A Laboratory Infrared Model of Astrophysical Pyrimidines

Caroline Antunes Rosa[1*], Alexandre Bergantini[2], Enio Frota da Silveira[3], Marcelo Emilio[4], Laerte Andrade[5], Eduardo Janot Pacheco[6], Nigel J. Mason[7,8] and Claudia Lage[1]

1 – Instituto de Biofísica Carlos Chagas Filho, Universidade Federal do Rio de Janeiro, Av. Carlos Chagas Filho, 373 - Cidade Universitária da Universidade Federal do Rio de Janeiro, CEP - 21941-170, Rio de Janeiro, RJ - Brazil.

2 –Centro Federal de Educação Tecnológica Celso Suckow da Fonseca, Av. Maracanã 229, CEP - 20271-110, Rio de Janeiro, RJ, Brazil;

3 – Departamento de Física, Pontifícia Universidade Católica do Rio de Janeiro, Rua Marquês de São Vicente 225, CEP - 22451-900, Rio de Janeiro, RJ – Brazil;

4 – Departamento de Geociências, Universidade Estadual de Ponta Grossa, Av. General Carlos Cavalcanti, 4748, CEP - 84030-900, Ponta Grossa, PR – Brazil;

5 - Laboratório Nacional de Astrofísica, R. dos Estados Unidos, 154, CEP - 37504-364, Itajubá, MG, Brazil;

6 - Instituto de Astronomia, Geofísica e Ciências Atmosféricas - IAG, Universidade de São Paulo, R. do Matão, 1226, CEP - 05508-090, São Paulo, SP, Brazil;

7 - Centre for Astrophysics and Planetary Science, School of Physics and Astronomy, University of Kent, Canterbury CT2 7NH, United Kingdom

8 - Atomic and Molecular Physics Laboratory, Atomki Institute for Nuclear Research, Debrecen H-4026, Hungary

*Corresponding author: cantunesbio@gmail.com

Abstract:

Nucleobases are essential molecules for life, forming integral parts of DNA and RNA in all terrestrial life forms. Despite evidence of their abiotic synthesis in meteorites and laboratory simulations of interstellar medium (ISM) conditions, nucleobases have not been detected in the ISM. This study investigates the infrared spectra of uracil, cytosine, and thymine—pyrimidine nucleobases—embedded in an ice mixture simulating common volatiles found in protostellar disks. Our objective was to explore the feasibility of identifying unique infrared bands of pyrimidines in the ISM, despite significant overlapping absorption features from simpler, more abundant interstellar species such as H_2O, CO, CH_3OH, and NH_3. Laboratory results revealed that although two common bands (1240 and 760 cm^{-1} in uracil; 1236 and 763 cm^{-1} in cytosine; 1249 and 760 cm^{-1} in thymine) were identified, the detection of these bands in space is challenged by overlapping absorption features. Recent observations with the JWST have shown that interstellar organic species

exhibit infrared signals within similar ranges, making it impossible to distinguish pyrimidine bands from these organics. Thus, detecting pyrimidines with current telescopes is infeasible, not due to sensitivity limitations or the need for more powerful instrumentation, but because of the intrinsic overlap in spectral features. This study complements previous research on purines by examining pyrimidines and including the impact of common ISM volatiles in the ice composition. The results highlight the significant challenges in detecting complex molecules in the ISM, underscoring the importance of understanding the spectral complexities and interactions to interpret astronomical observations accurately.

Key-words: Astrochemistry, Infrared: general, ISM: molecules, methods: laboratory: solid state.

1. Introduction

Nucleobases are essential molecules for life as we know it. They are key molecules in evolutionary, genetic, and hereditary processes in all life forms on Earth, being the aromatic nitrogen heterocyclic (N-heterocycles) chemical species present in DNA and RNA. There are two groups of nucleobases: i) purines, which are formed by the fusion of two rings - a pyrimidine ring and an imidazole ring, and ii) pyrimidines, which are made of a pyrimidine ring only. For life on Earth, the purines present in nucleic acids are adenine ($C_5H_5N_5$) and guanine ($C_5H_5N_5O$), the pyrimidines present in nucleic acids of living organisms are cytosine ($C_4H_5N_3O$ - DNA and RNA), thymine ($C_5H_6N_2O_2$ - DNA only), and uracil ($C_4H_4N_2O_2$ - RNA only). However, it is worth mentioning that there are other purines and pyrimidines in nature, besides the ones present in nucleic acids, which may or may not participate in biological functions. Nucleobases from each group are differentiated by the functional groups hanging off the ring (Figure 1).

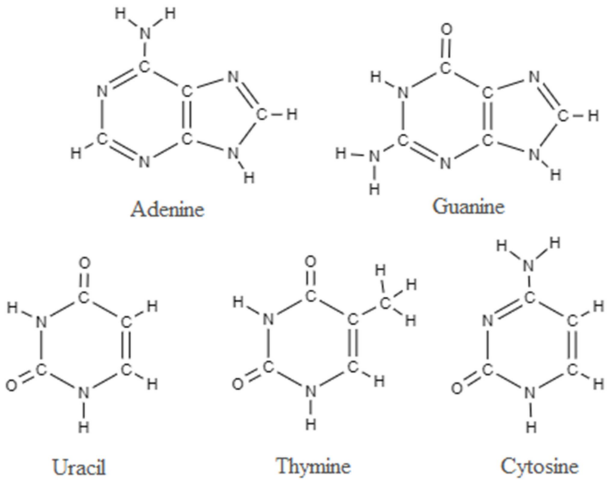

Figure 1 – Nucleobases found in nucleic acids on Earth.

How complex molecules such as nucleobases were formed and incorporated into the first living organisms on Earth remains a matter of debate (Saladino *et al.* 2012). Early hypothesis suggested that the building blocks of life were formed via chemical reactions between simple molecules present in the early Earth's atmosphere and oceans (Kobayashi 2019). Indeed, a series of experiments have demonstrated that nucleobases can be formed in conditions analogue to those found in primitive Earth (Kitadai & Maruyama 2018). It is known that pyrimidines can be abiotically synthesized. For example, uracil and cytosine can be synthesized from cyanoacetylene (C_3HN) (Fox & Harada 1961, Ferris *et al.* 1968) – a chemical species which has been detected in the interstellar medium (ISM) (Turner 1971). According to Choughuley *et al.* (1977) and Schwartz & Chittenden (1977), thymine can be formed from reactions between uracil and formaldehyde (CH_2O) (a species found in the ISM as well). Among the several experiments reporting the abiotic synthesis of nucleobases, the ones that seem to yield the most promising results use formamide (NH_2CHO, detected in the ISM by Rubin *et al.* 1971) as the main precursor. Indeed, it has been shown that it is possible to synthetize all nucleobases found both in DNA and RNA from the polymerization of formamide under high temperature in early-Earth-like scenarios (Saladino *et al.* 2001. Saladino *et al.* 2003, Saladino *et al.* 2004, Saladino *et al.* 2006, Saladino *et al*. 2012).

Besides experimental evidence for routes of terrestrial abiotic synthesis of nucleobases, an alternative hypothesis suggests that the ingredients for early-life forms were produced on dust grain mantles found in cold and dense gas envelopes in the protostellar stage and delivered to Earth by meteorites and comets that collided with the planet, mainly during the "late heavy bombardment" epoch, about 4 billion years ago (Chyba & Sagan 1992, Sandford *et al*. 2020). This hypothesis is supported by the observation of an extensive number of biomolecules, such as amino acids, nucleobases, sugars, and aliphatic molecules (bio-membrane precursors) found in meteorites (Hayatsu *et al.* 1975, Stoks & Schwartz 1979, Deamer & Pashley 1989, Shock & Schulte 1990, Cooper *et al.* 2001, Sephton 2002, Martins *et al.* 2008, Callahan *et al.* 2011, Burton *et al.* 2012, Cooper & Rios 2016, Furukawa *et al.* 2019). Calculations made by Chyba and Sagan (1992) suggested that the amount of organic material produced on Earth (i.e., endogenous material) and delivered from outside Earth (i.e., exogenous material) may have contributed equally to the origin of life in the planet.

