to misidentifications when these molecules are detected in unknown samples. The presence of the $-NH_2$ group also induces a tautomerization of the molecule and therefore multiplies the number of methylated derivatives detected for each of these nitrogenous bases. Based on the values of LOD and LOQ for each nucleobase, adenine, thymine, and uracil are easier to be detected, as they can be detected at about 0.15 nmol. It is possible, though more difficult, to detect the guanine, cytosine and hypoxanthine (LOD=0.40, 0.55 and 0.75 nmol, respectively). However, the complex matrix needs to be pretreated with the extraction process in order to detect the nucleobases on Mars, the LOD may be higher than value of our study.

According to the observed retention times, potential coelutions may occur in studies of mixtures of different nucleobases at high concentrations. This was mainly due to the spread of cytosine peaks (between 30 and 37 min), guanine (between 30.5 and 32.2 min) and, to a lesser extent, adenine (around 30.7 min). There was also a coelution between dimethyl-uracil (at the retention time of 20.40 min) and dimethyl cytosine (at the retention time of 20.33 min), though these problems can be avoided by using a lower column heating rate. In addition, some nucleobases with a primary amine function ($-NH_2$), such as adenine, guanine, and cytosine, displayed a spreading of the chromatographic peaks because of the interaction of the amine group with the stationary phase of the column.

All nucleobases detected here had relatively high retention times in the range of 27 to 47 minutes of elution with the temperature programming of the column used here (50°C for 2 min, then 3°C min^{-1} up to 170 °C then 10 °C.min^{-1} up to 300°C, held 3 min), the column of MOMA flight instruments is shorter than the column used herein; Therefore, it would be more easier to detect the nucleobases related compounds *in situ*, such as trimethyl-guanine, trimethyl-adenine and dimethyl-hypoxanthine; and signs of adenine and hypoxanthine can be detected when the temperature is lower than 250°C. What's more, the separation and detection abilities of the device on SAM and MOMA will be tested in the coming missions, and SAM is ready to do the TMAH analysis experiments soon; MOMA will be ready in 2020. In a word, it would be possible to detect the nucleobases if the samples are fresh and their content are higher than the limit of detection of the device on Mars and MOMA, and our results laid the data foundation for the implementation of SAM and MOMA missions.

Acknowledgements

The authors are grateful for the support of the French Space Agency (Centre National d'Etudes Spatiales) and SAM/MOMA funding. R. Navarro Gonzalez has received funds from the Universidad Nacional Autónoma de México (PAPIIT IN111619).

References

[1] Z. Martins, O. Botta, M.L. Fogel, M.A. Sephton, D.P. Glavin, J.S. Watson, et al., Extraterrestrial nucleobases in the Murchison meteorite, Earth Planet. Sci. Lett. 270 (2008) 130–136.

[2] O. Botta, J.L. Bada, Extraterrestrial organic compounds in meteorites, Surv. Geophys. 23 (2002) 411–467.

[3] M.A. Sephton, Organic compounds in carbonaceous meteorites, Nat. Prod. Rep. 19 (2002) 292–311.

[4] R. Hayatsu, Orgueil Meteorite: Organic Nitrogen Contents., Science (80-.). 146 (1964) 1291–3.

[5] R. Hayatsu, M.H. Studier, A. Oda, K. Fuse, E. Anders, Origin of organic matter in early solar system - II. Nitrogen compounds, Geochim. Cosmochim. Acta. 32 (1968) 175–190.

[6] R. Hayatsu, M.H. Studier, L.P. Moore, E. Anders, Purines and triazines in the Murchison meteorite, Geochim. Cosmochim. Acta. 39 (1975) 471–488.

[7] W. van der Velden, A.W. Schwartz, Search for purines and pyrimidines in the Murchison meteorite, Geochim. Cosmochim. Acta. 41 (1977) 961–968.

[8] P.G. Stoks, A.W. Schwartz, Uracil in carbonaceous meteorites, Nature. 282 (1979) 709–710.

[9] P.G. Stoks, A.W. Schwartz, Nitrogen-heterocyclic compounds in meteorites: significance and mechanisms of formation, Geochim. Cosmochim. Acta. 45 (1981) 563–569.

[10] M. a. Sephton, Organic Geochemistry and the Exploration of Mars, J. Cosmol. 5 (2010) 1141–1149.

[11] M.P. Callahan, A.S. Burton, J.E. Elsila, E.M. Baker, K.E. Smith, D.P. Glavin, et al., A search for amino acids and nucleobases in the Martian meteorite Roberts Massif 04262 using liquid chromatography-mass spectrometry, Meteorit. Planet. Sci. 48 (2013) 786–795.

[12] V.A. Basiuk, J. Douda, Pyrolysis of simple amino acids and nucleobases: survivability limits and Implications for extraterrestrial dielivery, Planet. Sp. Sci. 36. 47 (1999) 577–584.

[13] D.P. Glavin, H.J. Cleaves, A. Buch, M. Schubert, A. Aubrey, J.L. Bada, et al., Sublimation extraction coupled with gas chromatography-mass spectrometry: A new technique for future in situ analyses of purines and pyrimidines on Mars, Planet. Space Sci. 54 (2006) 1584–1591.

[14] C. AbbasHawks, K.J. Voorhees, T.L. Hadfield, In situ methylation of nucleic acids using pyrolysis mass spectrometry, Rapid Commun. Mass Spectrom. 10 (1996) 1802–1806.

[15] A.J. Kossa, W. C., MacGee, J., Ramachandran, S., Webber, Pyrolytic Methylation/Gas Chromatography: A Short Review, J. Chromatogr. Sci. 17 (1979)

177–187.

[16] S.W. Frazier, K.O. Nowack, K.M. Goins, F.S. Cannon, L.A. Kaplan, P.G. Hatcher, Characterization of organic matter from natural waters using tetramethylammonium hydroxide thermochemolysis GC-MS, J. Anal. Appl. Pyrolysis. 70 (2003) 99–128.

[17] M. Le Meur, L. Mansuy-Huault, C. Lorgeoux, A. Bauer, R. Gley, D. Vantelon, et al., Spatial and temporal variations of particulate organic matter from Moselle River and tributaries: A multimolecular investigation, Org. Geochem. 110 (2017) 45–56.

[18] J. Templier, F. Miserque, N. Barré, F. Mercier, J.P. Croué, S. Derenne, Is nitrogen functionality responsible for contrasted responses of riverine dissolved organic matter in pyrolysis?, J. Anal. Appl. Pyrolysis. 97 (2012) 62–72.