In a recent study, Oba *et al*. (2022), using new analytical techniques, reported the detection of all terrestrial life nucleobases – adenine, guanine, thymine, cytosine and uracil – in the fragments of the meteorites *Murray*, *Tagish Lake*, and *Murchison*. Furthermore, laboratory experiments simulating ISM conditions have demonstrated that the synthesis of nucleobases on the surface of ices under electromagnetic and/or ion irradiation is plausible (Saladino *et al.* 2005, Nuevo & Sandford 2014,

Materese *et al.* 2018, Ruf *et al.* 2019, Oba *et al.* 2019). For example, from the UV (Lyman-α) irradiation of purine ($C_5H_4N_4$) and pyrimidine ($C_4H_4N_2$) with different mixtures of commonly found ISM species (including H_2O, CH_3OH, NH_3 and CH_4), the synthesis of adenine, guanine, uracil, cytosine, and thymine was obtained (Nuevo *et al.* 2009, Nuevo *et al.* 2012, Materese *et al.* 2013, Materese *et al.* 2017). Oba *et al*. (2019) also performed an experiment with an interstellar ice analogue containing a mixture of H_2O:CO:CH_3OH:NH_3 (5:2:2:2) irradiated with UV (Lyman-α) at 10K in ultra-high vacuum. They were able to produce cytosine, uracil, thymine, adenine, hypoxanthine, xanthine, and nitrogen heterocycles. Ruf *et al*. (2019) also demonstrated the synthesis of cytosine from an ice mixture containing H_2O:CH_3OH:NH_3 (2:1:1) irradiated with Lyman-α in a simulated astrophysical environment. Both experiments demonstrated that nucleobases could be formed from noncyclic molecules in realistic ISM conditions. Indeed, the simple molecular species used in these experiments are abundant in protostellar disks (Öberg *et al.* 2011), making these regions favorable for the synthesis of biomolecules (Bergantini *et al*. 2017).

From an astronomical point of view, nucleobases are labeled as complex organic molecules (COMs) since they contain more than six atoms (Herbst & Van Dishoeck 2009) and the nitrogen incorporated to the ring increases the complexity of the astrophysical synthesis of these molecules (Parker *et al*. 2015). According to Sandford et al. (2020) and Dullemond et al. (2020) the region known as the midplane of the middle of the protostellar disk can reach temperatures lower than 20 K and COMs are believed to be produced on the dust grain mantles in these disks (Öberg 2016, Danger *et al.* 2011, Maity *et al.* 2015). Due to temperature increase and/or shocks associated with accretion and outflow activities (among other phenomena) around protostellar disks, the COMs produced in the solid phase can be liberated from the dust grains (Zhang *et al*. 2023), thus going to gas phase. Thanks to the Atacama Large Millimeter/submillimeter Array and other large radio facilities (e.g., IRAM 30 m) COMs and many organic molecules are being observed in several of these objects (e.g. Baek *et al*. 2022; Guélin & Cernicharo 2022; Coutens et al. 2022). However, no purine or pyrimidine has been identified, in either phase, towards objects of the interstellar medium (ISM) so far, including attempts made by Simon & Simon (1973), Charnley *et al.* (2005), and Brünken *et al.* (2006). Identifying complex molecules in the ISM is a major challenge, since the larger the molecule, the more complex is its spectral signature. It is especially difficult to detect species which are in low concentration, since many of the distinctive spectral features (i.e., their "*fingerprint*") may be concealed by bands of higher-concentration species (Janot-Pacheco *et al.* 2018).

However, the detection of pyrimidines or other complex molecules in the solid phase toward ISM objects is highly challenging if not impossible. This is due to the significant overlap of their infrared bands with those of many other molecules that have similar functional groups. Specifically,

the C–H and C–O absorption features of various possible molecules contribute at 7.8 micrometers, and ice features are present in the range between 6.8 and 8.5 micrometers. Additionally, molecules such as CH_4, SO_2, HCOOH, CH_3CHO, and CH_3CH_2OH are expected to absorb in the targeted wavelength range. The absorption profiles of H_2O and CH_3OH around 6.8 micrometers from ice mixtures, along with the shared band at 7.2 micrometers between HCOO− and HCOOH, further complicate the detection. Therefore, even with more powerful instrumentation, the overlapping absorption features make it impossible to distinctly determine the characteristic infrared bands of pyrimidines or other complex molecules.

Therefore, in this study, the infrared spectra of a mixture of molecules, at cryogenic temperatures, containing H_2O:CO:CH_3OH:NH_3 deposited on top of uracil, cytosine, and thymine, were collected, as these volatiles are potential precursors of interstellar nucleobases (Nuevo & Sandford 2014, Materese *et al*. 2018, Oba *et al*. 2019, Ruf *et al*. 2019). Data was collected both at room (300 K) and cryogenic temperatures (15 K). Room-temperature spectra are compatible with the later stages and inner portions of the protostellar disk, where the temperatures can easily reach 300 K and more (Boss 1998). Conversely, the outer regions, as well as the midplane of the middle of the protostellar disk, are found in temperatures ranging from 10 K to 100 K (Sandford *et al*. 2020). Our main objective is to explore the feasibility of identifying any unique infrared bands of pyrimidines related to nucleobases in the ISM, despite the presence of overlapping absorption features from common ISM species (i.e., H2O:CO:CH3OH:NH3). This study complements a previous study on purines which proposed their interstellar infrared spectral signature (Rosa *et al*. 2023). Although distinct identification may be impossible, understanding the spectral complexities and interactions provides valuable insights for interpreting astronomical observations and the chemical composition of the ISM.

2. Methodology

2.1 Experimental apparatus

The experiments were performed at the *Van de Graaff Laboratory* of the *Pontifical Catholic University of Rio de Janeiro* (PUC-Rio). The experimental apparatus consists of a high-vacuum chamber, a Fourier Transform Infrared Spectrometer (FTIR – Jasco FT-IR-4200), a sample holder which can be cooled down to 15 K using a closed-cycle helium cryostat (CCS-UHV/204 Janis Research Company cold head and Sumitomo HC-4E compressor). The base pressure of the chamber during the experiments was of the order of 10^{-8} mbar. The nucleobases cytosine, thymine, and uracil (≥99% purity) in powder form were purchased from Sigma-Aldrich. The powdered pyrimidines were deposited onto ZnSe substrates as a thin film produced by sublimation inside a separated high-vacuum chamber (Edwards E306 model). The thickness of each pyrimidine sample

was measured *in-situ* by a quartz crystal microbalance. After producing the pyrimidine thin film, the sample was then transferred to the main experimental high-vacuum chamber, which was then pumped down for 12 to 24 hours, after which the substrate was cooled down to approximately 15 K. Then, a H_2O:CO:CH_3OH:NH_3 gas mixture (prepared in a separated chamber) was deposited on top of the pyrimidine film. Since the pyrimidines of interest are solid at room temperature, we have used the methodology described by Saïagh et al. (2014) to prepare our samples, which consisted of a condensed layer of gases on top of a previously evaporated layer of nucleobase. The final result is a sample made of two main layers: the bottom layer is the nucleobase, and the top layer is the gas mixture. The gas mixture was prepared as follows: the gases were added one at a time in a separated gas-mixing chamber (GMC). Firstly, 16 mbar of a 2:1 solution of H_2O and CH_3OH was added. Then, 4 mbar of CO and 4 mbar of NH_3 were added, respectively, resulting in the final total pressure of 24 mbar. The GMC was otherwise kept at 10^{-6} mbar to avoid contamination. The 2:1 solution of H_2O and CH_3OH was determined empirically to produce the final desired ratio of 5:2 in ice form.

Once the mixture was prepared, a leak valve was used to let the gas mixture into the main experimental chamber. During the mixture deposition, the main chamber pressure was kept at (5±3) x10^{-7} mbar. Each mixture deposition lasted for approximately 10 minutes. The chemical composition of the sample was analyzed *online* and *in-situ* immediately after deposition using FTIR spectroscopy.

Infrared spectra (120 scans per spectrum at 4 cm^{-1} of resolution) in the 4500 to 650 cm^{-1} range (~2 to 15 μm) were collected in each of the following steps: i) when the substrate containing the evaporated pyrimidine nucleobase was loaded in the experimental chamber at room pressure and temperature; ii) when the pressure of the chamber was evacuated to ~10^{-8} mbar (high vacuum) at room temperature; iii) in high vacuum and at 15 K, before and after the deposition of the gas mixture.

2.2 Column density calculation

The column densities (*N*) and the ratios between the ices and the pyrimidine nucleobases were calculated using a modified Lambert-Beer equation (eq. 1) (Bergantini *et al.* 2018):

$$N = ln10 \frac{\int_{\nu 1}^{\nu 2} | \quad | d\nu}{Avalue} (1)$$

where $\int_{\nu 1}^{\nu 2} | \quad | d\nu$ is the integrated peak area the IR band of interest, and the *A value* is the infrared absorption coefficient (also called band strength). The *A values* of the water, methanol, ammonia, and carbon monoxide bands used in the calculations were extracted from Bouilloud *et al.* (2015). The pyrimidines column densities were measured as follows: The *A value* used for the uracil sample

was the one from the β band, which is defined as the integrated band strength of the bands found between 3400 to 1890 cm^{-1} of uracil (Saïagh *et al.* 2015). Similarly for cytosine, the *A value* used was the α band, which is defined as integrated band strength between 3500 to 2000 cm^{-1} of the cytosine spectrum (Vignoli Muniz *et al.* 2022). To characterize thymine the 940 cm^{-1} band was used (Mejía *et al.* 2023). These considerations resulted in the following ratio calculated from the column densities: 10:5:0.2:19:5 (H_2O:CO:NH_3:CH_3OH:Uracil), 10:3:2:8:0.7 (H_2O:CO:NH_3:CH_3OH:Cytosine), and 10:0.8:0.1:3:0.3 (H_2O:CO:NH_3:CH_3OH:Thymine).

3. Results

3.1 The infrared signatures of the Pyrimidines Uracil, Cytosine, and Thymine

The crystalline IR spectra of the neat uracil, cytosine, and thymine obtained at 300 K and 15 K are presented in Figure 2 and 3, respectively. The crystallinity of the samples was verified by comparison with data from the literature (Szczesniak *et al.* 1985). The assignments of the observed infrared bands are shown in Table 1. The observed bands were characterized according to the literature (Susi & Ard 1971, Radchenko *et al*. 1984, Mathlouthi *et al.* 1986, Szczesniak *et al*. 1988, Nowak *et al*. 1989, Leś *et al.* 1992, Kwiatkowski & Leszczyński 1996, Colarusso *et al.* 1997, Szczepaniak *et al.* 2000, Singh 2008, Fornaro *et al.* 2014, Saïagh *et al.* 2015).

The spectral region between ~3360 cm^{-1} and ~2650 cm^{-1} is characterized by C-H and N-H out of the ring stretching vibration modes for all three of the pyrimidines analyzed, in addition to the stretching modes of the functional groups attached to the ring, such as the cytosine's amino group and the thymine's methyl group. The region between ~1780 cm^{-1} and 650 cm^{-1} is characterized by νC-C, νC-N, νC=C, νC=N and νC=O modes as well as by the bending modes out of the rings' functional groups. In this region, the vibration modes of the whole molecule, such as the βRing, the νRing and the ring-breathing modes, are also observed. There are similarities between the spectra of uracil, cytosine, and thymine since all these species belong to the pyrimidines group. Some bands are found near the same position in all the three pyrimidines' spectra. There are six such bands with similar wavenumbers: uracil's 1674, 1462, 1240, 1012, 806, and 760 cm^{-1} bands; cytosine's 1661, 1467, 1236, 1013, 795, and 763 cm^{-1} bands; and thymine's 1673, 1459, 1249, 1027, 816, and 760 cm^{-1} bands. For all three pyrimidines, the ~760 cm^{-1} band is attributed to the ring-breathing mode, a characteristic vibration for the whole ring. For this reason, this band could be used as a signature of the three pyrimidines present in DNA and RNA.