[19] C. Rodier, R. Sternberg, C. Szopa, A. Buch, M. Cabane, F. Raulin, Search for organics in extraterrestrial environments by in situ gas chromatography analysis, Adv. Sp. Res. 36 (2005) 195–200.

[20] D.J. Clifford, D.M. Carson, D.E. McKinney, J.M. Bortiatynski, P.G. Hatcher, A new rapid technique for the characterization of lignin in vascular plants: thermochemolysis with tetramethylammonium hydroxide (TMAH), Org. Geochem. 23 (1995) 169–175.

[21] A.D. Hendricker, K.J. Voorhees, Amino acid and oligopeptide analysis using Curie-point pyrolysis mass spectrometry with in-situ thermal hydrolysis and methylation: Mechanistic considerations, J. Anal. Appl. Pyrolysis. 48 (1998) 17–33.

[22] J. Templier, N. Gallois, S. Derenne, Analytical TMAH pyrolysis of dipeptides: Formation of new complex cyclic compounds related to the presence of the peptide bond, J. Anal. Appl. Pyrolysis. 104 (2013) 684–694.

[23] N. Gallois, J. Templier, S. Derenne, Pyrolysis-gas chromatography–mass spectrometry of the 20 protein amino acids in the presence of TMAH, J. Anal. Appl. Pyrolysis. 80 (2007) 216–230.

[24] S.S. Choi, J.E. Ko, Dimerization reactions of amino acids by pyrolysis, J. Anal. Appl. Pyrolysis. 89 (2010) 74–86.

[25] N. Gallois, J. Templier, S. Derenne, Limitations in interpreting TMAH thermochemolysis of natural organic matter via consideration of glycine and alanine derivatives, Org. Geochem. 41 (2010) 1338–1340.

[26] X. Zang, J.C. Brown, J.D.H. Van Heemst, A. Palumbo, P.G. Hatcher, Characterization of amino acids and proteinaceous materials using online tetramethylammonium hydroxide (TMAH) thermochemolysis and gas

chromatography-mass spectrometry technique, J. Anal. Appl. Pyrolysis. 61 (2001) 181–193.

[27] C.A. Joll, T. Huynh, A. Heitz, Off-line tetramethylammonium hydroxide thermochemolysis of model compound aliphatic and aromatic carboxylic acids: Decarboxylation of some ortho- and/or para- substituted aromatic carboxylic acids, J. Anal. Appl. Pyrolysis. 70 (2003) 151–167.

[28] J.C. Del Río, M.A. Olivella, H. Knicker, F.X.C. De Las Heras, Preservation of peptide moieties in three Spanish sulfur-rich Tertiary kerogens, Org. Geochem. 35 (2004) 993–999.

[29] A. Riboulleau, T. Mongenot, F. Baudin, S. Derenne, C. Largeau, Factors controlling the survival of proteinaceous material in Late Tithonian kerogens (Kashpir Oil Shales, Russia), Org. Geochem. 33 (2002) 1127–1130.

[30] H. Knicker, J.C.D. Río, P.G. Hatcher, R.D. Minard, Identification of protein remnants in insoluble geopolymers using TMAH thermochemolysis/GC-MS, Org. Geochem. 32 (2001) 397–409.

[31] A. Buch, D.P. Glavin, R. Sternberg, C. Szopa, C. Rodier, R. Navarro-González, et al., A new extraction technique for in situ analyses of amino and carboxylic acids on Mars by gas chromatography mass spectrometry, Planet. Space Sci. 54 (2006) 1592–1599.

[32] P.R. Mahaffy, C.R. Webster, M. Cabane, P.G. Conrad, P. Coll, S.K. Atreya, et al., The sample analysis at mars investigation and instrument suite, Space Sci. Rev. 170 (2012) 401–478.

[33] F. Goesmann, W.B. Brinckerhoff, F. Raulin, W. Goetz, R.M. Danell, S.A. Getty, et al., The Mars Organic Molecule Analyzer (MOMA) Instrument: Characterization of Organic Material in Martian Sediments, Astrobiology. 17 (2017) 655–685.

[34] A. Venema, R.C.A. Boom-Van Geest, In-situ hydrolysis/methylation pyrolysis CGC for the characterization of polyaramides, J. Microcolumn Sep. 7 (1995) 337–343.

[35] P.M. Buch, A., Morisson, M., Szopa, C., Millan, M., Freissinet, C., He, Y., D. Glavin, J.Y. Bonnet, D. Coscia, A.J. Williams, F. Stalport, F. Raulin, M. Stambouli, S. Teinturier, R. Navarro-González, C. Malespin, Optimization of the TMAH thermochemolysis technique for the detection of trace organic matter on Mars bythe SAM and MOMA-Pyr-GC-MS experiment., in: 49th LPSC Lunar Planet. Sci. Conf., 2018: pp. 5–6.

[36] M. Hanus, M. Kabeláč, J. Rejnek, F. Ryjáček, P. Hobza, Correlated ab Initio Study of Nucleic Acid Bases and Their Tautomers in the Gas Phase, in a Microhydrated Environment, and in Aqueous Solution. Part 3. Adenine, J. Phys. Chem. B. 108 (2004) 2087–2097.

[37] C. Fonseca Guerra, F.M. Bickelhaupt, S. Sana, F. Wang, Adenine tautomers: Relative stabilities, ionization energies, and mismatch with cytosine, J. Phys. Chem. A. 110 (2006) 4012–4020.

[38] H. Kim, D. Ahn, S. Chung, S.K. Kim, S. Lee, Tautomerization of Adenine Facilitated by Water : Computational Study of Microsolvation, 9 (2007) 8007–8012.

[39] M.A. Morsy, A.M. Al-Somali, A. Suwaiyan, Fluorescence of Thymine Tautomers at Room Temperature in Aqueous Solutions, J. Phys. Chem. B. 103 (1999) 11205–11210.

[40] M.K. Shukla, J. Leszczynski, Interaction of water molecules with cytosine tautomers: An excited-state quantum chemical investigation, J. Phys. Chem. A. 106 (2002) 11338–11346.

[41] Y. Tsuchiya, T. Tamura, M. Fujii, M. Ito, Keto-enol tautomer of uracil and thymine, J. Phys. Chem. 92 (1988) 1760–1765.

[42] B.B. Brady, L.A. Peteanu, D.H. Levy, The electronic spectra of the pyrimidine bases uracil and thymine in a supersonic molecular beam, Chem. Phys. Lett. 147 (1988) 538–543.