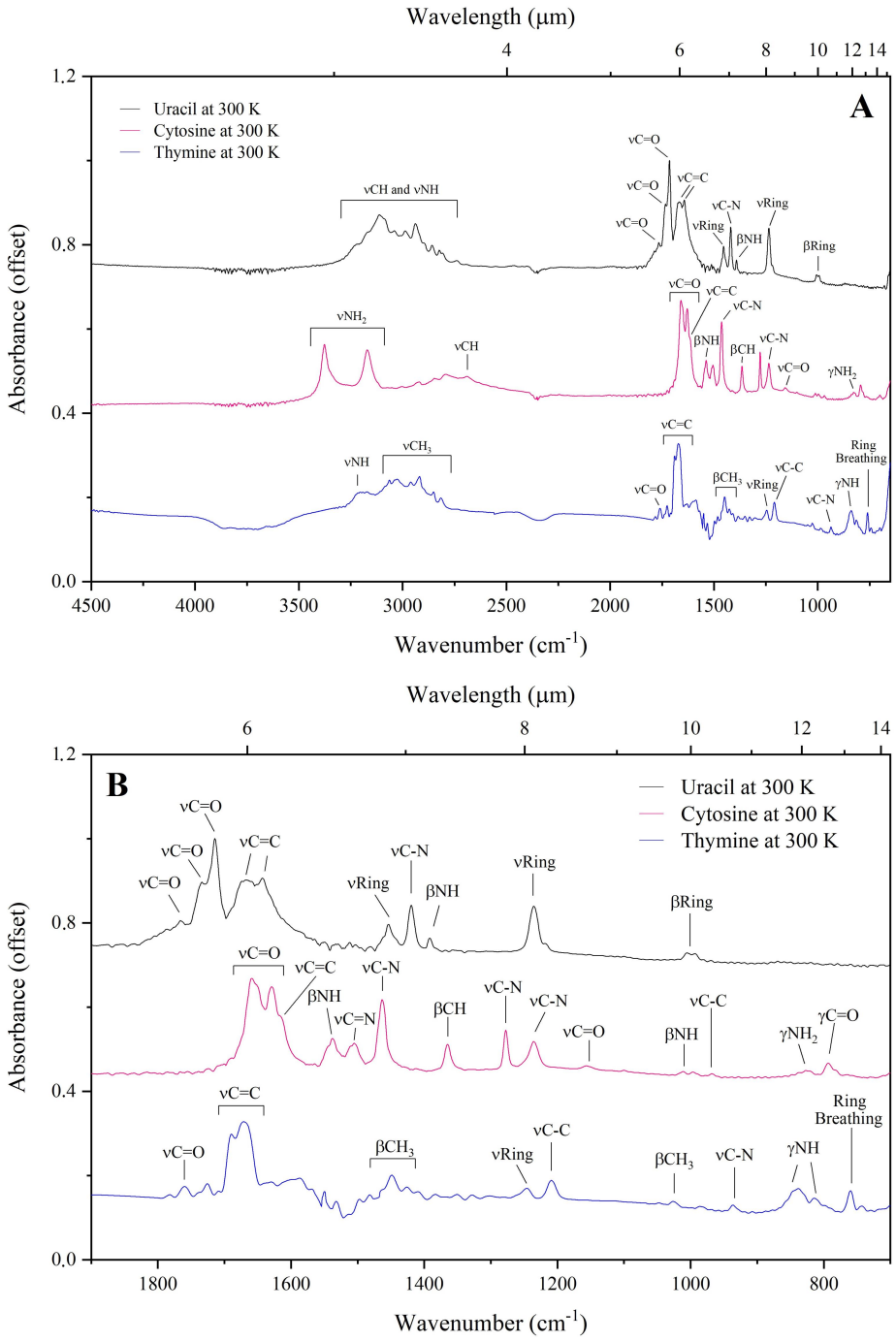

Figure 2 – (A) Mid-IR spectra of uracil (black line), cytosine (pink line), and thymine (blue line) at 300 K, in high-vacuum. The main bands are identified by their assignments. (B) The fingerprint region (1900-700 cm^{-1}) of pyrimidines uracil (black line) cytosine (pink line), and thymine (blue line). The top scale is in wavelength (μm), while the bottom scale is in wavenumber (cm^{-1}) a common unit used in laboratory IR spectroscopy. Abbreviations: ν – stretching; β – in plane bending; γ – out of plane bending; δ – scissoring.

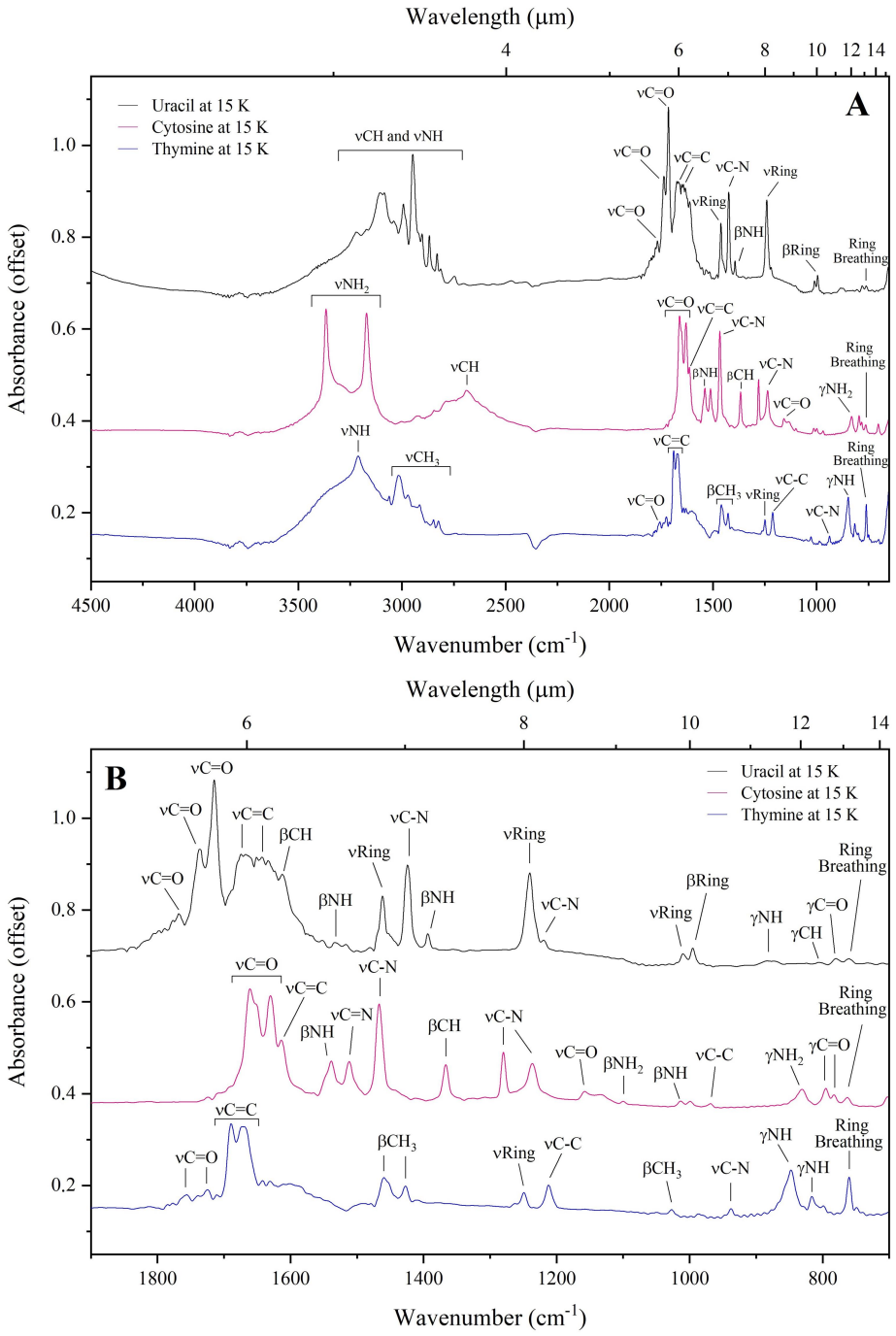

Figure 3 – (A) – Mid-IR spectra of uracil (black line), cytosine (pink line), and thymine (blue line) at 15 K, in high-vacuum. The main bands are identified by their assignments. (B) – The fingerprint region (1900-700 cm^{-1}) of pyrimidines uracil (black line) cytosine (pink line), and thymine (blue line). The top scale is in wavelength (μm), while the bottom scale is in wavenumber (cm^{-1}) a common unit used in laboratory IR spectroscopy. Abbreviations: ν – stretching; β – in plane bending; γ – out of plane bending; δ – scissoring.

Table 1 – Assignment of the IR bands observed in uracil, cytosine, and thymine's spectra at 300 and 15 K.

Uracil				**Cytosine**				**Thymine**			
300 K ν (cm^{-1})	15 K ν (cm^{-1})	Assignment	Ref.	300 K ν (cm^{-1})	15 K ν (cm^{-1})	Assignment	Ref.	300 K ν (cm^{-1})	15 K ν (cm^{-1})	Assignment	Ref.
3112	3107			3376	3366	ν_{asym} NH_2	6	3205	3211	νN-H	11
3087	3085			3170	3171	ν_{sym} NH_2	6	3063	3062	νC-H	11; 12
2988	2993			2692	2688	νC-H	6	3026	3016	$\nu_{asym}CH_3$	11
2938	2948	νC-H; νN-H	1; 2; 3;4	1659	1661	νC=O	6	2963	2972	νCH_3	1; 2; 5
2898	2903			1629	1630		6	2919	2918	νCH_3	11; 12
2858	2868			1616	1614	νC=C	6	2851	2848	νCH_3	1
2823	2831			1538	1539	βN-H	6; 9; 10	1760	1756	νC=O	1; 2; 5; 12
1768	1768	νC=O	1; 2; 5	1505	1512	νC=N and νC-N	6	1726	1725	νC=O	1; 2; 5; 12
1735	1736	νC=O	1; 2; 5	1463	1467	νC-N	5; 6; 7; 9;10	1689	1689	νC=C	1; 2
1714	1715	νC=O	1; 2; 3; 4	1365	1366	βC-H	5; 6; 8; 9	1671	1673	νC=C	1; 2; 5; 12
1667	1674	νC=C	3; 4	1278	1280	νC-N	6	1449	1459	βCH_3	1; 2; 5; 11; 12
1643	1643	νC=C	1; 2; 5	1235	1236	νC-N	5; 6; 7; 9; 10	1426	1427	βCH_3	1; 2; 5; 11; 12
-	1616	βC-H	2	1157	1158	νC=O	6	1246	1249	νRing	11
1533	1535	βN-H	1; 2; 3; 4	1100	1099	$\beta_{rock}NH_2$	5; 6; 8; 9; 10	1209	1212	βC-H, νC-N, νC-C, $\nu C\text{-}CH_3$	1; 2; 5; 12
1454	1462	νRing, νC-N and βN-H	1; 2; 3; 4; 5	1011	1013	βN-H and βC-H	6	1026	1027	$\beta_{rock}CH_3$ and νRing	1; 2; 5; 11; 12
1419	1424	νC-N and βN-H	1; 2; 3; 4; 5	998	999	βC-H	6	936	938	νC-N	1; 2
1392	1394	βN-H; βC-H and νC-N	1; 2; 3; 4; 5	969	969	νC-C	6; 9; 10	838	847	γN-H	11
1235	1240	νRing	3; 4	827	831	γNH_2	6; 8	814	816	γN-H	11
1220	1221	νC-N, βN-H and βC-H	1; 2; 3; 5	793	795	γC=O	5	760	760	Ring breathing and γC=O	1; 2; 5; 11; 12
1005	1012	νRing	3	783	783	γC=O	5; 7; 10	743	749	γC=O	1; 2; 5; 11; 12
-	995	βRing	1; 2; 3; 5	767	763	Ring breathing	5; 9; 10				

-	881	γNH	3	700	702	βRing	5; 6; 7; 10
-	806	γC-H and γC=O	1; 2; 5				
-	780	γC=O	1; 2; 5				
-	760	Ring breathing	1; 2; 5				

Ref. 1. 1 - Colarusso *et al.* 1997, 2 - Leś *et al*. 1992, 3 - Susi & Ard 1971, 4 - Saïagh *et al*. 2015, 5 - Fornaro *et al*. 2014, 6 - Mathlouthi *et al.* 1986, 7 - Kwiatkowski & Leszczyński 1996, 8 - Radchenko *et al.* 1984, 9 - Szczesniak *et al.* 1988, 10 - Nowak *et al.* 1989, 11 - Singh 2008, 12 - Szczepaniak *et al.* 2000.

To understand how the infrared bands of uracil, cytosine, and/or thymine contribute to the observed spectra in cold clouds and protostellar disks, it is necessary to take into account the chemical environment within which the molecules are found. According to recent experimental and theoretical studies (e.g., Bergantini *et al*. 2018; Bergantini *et al.* 2017; Ioppolo *et al*. 2021; Vasyunin *et al*. 2017), the interstellar synthesis of complex organic molecules, which potentially includes pyrimidines, is likely to occur in gas and dust-enriched cold regions, via non-thermal, energetic, and non-energetic mechanisms. These conditions are mainly found in the outer regions, as well as the midplane of the middle of the protostellar disks. Therefore, the infrared spectra of uracil, cytosine, and thymine was collected embedded in a realistic interstellar ice analogue, which contained some of the most abundant species found in polar ices found within protostellar disks (McClure *et al.* 2023, Öberg *et al.* 2011, Boogert *et al.* 2015). Our aim was to determine if, and how, the spectrum of uracil, cytosine, and thymine would change when common molecular species, such as H_2O:CO:CH_3OH:NH_3, are condensed onto it, thus possibly mimicking the ice mantles found in the coldest regions of protostellar disks. The infrared spectra of these three pyrimidines embedded in a H_2O:CO:CH_3OH:NH_3 mixture at 15 K is shown in Figure 4. The IR modes of uracil, cytosine and thymine are indicated by red arrows and the modes of H_2O, CO, NH_3, and CH_3OH by black arrows. The icy samples whose spectra are shown in Figure 4 present a layered structure. Even though the reaction mechanisms that take place in ice mixtures versus layered ices are mostly likely different from each other, for the purposes of this investigation (i.e., the characterization of infrared bands of pyrimidines in space), the spectra of layered ices can be considered a valid effort to be used in astrophysical observations.

Figure 4 shows that most infrared bands of uracil, cytosine, and thymine are overlapped by the volatiles bands. As a result, even the weak lower energy transitions, which are less affected by overlapping, are not sufficiently distinguishable to be detected in future observations. This is because the overlapping absorption features from common ISM species, such as H_2O, CO, CH_3OH, and NH_3, create a complex background that masks the unique spectral signatures of the pyrimidines. The region between 3500 and 2700 cm^{-1} is completely dominated by the strong water, ammonia, and methanol bands, and the stretching mode of CO occurs in a region without pyrimidines bands. The bending mode of water is found in the same region where one of the strongest bands of pyrimidines (~1700 cm^{-1}) is expected to be found. In the region near ~1460 cm^{-1}, the CH and OH bending modes of CH_3OH are observed. This band also overlaps some important pyrimidine bands; the same occurs in the region between ~1100 cm^{-1} and 1000 cm^{-1}, where the umbrella mode of NH_3 and the stretching mode of CO (CH_3OH) bands overlap the signatures of the pyrimidine's nucleobases. Only a few pyrimidine bands are observed between 1600 and 650 cm^{-1}. The assignments of the infrared bands observed in Figure 4 are presented in Table 2.