[43] J. Rejnek, M. Hanus, M. Kabeláč, F. Ryjáček, P. Hobza, Correlated ab initio study of nucleic acid bases and their tautomers in the gas phase, in a microhydrated environment and in aqueous solution. Part 4. Uracil and thymine, Phys. Chem. Chem. Phys. 7 (2005) 2006–2017.

[44] M.K. Shukla, J. Leszczynski, Tautomerism in nucleic acid bases and base pairs: A brief overview, Wiley Interdiscip. Rev. Comput. Mol. Sci. 3 (2013) 637–649.

[45] G. Fogarasi, P.G. Szalay, The interaction between cytosine tautomers and water: An MP2 and coupled cluster electron correlation study, Chem. Phys. Lett. 356 (2002) 383–390.

[46] M.N. Manalo, A.C. de Dios, R. Cammi, Solvent Effects on 15 N NMR Shielding of 1,2,4,5-Tetrazine and Isomeric Tetrazoles: Continuous Set Gauge Transformation Calculation Using the Polarizable Continuum Model, J. Phys. Chem. A. 104 (2000) 9600–9604.

[47] S.A. Trygubenko, T. V. Bogdan, M. Rueda, M. Orozco, F.J. Luque, J. Šponer, et al.,

Correlated ab initio study of nucleic acid bases and their tautomers in the gas phase, in a microhydrated environment and in aqueous solution : Part 1. Cytosine, Phys. Chem. Chem. Phys. 4 (2002) 4192–4203.

[48] I.R. Gould, I.H. Hillier, Accurate calculations of the relative energies of the tautomers of cytosine and guanine, Chem. Phys. Lett. 161 (1989) 185–187.

[49] G.G. Sheina, S.G. Stepanian, E.D. Radchenko, Y.P. Blagoi, IR spectra of guanine and hypoxanthine isolated molecules, J. Mol. Struct. 158 (1987) 275–292.

[50] S.M. Szczepaniak K, Matrix isolation infrared studies of nucleic acid constituents: Part 4. Guanine and 9-methylguanine monomers and their keto-enol tautomerism, J. Mol. Struct. 156 (1987) 29–42.

[51] M.S. and W.B.P. Krystyna Szczepaniak, Infrared Studies and the Effect of Ultraviolet Irradtation on the tautomers of 9-methylguanine isolated in an argon matrix, Chem. Phys. Lett. 153 (1988) 39–44.

[52] R. Ramaekers, G. Maes, L. Adamowicz, A. Dkhissi, Matrix-isolation FT-IR study and theoretical calculations of the vibrational, tautomeric and H-bonding properties of hypoxanthine, J. Mol. Struct. 560 (2001) 205–221.

[53] G. Bazsó, G. Tarczay, G. Fogarasi, P.G. Szalay, Tautomers of cytosine and their excited electronic states: a matrix isolation spectroscopic and quantum chemical study., Phys. Chem. Chem. Phys. 13 (2011) 6799–807.

[54] D.A. Arinbruster, D. Tillman, L.M. Hubbs, Limit of Detection (LOD)/ Limit of Quantitation (LOQ): Comparison of the Empirical and the Statistical Methods Exemplified with GC-MS Assays of Abused Drugs, Clin. Chem. 40 (1994) 1233–1238.

[55] H.Y. Shen, Simultaneous screening and determination eight phthalates in plastic products for food use by sonication-assisted extraction/GC-MS methods, Talanta. 66 (2005) 734–739.

[56] K.G.J. Nierop, T.R. Filley, Simultaneous analysis of tannin and lignin signatures in soils by thermally assisted hydrolysis and methylation using13C-labeled TMAH, J. Anal. Appl. Pyrolysis. 83 (2008) 227–231.

[57] F.Q. Yang, J. Guan, S.P. Li, Fast simultaneous determination of 14 nucleosides and nucleobases in cultured Cordyceps using ultra-performance liquid chromatography, Talanta. 73 (2007) 269–273.

[58] F.Q. Yang, D.Q. Li, K. Feng, D.J. Hu, S.P. Li, Determination of nucleotides, nucleosides and their transformation products in Cordyceps by ion-pairing reversed-

phase liquid chromatography-mass spectrometry, J. Chromatogr. A. 1217 (2010) 5501–5510.

Table captions:

Table 1: Detailed information of nucleobases used in this study.

Table 2: The thermochemolysis products detected from different nucleobases.

Table 3: The characteristic peaks of the seven nucleobases in TMAH.

Table 4: The methylated components detected from pyrolysis of the nucleobases mixture.

Table 5: LOD and LOQ values and the standard curves of each nucleobase.

Table 1: Detailed information of nucleobases used in this study.

Nucleobases		Amount (nmol)	Injection volume (µl)	Concentration ($mol \cdot L^{-1}$)
Adenine		0.125	0.05	2.5×10^{-3}
		0.25	0.10	2.5×10^{-3}
		2.50	0.01	0.25
Thymine		0.26	0.10	2.6×10^{-3}
Uracil		0.26	0.10	2.6×10^{-3}
Cytosine		15.0	0.05	0.3
Guanine		11.0	0.05	0.22
Xanthine		0.22	0.10	2.2×10^{-3}
Hypoxanthine		1.20	0.10	1.2×10^{-2}
	Adenine	18	0.10	0.18
	Guanine	31	0.10	0.31
	Cytosine	43	0.10	0.43
Mixtures	Thymine	0.19	0.10	1.86×10^{-3}
	Uracile	0.19	0.10	1.86×10^{-3}
	Xanthine	0.18	0.10	1.57×10^{-3}
	Hypoxanthine	0.86	0.10	8.57×10^{-3}

Table 2: The thermochemolysis products detected from different nucleobases.