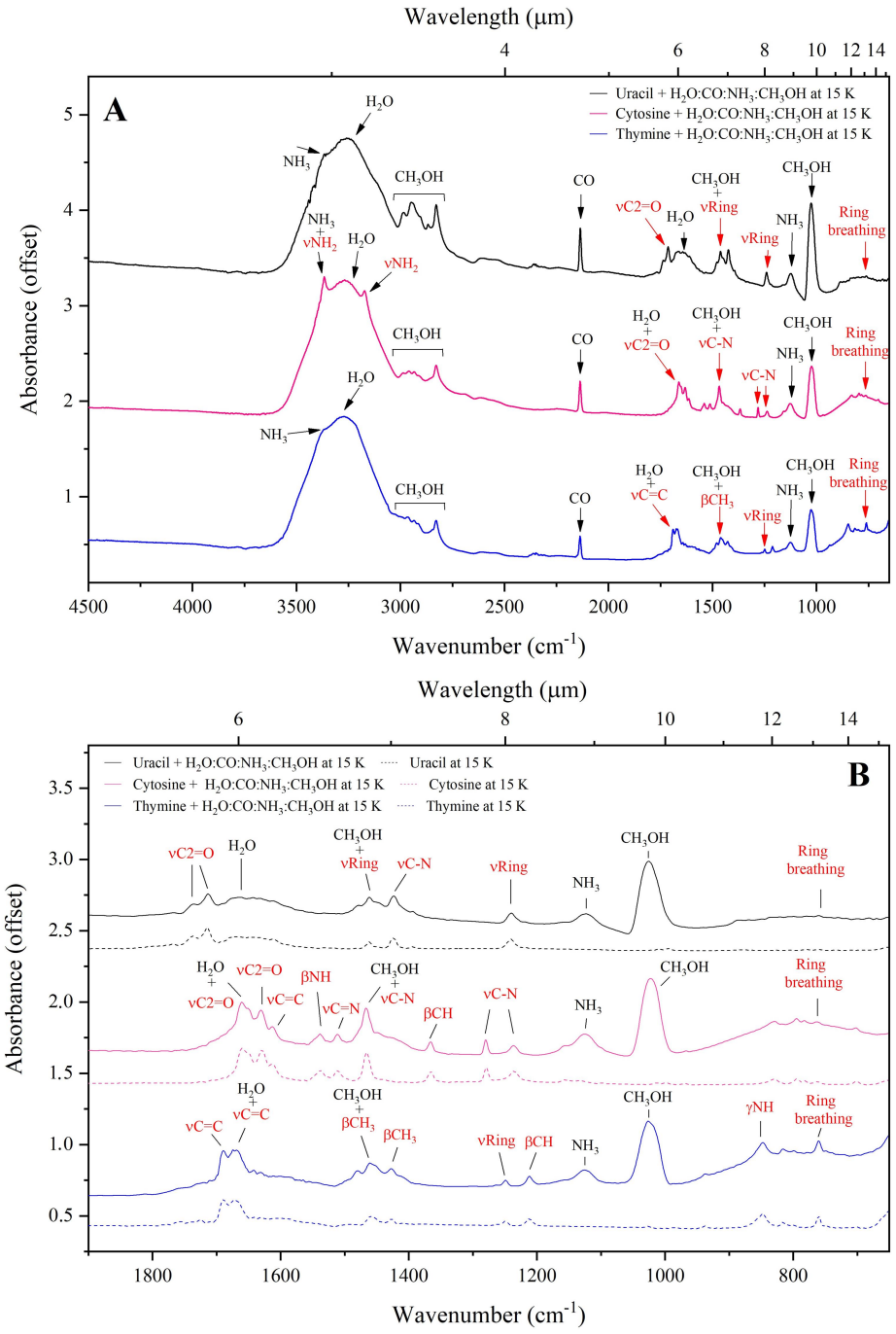

Figure 4 – (A) Mid-IR spectra of uracil (black line), cytosine (pink line), and thymine (blue line) covered by common ISM volatiles at 15 K. (B) The fingerprint region (1900-650 cm^{-1}) of pyrimidines uracil (black line), cytosine (pink line), and thymine (blue line) is covered by common ISM volatiles at 15 K. Red arrows indicate transitions assigned to the nucleobases and black arrows indicate transitions assigned to the ice mixture. Based on the column densities, the calculated mixture ratios are: H_2O:CO:NH_3:CH_3OH:Uracil (10:5:2:19:5); H_2O:CO:NH_3:CH_3OH:Cytosine (10:3:2:8:0.7); H_2O:CO:NH_3:CH_3OH:Thymine (10:0.8:0.1:3:0.3). The top scale is in wavelength (μm), while the bottom scale is in wavenumber (cm^{-1}). Abbreviations: ν – stretching; β – in plane bending; γ – out of plane bending; δ – scissoring.

Table 2 - Assignment of observed bands in uracil, cytosine, and thymine spectrum covered with common ISM volatiles at 15 K and high-vacuum.

Uracil + H_2O:CO:NH_3:CH_3OH		**Cytosine + H_2O:CO:NH_3:CH_3OH**		**Thymine + H_2O:CO:NH_3:CH_3OH**	
ν (cm^{-1})	Assignment	ν (cm^{-1})	Assignment	ν (cm^{-1})	Assignment
3370	ν_3 of NH_3 (ν_{asym})[a]	3366	ν_{asym} NH_2 of Cytosine + ν_3 of NH_3 (ν_{asym})[a]	3373	ν_3 of NH_3 (ν_{asym})[a]
3259	ν_3 of H_2O (ν_{asym})[a]	3268	ν_3 of H_2O (ν_{asym})[a]	3270	ν_3 of H_2O (ν_{asym})[a]
2948	νC-H; νN-H of Uracil	3172	ν_{sym} NH_2 of Cytosine	2970	νCH_3 of Thymine
2867	νC-H; νN-H of Uracil	2959	ν_9 of CH_3OH ($\nu_{asym}CH_3$)[a]	2918	νCH_3 of Thymine
2828	ν_3 of CH_3OH ($\nu_{sym}CH_3$)[a]	2828	ν_3 of CH_3OH ($\nu_{sym}CH_3$)[a]	2829	ν_3 of CH_3OH ($\nu_{sym}CH_3$)[a]
2136	1-0 mode of CO (ν)[a]	2136	1-0 mode of CO (ν)[a]	2137	1-0 mode of CO (ν)[a]
1737	νC=O of Uracil	1661	νC=O of Cytosine + ν_2 of H_2O (β)[a]	1689	νC=C + ν_2 of H_2O (β)[a]
1713	νC=O of Uracil	1632	νC=O of Cytosine + ν_2 of H_2O (β)[a]	1674	νC=C + ν_2 of H_2O (β)[a]
1664	ν_2 of H_2O (β)[a]	1616	νC=C of Cytosine		
1644	νC=C of Uracil + ν_2 of H_2O (β)[a]	1540	βN-H of Cytosine	1460	βCH_3 of Thymine + ν_4 and ν_{10} of CH_3OH (βCH_3 and βOH)[a]
1613	βC-H of Uracil	1515	νC=N and νC-N of Cytosine	1427	βCH_3 of Thymine
1461	νRing, νC-N and βN-H of Uracil + ν_4 and ν_{10} of CH_3OH (βCH_3 and βOH)[a]	1467	νC-N of Cytosine + ν_4 and ν_{10} of CH_3OH (βCH_3 and βOH)[a]	1249	νRing of Thymine
1423	νC-N and βN-H of Uracil	1367	βC-H of Cytosine	1212	βC-H, νC-N, νC-C, νCH_3 of Thymine
1395	βN-H; βC-H and νC-N of Uracil	1281	νC-N of Cytosine	1126	ν_2 of NH_3 (umbrella mode)[a]
1240	νRing of Uracil	1236		1026	ν_8 of CH_3OH (νCO)[a]
1124	ν_2 of NH_3 (umbrella mode)[a]	1161	νC=O of Cytosine	939	νC-N of Thymine
1025	ν_8 of CH_3OH (νCO)[a]	1126	ν_2 of NH_3 (umbrella mode)[a]	847	γN-H of Thymine
886	γN-H of Uracil	1024	ν_8 of CH_3OH (νCO)[a]	816	
760	Ring breathing of Uracil	834	γNH_2 of Cytosine	760	Ring breathing and γC=O of Thymine
		796	γC=O of Cytosine		
		785	γC=O of Cytosine		

764	Ring breathing of Cytosine
705	βRing of Cytosine

[a]Bouilloud *et al.* (2015)

Some of the pyrimidine's bands have similar wavenumbers, such as the bands 1240 cm^{-1} (νRing) and 760 cm^{-1} (Ring breathing) in uracil's spectrum, 1236 cm^{-1} (νC-N) and 763 cm^{-1} (Ring breathing) in cytosine's spectrum, and 1249 cm^{-1} (νRing) and 760 cm^{-1} (Ring breathing) in thymine's spectrum (Figure 5).

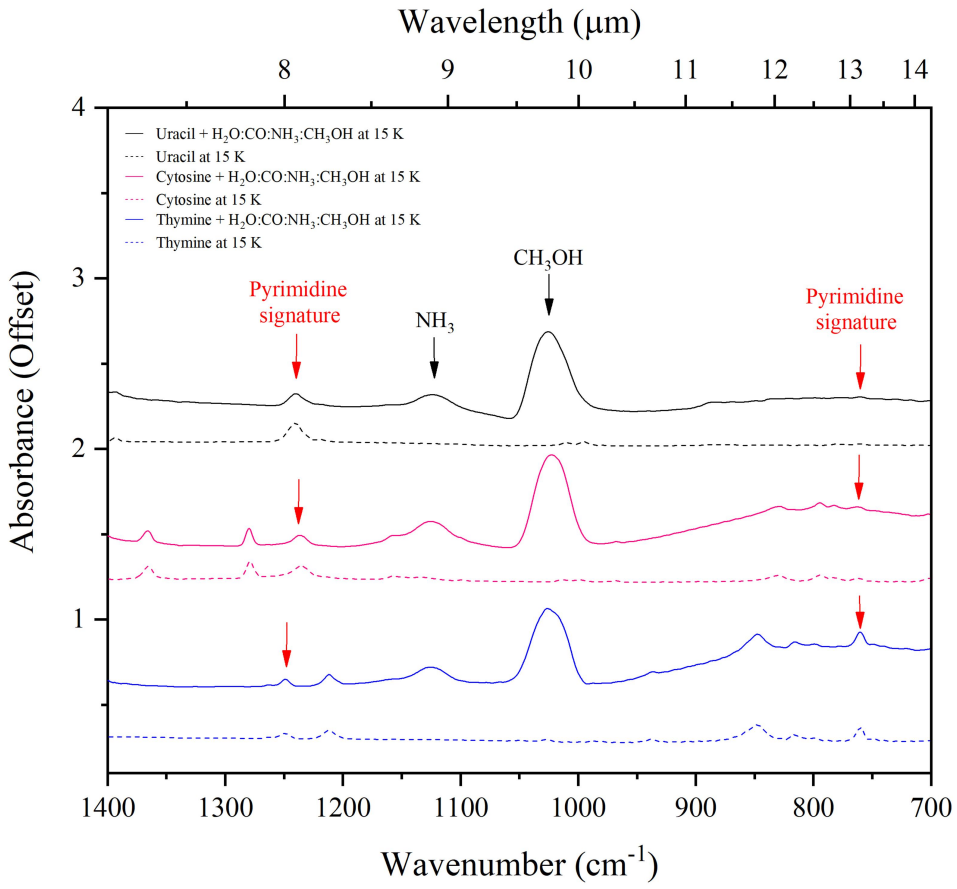

Figure 5 – The Mid-IR spectra of uracil (black line), cytosine (pink line), and thymine (blue line) in the region where are found common bands of pyrimidines (1400-700 cm^{-1}). The red arrows indicate the two pyrimidines bands that are not covered by the ice mixture containing H_2O:CO:NH_3:CH_3OH. The bands of interest are: 1242 ± 6 and 761 ± 2 cm^{-1}.

The common pyrimidines bands presented in Figure 5 are characteristic of vibration modes of the whole ring, rendering a specific signature for the ensemble of pyrimidines. Therefore, their detection is complicated by overlap with simpler interstellar organic species, which are expected to be more abundant. For example, the vibrational transitions of CH_3OCHO (1211 cm^{-1}) and H_2CO (1244 cm^{-1}) fall within the same spectral region. Additionally, the weak ring breathing mode of thymine at 760 cm^{-1} coincides with the absorption region of H_2O, further complicating the detection. Even if these bands are weak in intensity, if uracil, cytosine, or thymine were present altogether in the same astrophysical target, these common bands with similar wavenumbers may blend, making the spectral profile broader and/or more prominent. For this reason, the detection and

identification of individual pyrimidines are challenging. The overlap of these common bands complicates the quantification of each species' contribution to the overall spectral profile.

4. Challenges in Identifying Pyrimidine Bands in Interstellar Ice Spectra

The IR bands discussed in this study are characterized in Table 3, which also includes the integrated molar absorptivity (ψ) for each band, as calculated by Iglesias-Groth and Cataldo (2023). Searching the interstellar ice inventory for spectroscopic features of other simple common ice species at wavelengths similar to those discussed here for pyrimidines did not yield any identifiable bands around 8.05 and 13.14 μm, the regions of interest for pyrimidines (Boogert et al. 2015; McClure et al. 2023; Gibb et al. 2004). Recent observations made by JWST revealed that interstellar organic species such as CH_3OCHO, CH_3COOH, and HCOOH, which are expected to be more abundant than pyrimidines, exhibit infrared signals between 7.8 – 8.6 μm (Rocha et al. 2024). Therefore, even JWST would not be able to distinguish the 8.05 ± 0.08 μm band of pyrimidines from these organic compounds. Additionally, although the 13.14 ± 0.10 μm band of pyrimidines is distinguishable from the libration mode of water (13.6 μm) in laboratory spectra, it would be difficult to separate these bands in observations. Thus, detecting pyrimidines with current telescopes is impossible due to overlapping absorption features, not a matter of sensitivity or more powerful instrumentation.

Table 3 - Characterization of IR bands for uracil, thymine, and cytosine within interstellar ices.

ν(cm-1)	λ (μm)	Assignment	ψ (km/mol)[a]
1242 ± 6	8.05 ± 0.08	νRing of uracil and thymine + νC-N of cytosine	47 (U); 14 (T); 19 (C)
761 ± 2	13.14 ± 0.10	Ring breathing of uracil, thymine, and cytosine	8.7 (T);

[a] **Iglesias-Groth & Cataldo 2023**

5. Conclusions

Nucleobases are essential molecules for life since they form an integral part of the DNA and RNA present in all life forms on Earth. There is evidence for the abiotic synthesis of nucleobases outside Earth as they have been detected in meteorites (Oba *et al.* 2022) and laboratory experiments have revealed that they can be synthesized in ISM ices (Ruf *et al.* 2019, Oba *et al*. 2019). Despite this, nucleobases have never been detected in the ISM. It is unlikely that the solar system is the only place where such molecules are formed. Therefore, it is almost certain that the detection of nucleobases outside Earth is hindered by the challenges inherently associated with the detection methods. Laboratory studies, such as the one presented here, reveal the challenges to observe these complex molecules in outer space.

In the current paper, we present the infrared spectra of uracil, cytosine, and thymine, the pyrimidines nucleobases found in DNA and in RNA, embedded in an ice mixture that represents the most common volatiles found protostellar disks. Pyrimidines infrared bands can potentially lead to their identification even in environments rich in H_2O, CO, CH_3OH, and NH_3. It was possible to identify two common bands which are not covered by volatiles at 1240 and 760 cm^{-1} in uracil's spectrum; 1236 and 763 cm^{-1} in cytosine's spectrum, and 1249 and 760 cm^{-1} in thymine's spectrum. However, even these bands are not proposed as definitive spectral signatures for pyrimidines in the ISM due to the complexity of distinguishing them in the presence of overlapping features.

In a previous study, our group performed a similar experiment and analysis for the purine family (Rosa *et al*. 2023). The present study complements the previous one, not only by investigating the other family of life-related nucleobases – the pyrimidines – but also by including the effects of the presence of common ISM volatiles in the composition of the ice. While searching for common bands from each group of nucleobases (purines and pyrimidines) in astrophysical targets could potentially increase the overall signal of these low-intensity bands, the inherent challenges in detection remain significant. The low abundance of nucleobases in such environments, compared to simpler species, further complicates their detection. Thus, detecting pyrimidines with current telescopes is impossible due to overlapping absorption features, and this issue is not merely a matter of sensitivity or more powerful instrumentation.

Acknowledgements

We thank Sao Paulo Research Foundation (FAPESP) support through grant 2016/13750-6 and the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brazil (CAPES) – Finance Code 001. L.A. thank to support with resources from Conselho Nacional de Pesquisa (CNPq) 300927/2023-1, and C.A.R. and A.B. thank the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for support through grants 401567/2022-2 and 309258/2023-5. A.B. thank the Instituto Serrapilheira Serra-1912-31843. NJM recogises support from the Europlanet 2024 RI which has been funded by the European Union's Horizon 2020 Research Innovation Programme under grant agreement No. 87114 A. M.E. gratefully acknowledges the financial support of the "Fenômenos Extremos do Universo" of the Fundação Araucária.

Autor contribution statement - Caroline Antunes Rosa: Conceptualization (Lead); Formal analysis (lead); Methodology (supporting); Software (equal); Visualization (lead); Writing – original draft (lead); Writing – review & editing (lead). Alexandre Bergantini: Methodology (lead); Software (equal); Supervision (equal); Formal analysis (supporting); Conceptualization (supporting); Writing – original draft (supporting); Writing – review & editing (supporting). Enio Frota da

Silveira: Methodology (supporting); Writing – original draft (supporting); Writing – review & editing (supporting). Marcelo Emilio: Conceptualization (supporting); Writing – original draft (supporting); Writing – review & editing (supporting). Laerte Andrade: Conceptualization (supporting); Writing – original draft (supporting); Writing – review & editing (supporting). Eduardo Janot-Pacheco: Conceptualization (supporting); Writing – original draft (supporting); Writing – review & editing (supporting). Nigel Mason: Conceptualization (supporting); Writing – original draft (supporting); Writing – review & editing (supporting). Claudia Lage: Conceptualization (supporting); Formal analysis (supporting); Methodology (supporting); Supervision (equal); Writing – original draft (supporting); Writing – review & editing (supporting).