Nucleobases	Retention time	Base peak	Masses of fragments* and relative abundance: *m/z* (%)	Compounds
Adenine	24.84	149	**149(100)**, 44(86),148(27), 42(26), 122(23), 121(11), 207(10), 53(10), 68(9)	7-methyl-adenine
	25.90	163	**163(100)**, 107(72), 135(69),134(55), 80(31), 133(30), 162(27)	N, 9-dimethyl-adenine
	26.54	148	**148(100)**, 162(38), 177(34), 107(22), 133(22), 135(15), 106(12)	N, N, 9-trimethyl-adenine
	27.93	162	**162(100)**, 163(45), 42(41), 82(17), 121(17), 108(15), 135(15)	dimethyl-adenine
	28.33	176	**176(100)**, 177(56), 148(33), 147(23), 67(24), 133(20), 42(16)	trimethyl-adenine
	28.54	162	**162(100)**, 177(91), 176(44), 148(37), 42(38), 133(30), 94(30), 121(29)	trimethyl-adenine
	29.47	177	**177(100)**, 135(93), 44(92), 42(83), 121(63), 162(44), 133(41), 107(32)	trimethyl-adenine
	30.54	177	**177(100)**, 135(81), 162(66), 44(50), 42(49), 108(21), 120(15), 133(10)	trimethyl-adenine
	30.72	148	**148(100)**, 162(59), 177(34), 134(30), 107(21), 135(21), 119(18)	N, N, 3-trimethyl-adenine
Thymine	21.49	154	**154(100)**, 68(93), 69(46), 42(41), 97(15), 96(11), 56(9.0)	1,3-dimethyl-thymine
Uracil	20,40	140	**140(100)**, 42(62), 55(53), 83(43), 82(29), 54(15), 56(11)	1,3-dimethyl-uracile
Cytosine	18.90	95	**95(100)**, 125(72), 124(50), 68(38), 67(19), 96(16), 41(15)	4-amino-2-methoxy-pyrimidine
	20.33	139	**139(100)**, 109(80), 108(63), 138(59), 95(53), 110(32), 81(22)	dimethyl-cytosine
	20.58	153	**153(100)**, 124(39), 152(7), 123(7), 138(5), 44(1.14), 95(1)	trimethyl-cytosine
	21.36	72	**72(100),** 44(68), 153(15), 71(17), 207(10)	2-(dimethylamino)-1-phenyl-1-propanone
	21.84	124	**124(100)**, 153(75), 55(44), 95(43), 96(36), 42(29), 54(13)	N, N, N'-trimethyl-cytosine
	24.40	123	**123(100),** 96(76), 42(72), 154(58), 44(46), 80(37)	2-amino-4,6-dimethyl-pyrimidine
	28.95	166	**166(100),** 44(65), 197(26), 139(23), 109(20), 83(19), 138(11)	3,4,5-trimethyoxybenzylamine

	30-37	153	**153(100)**, 124(44), 42(13), 138(11), 125(6), 154(5), 123(4), 152(3)	trimethyl-cytosine
Guanine	19.98	122	**122(100),** 123(7)	Glycocyamine
	21.28	136	**136(100),** 108(7)	Mesalamine
	28.88	194	**194(100)**, 109(62), 55(44), 82(29), 44(27), 67(26), 193(20),	Caffeine (1, 3, 7-trimethyl-xanthine)
	29.18	179	**179(100)**, 150(49), 42(45), 178(40), 149(37), 107(27), 44(24)	dimethyl-guanine
	30.01	193	**193(100)**, 149(65), 164(51), 165(52), 42(48), 192(21), 163(28)	trimethyl-guanine
	30.20	207	**207(100)**, 178(94), 163(71), 192(54), 42(54), 149(48), 44(31), 164(28)	tetramethyl-guanine
	30.51-32.21	193	**193(100)**, 82(32), 55(31), 42(29), 67(27), 164(24), 109(22)	trimethyl-guanine
	31.25	178	**178(100)**, 67(70), 207(67), 109(52), 55(41), 82(40),137(36)	tetramethyl-guanine
	32.99	164	**164(100)**, 207(65) , 123(64), 163(60), 178(50), 136(40), 67(23)	tetramethyl-guanine
	37.10-38.00	207	**207(100)**, 44(66), 163(49), 178(30), 42(30), 164(29), 67(24), 109(14)	tetramethyl-guanine
	35.12	193	**193(100)**, 164(92), 123(60), 67(32), 163(24), 110(28), 136(20)	trimethyl-guanine(400°C)
Xanthine	28.90	194	**194(100)**, 109(52), 67(47), 55(38), 82(32), 193(21), 42(13)	1, 3,7-trimethyl-xanthine (caffeine)
Hypoxanthine	29.60	164	**164(100)**, 163(58), 42(44), 110(14), 68(14), 67(13), 53(9)	1,7-dimethyl-hypoxanthine
	31.52	164	**164(100)**, 135(56), 42(56), 82(21), 163(16), 67(10), 108(9)	dimethyl-hypoxanthine

Note: * the masses of the main fragments are in order of decreasing abundance; the mass marked in bold is the main peak.

Table 3: The characteristic peaks of the seven nucleobases in TMAH.

Nucleobase	Retention time (min)	Characteristic peak
Uracil	20.40	1, 3-dimethyl-uracil
Thymine	21.49	1, 3-dimethyl-thymine
Cytosine	21.84	N, N, N'-trimethyl-cytosine
Adenine	26.54	N, N, 9-trimethyl-adenine
Xanthine	28.90	1, 3, 7-trimethylxanthine
Hypoxanthine	29.60	1,7-dimethyl-hypoxanthine
Guanine	30.01	trimethyl-guanine

Table 4: The methylated components detected from pyrolysis of the nucleobases mixture.

Retention time (min)	Component	Parent compound
27.79	1, 3-dimethyl-uracil	Uracil
29.94	1, 3-dimethyl-thymine	Thymine
31.19	N, N, N'-trimethyl-cytosine	Cytosine
39.81	N, N, 9-trimethyl-adenine	Adenine
43.78	1, 3, 7-trimethyl-xanthine (caffeine)	Xanthine
44.34	1,7-dimethyl-hypoxanthine	Hypoxanthine
45.66	trimethyl-guanine	Guanine
45.91	trimethyl-adenine	Adenine

Table 5 LOD and LOQ values and the standard curves of each nucleobase.

Nucleobases	LOD(nmol)	LOQ(nmol)	Linear regression equation	Correlation coefficient
Adenine	0.075	0.15	$y = 0.0190x - 0.00001$	0.9975
Thymine	0.075	0.15	$y = 0.0250x + 0.0259$	0.9989
Uracil	0.075	0.30	$y = 0.0023x - 0.0038$	0.9973
Cytosine	0.55	0.65	$y = 0.0004x - 0.0015$	0.9898
Guanine	0.40	0.675	$y = 0.0002x - 0.0004$	0.9958
Xanthine	0.15	0.275	$y = 0.0090x + 0.0018$	0.9992
Hypoxanthine	0.75	0.95	$y = 0.0006x - 0.0019$	0.9805

Figure captions:

Figure 1: Chromatograms of adenine thermochemolysis in TMAH performed at different temperature (A: 600°C, 500°C, and 400°C.) and different amount (B: (a) 0.125 nmol (b) 0.25 nmol (c) 2.5 nmol). Chromatogram A: Peak 1: 3-methyl-3H-purin-6-amine; Peak 2: N, 9-dimethyl-9H-purin-6-amine; Peak 3: N, N, 9-trimethyl-9H-purin-6-amine; Peak 4: N, N, 3-trimethyl-3H-purin-6-amine; Peak 5: dimethyl adenine; Peak 6,7,8,9: trimethyl adenine. Note the break in the intensity scale to show the height of peak 3; Chromatogram B: Peak 1: 3-methyl-adenine; Peak 2: N, 9-dimethyl-adenine; Peak 3: N, N, 9-trimethyl adenine; Peak 4: N, N, 3-trimethyl-adenine; Peak 5: dimethyl adenine; Peak 6,7,8,9: trimethyl adenine. Note: the break in the intensity scale to show the height of peak 3 in the 2.5 nmol chromatogram.

Figure 2: Chromatograms of thymine (0.26 nmol) and uracil (0.26 nmol) thermochemolysis in TMAH at different temperatures.

Figure 3: Chromatograms of cytosine (15 nmol) thermochemolysis in TMAH thermochemolysis at different temperatures (1: 4-amino-2-methoxy-pyrimidine; 2: dimethyl cytosine; 3: trimethyl-cytosine; 4: N, N, N'-trimethyl-cytosine; 5: trimethyl-cytosine; 6: siloxane).

Figure 4: Chromatograms of guanine (11 nmol) thermochemolysis in TMAH at 600 °C, 500 °C, and 400 °C. (1: caffeine; 2: dimethyl-guanine; 3 and 5: trimethylguanine; 4, 6, 7, 8: tetramethylguanine; 9: trimethylguanine).

Figure 5: Chromatograms of xanthine (0.22 nmol) and hypoxanthine (1.2 nmol) thermochemolysis in TMAH at different temperatures.

Figure 6: The TMAH thermochemolysis schemes of seven nucleobases.

Figure 7: Chromatogram of seven nucleobases mixtures dissolved in TMAH thermochemolysis at 600°C.Peak 1: 1, 3-dimethyl-uracil ; Peak 2 : 1, 3-dimethyl-thymine; Peak 3 : N, N, N'-trimethyl-cytosine; Peak 4 : N, N, 9-trimethyl-adenine; Peak 5 : 1, 3, 7-trimethyl-xanthine (caffeine); Peak 6 : 1,7-dimethyl-hypoxanthine; Peak 7: methyl ester hexadecanoic acid; Peak 8 : trimethyl-guanine; Peak 9 : trimethyl-adenine Peak X: Siloxane.

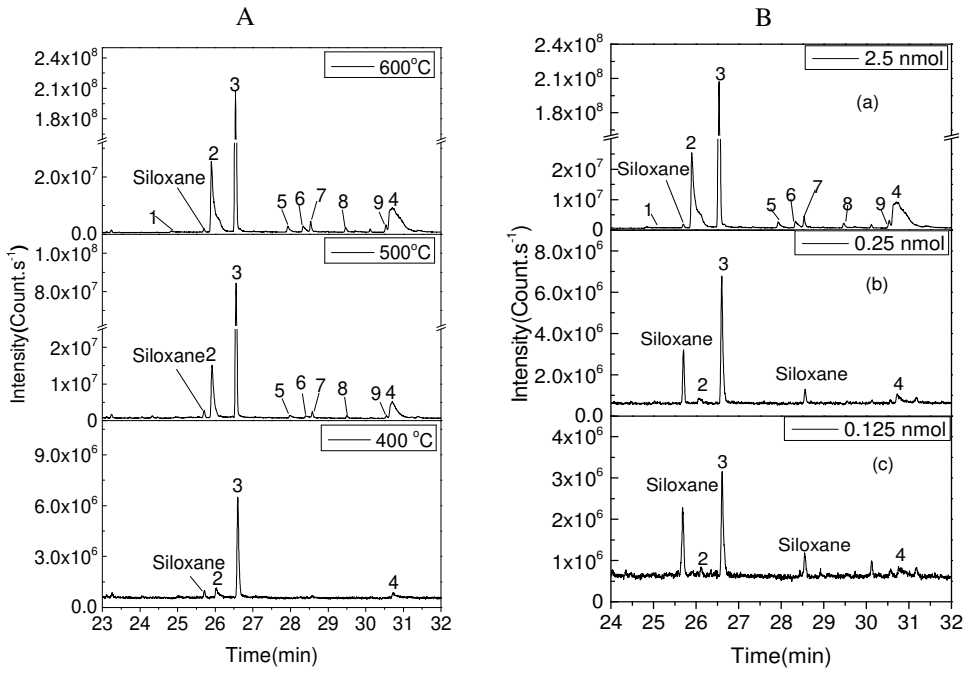


Figure 1: Chromatograms of adenine thermochemolysis in TMAH performed at different temperature (A: 600°C, 500°C, and 400°C.) and different amount (B: (a) 0.125 nmol (b) 0.25 nmol (c) 2.5 nmol). Chromatogram A: Peak 1: 3-methyl-3H-purin-6-amine; Peak 2: N, 9-dimethyl-9H-purin-6-amine; Peak 3: N, N, 9-trimethyl-9H-purin-6-amine; Peak 4: N, N, 3-trimethyl-3H-purin-6-amine; Peak 5: dimethyl adenine; Peak 6,7,8,9: trimethyl adenine. Note the break in the intensity scale to show the height of peak 3;
Chromatogram B: Peak 1: 3-methyl-adenine; Peak 2: N, 9-dimethyl-adenine; Peak 3: N, N, 9-trimethyl adenine; Peak 4: N, N, 3-trimethyl-adenine; Peak 5: dimethyl adenine; Peak 6,7,8,9: trimethyl adenine. Note the break in the intensity scale to show the height of peak 3 in the 2.5 nmol chromatogram.

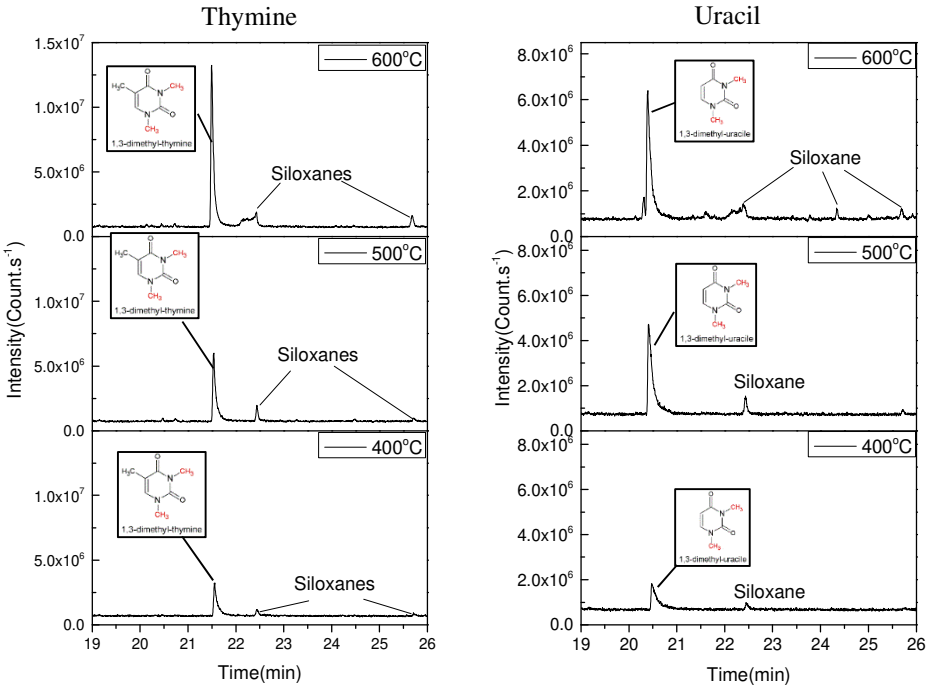


Figure 2: Chromatograms of thymine (0.26 nmol) and uracil (0.26 nmol) thermochemolysis in TMAH at different temperatures.

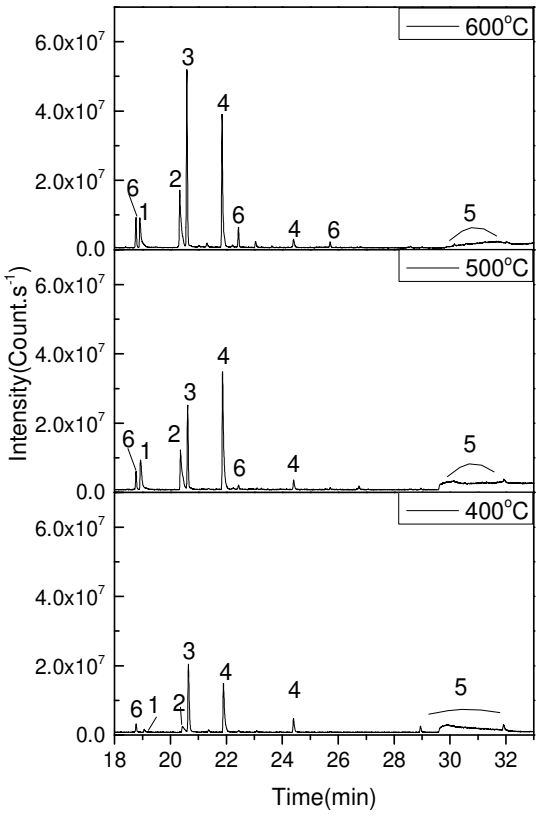


Figure 3: Chromatograms of cytosine (15 nmol) thermochemolysis in TMAH thermochemolysis at different temperatures (1: 4-amino-2-methoxy-pyrimidine; 2: dimethyl cytosine; 3: trimethyl-cytosine; 4: N, N, N'-trimethyl-cytosine; 5: trimethyl-cytosine; 6: siloxane).

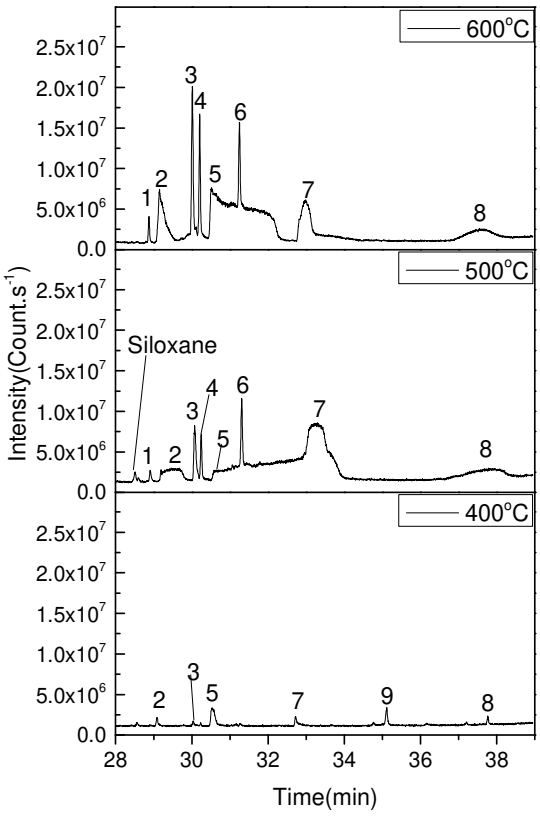


Figure 4: Chromatograms of guanine (11 nmol) thermochemolysis in TMAH at 600 °C, 500 °C, and 400 °C. (1: caffeine; 2: dimethyl-guanine; 3 and 5: trimethylguanine; 4, 6, 7, 8: tetramethylguanine; 9: trimethylguanine).

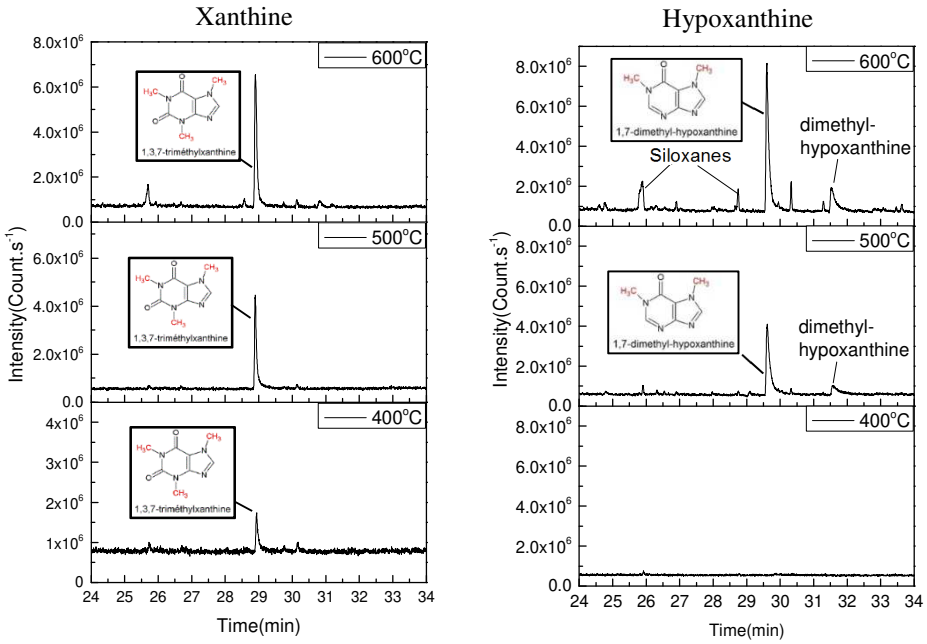


Figure 5: Chromatograms of xanthine (0.22 nmol) and hypoxanthine (1.2 nmol) thermochemolysis in TMAH at different temperatures.

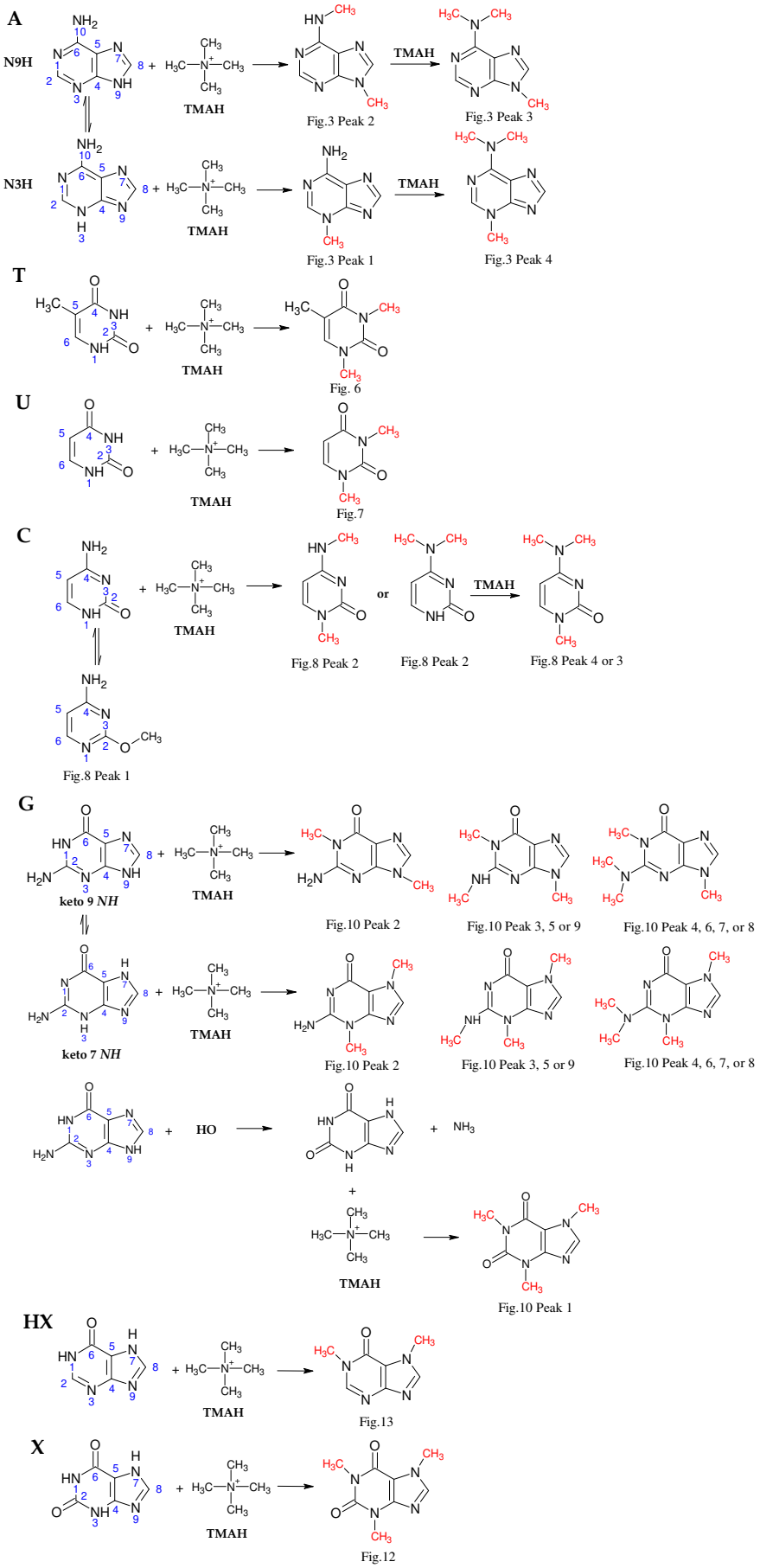


Figure 6: The TMAH thermochemolysis schemes of seven nucleobases.

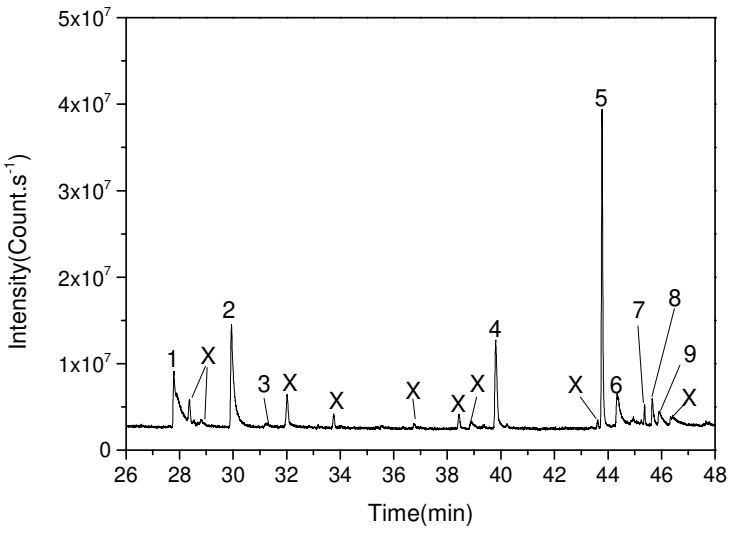
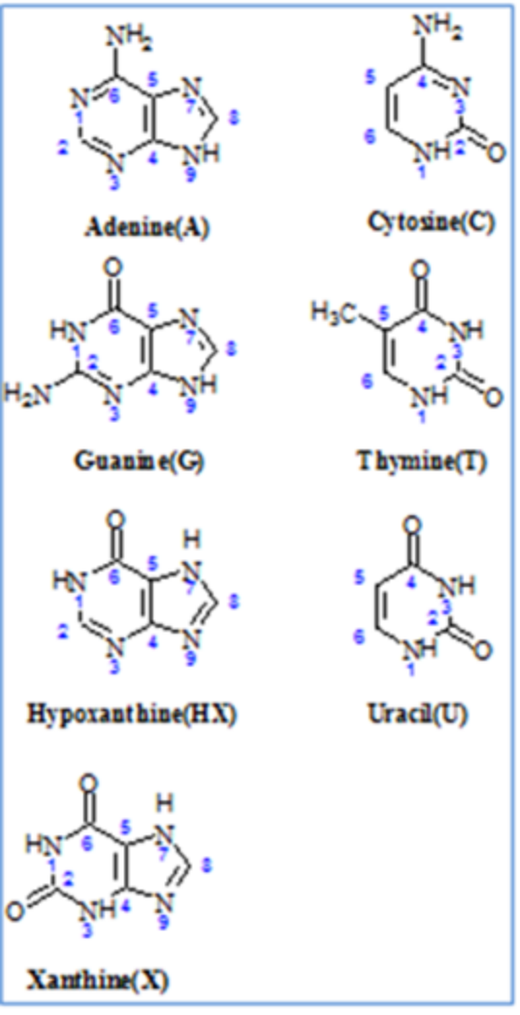


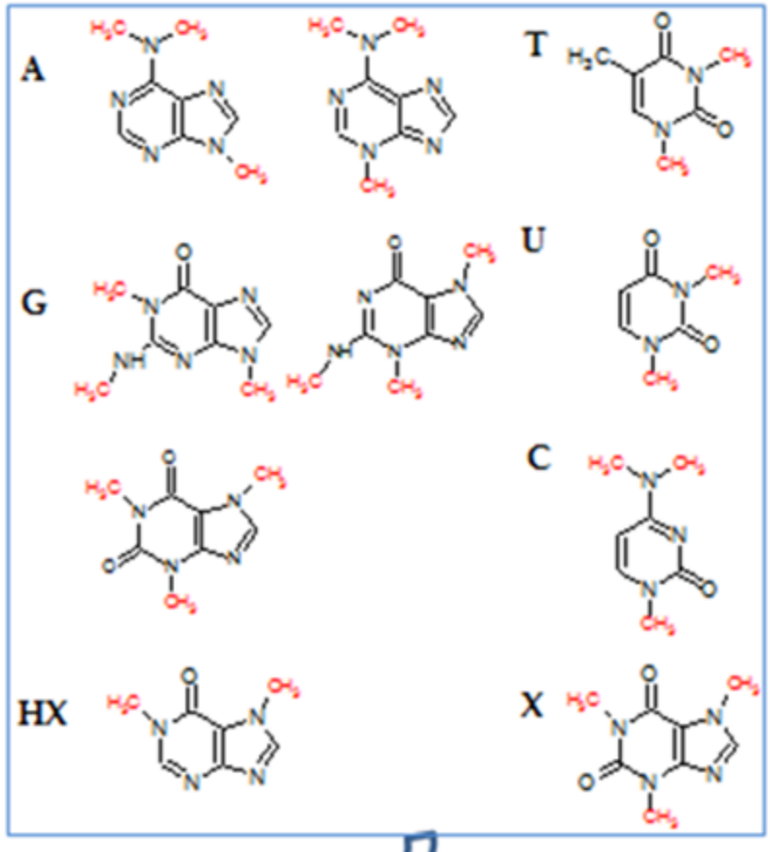
Figure 7: Chromatogram of seven nucleobases mixtures dissolved in TMAH thermochemolysis at 600°C. Peak 1: 1, 3-dimethyl-uracil ; Peak 2 : 1, 3-dimethyl-thymine; Peak 3 : N, N, N'-trimethyl-cytosine; Peak 4 : N, N, 9-trimethyl-adenine; Peak 5 : 1, 3, 7-trimethyl-xanthine (caffeine); Peak 6 : 1,7-dimethyl-hypoxanthine; Peak 7: methyl ester hexadecanoic acid; Peak 8 : trimethyl-guanine; Peak 9 : trimethyl-adenine Peak X: Siloxane.

Nucleobases (NB)

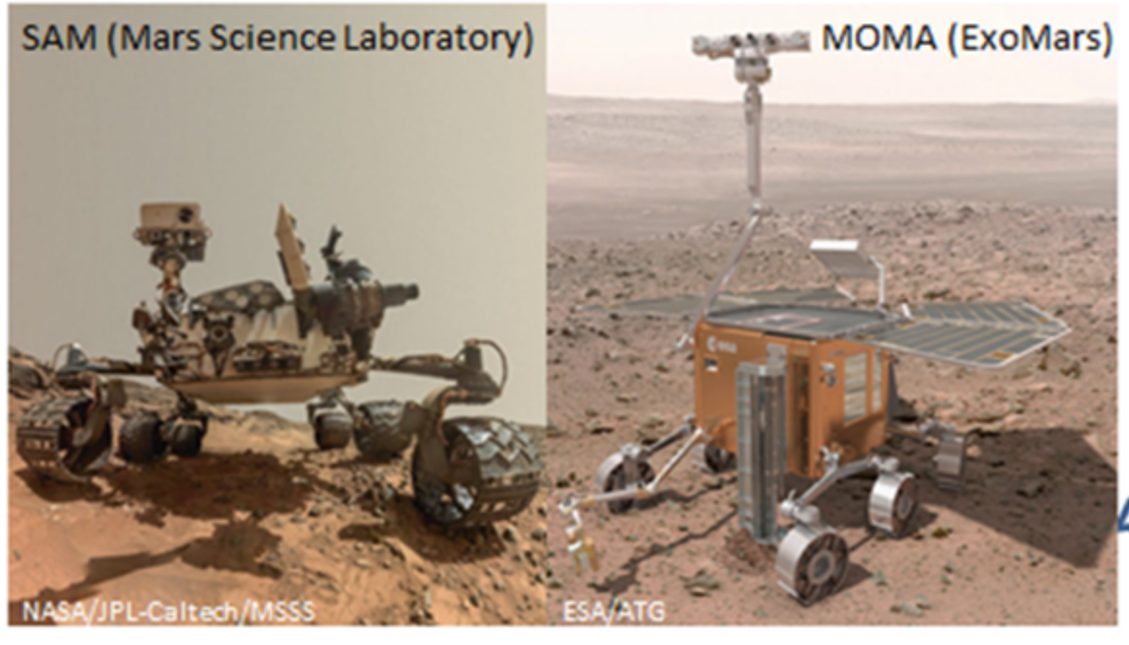


TMAH Thermochemolysis

Methylated NB
LOD
LOQ



Identification



Use lab experiment on Mars