Data Availability

The data underlying this article will be shared on reasonable request to the corresponding author.

References

Baek G., Lee J.-E., Hirota T., Kim K.-T., Kim M.K., 2022, The Astrophysical Journal, 939, 84

Bergantini A., Maksyutenko P., Kaiser R.I., 2017, The Astrophysical Journal, 841, 96

Bergantini A., Góbi S., Abplanalp M.J., Kaiser R.I., 2018, The Astrophysical Journal, 852, 70

Bergantini A., Abplanalp, M. J., Pokhilko, P. l., Krylov, A. I., Shingledecker, C. N., Herbst, E., & Kaiser, R. I. 2018, ApJ, 860, 108

Boogert A.A., Gerakines P.A., Whittet D.C., 2015, Annual Review of Astronomy and Astrophysics, 53, 541

Boss A.P., 1998, Annual Review of Earth and Planetary Sciences, 26, 53

Bouilloud M., Fray N., Bénilan Y., Cottin H., Gazeau M.-C., Jolly A., 2015, Monthly Notices of the Royal Astronomical Society, 451, 2145

Brünken S., McCarthy M., Thaddeus P., Godfrey P.D., Brown R., 2006, Astronomy & Astrophysics, 459, 317

Burton A.S., Stern J.C., Elsila J.E., Glavin D.P., Dworkin J.P., 2012, Chemical Society Reviews, 41, 5459

Callahan M.P., Smith K.E., Cleaves H.J., Ruzicka J., Stern J.C., Glavin D.P., House C.H., Dworkin J.P., 2011, Proceedings of the National Academy of Sciences, 108, 13995

Charnley S.B., et al., 2005, Advances in Space Research, 36, 137

Choughuley A., Subbaraman A., Kazi Z., Chadha M., 1977, Biosystems, 9, 73

Chyba C., Sagan C., 1992, Nature, 355, 125

Colarusso P., Zhang K., Guo B., Bernath P.F., 1997, Chemical Physics Letters, 269, 39

Cooper G., Kimmich N., Belisle W., Sarinana J., Brabham K., Garrel L., 2001, Nature, 414, 879

Cooper G., Rios A.C., 2016, Proceedings of the National Academy of Sciences, 113, E3322
Coutens A., Loison J.-C., Boulanger A., Caux E., Müller H., Wakelam V., Manigand S., Jørgensen J., 2022, Astronomy & Astrophysics, 660, L6
Danger G., Bossa J.-B., De Marcellus P., Borget F., Duvernay F., Theulé P., Chiavassa T., d'Hendecourt L., 2011, Astronomy & Astrophysics, 525, A30
Deamer D.W., Pashley R., 1989, Origins of Life and Evolution of the Biosphere, 19, 21
Dullemond C., Isella A., Andrews S., Skobleva I., Dzyurkevich N., 2020, Astronomy & Astrophysics, 633, A137
Ferris J.P., Sanchez R.A., Orgel L.E., 1968, Journal of molecular biology, 33, 693
Fornaro T., Biczysko M., Monti S., Barone V., 2014, Physical Chemistry Chemical Physics, 16, 10112
Fox S.W., Harada K., 1961, Science, 133, 1923
Furukawa Y., Chikaraishi Y., Ohkouchi N., Ogawa N.O., Glavin D.P., Dworkin J.P., Abe C., Nakamura T., 2019, Proceedings of the National Academy of Sciences, 116, 24440
Gibb E., Whittet D., Boogert A., Tielens A., 2004, The astrophysical journal supplement series, 151, 35
Guélin M., Cernicharo J., 2022, Frontiers in Astronomy and Space Sciences, 9, 787567
Hayatsu R., Studier M.H., Moore L.P., Anders E., 1975, Geochimica et Cosmochimica Acta, 39, 471
Herbst E., Van Dishoeck E.F., 2009, Annual Review of Astronomy and Astrophysics, 47, 427
Iglesias-Groth S., Cataldo F., 2023, Monthly Notices of the Royal Astronomical Society, 523, 1756
Ioppolo, S., et al. 2021, Nature Astronomy, 5, 197
Janot-Pacheco E., Rachid M., Bendjoya P., Domiciano A., Antunes-Rosa C., Emilio M., Lage C., 2018, International Astronomical Union. Proceedings of the International Astronomical Union, 14, 215
Kitadai N., Maruyama S., 2018, Geoscience Frontiers, 9, 1117
Kobayashi K., 2019, Astrobiology. Springer, p. 43
Kuan Y.-J., Charnley S.B., Huang H.-C., Kisiel Z., Ehrenfreund P., Tseng W.-L., Yan C.-H., 2004, Advances in Space Research, 33, 31
Kwiatkowski J.S., Leszczyński J., 1996, The Journal of Physical Chemistry, 100, 941
Leś A., Adamowicz L., Nowak M.J., Lapinski L., 1992, Spectrochimica Acta Part A: Molecular Spectroscopy, 48, 1385
Maity S., Kaiser R.I., Jones B.M., 2015, Physical Chemistry Chemical Physics, 17, 3081
Martins Z., et al., 2008, Origins of Life and Evolution of the Biosphere, 30, 214
Materese C.K., Nuevo M., Bera P.P., Lee T.J., Sandford S.A., 2013, Astrobiology, 13, 948

Materese C.K., Nuevo M., McDowell B.L., Buffo C.E., Sandford S.A., 2018, The Astrophysical Journal, 864, 44
Materese C.K., Nuevo M., Sandford S.A., 2017, Astrobiology, 17, 761
Mathlouthi M., Seuvre A.M., Koenig J.L., 1986, Carbohydrate research, 146, 1
McClure M.K., et al., 2023, Nature astronomy, 1
Mejía C., et al., 2023, Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms, 534, 11
Nowak M., Lapinski L., Fulara J., 1989, Spectrochimica Acta Part A: Molecular Spectroscopy, 45, 229
Nuevo M., Milam S.N., Sandford S.A., 2012, Astrobiology, 12, 295
Nuevo M., Milam S.N., Sandford S.A., Elsila J.E., Dworkin J.P., 2009, Astrobiology, 9, 683
Nuevo M., Sandford S.A., 2014, The Astrophysical Journal, 793, 125
Oba Y., Takano Y., Furukawa Y., Koga T., Glavin D.P., Dworkin J.P., Naraoka H., 2022, Nature communications, 13, 2008
Oba Y., Takano Y., Naraoka H., Watanabe N., Kouchi A., 2019, Nature communications, 10, 1
Öberg K.I., Boogert A.A., Pontoppidan K.M., Van den Broek S., Van Dishoeck E.F., Bottinelli S., Blake G.A., Evans II N.J., 2011, The Astrophysical Journal, 740, 109
Öberg K.I., 2016, Chemical reviews, 116, 9631
Parker D.S., Kaiser R.I., Kostko O., Troy T.P., Ahmed M., Mebel A.M., Tielens A.G., 2015, The Astrophysical Journal, 803, 53
Radchenko E., Sheina G., Smorygo N., Blagoi Y.P., 1984, Journal of molecular structure, 116, 387
Rocha W. R. M., et al., 2024, Astronomy and Astrophysics, 683, A124
Rosa C. A., et al., 2023, Life, 13, 2208
Rubin R.H., Swenson Jr G.W., Benson R.C., Tigelaar H.L., Flygare W.H., 1971, The Astrophysical Journal, 169, L39
Ruf A., Lange J., Eddhif B., Geffroy C., d'Hendecourt L.L.S., Poinot P., Danger G., 2019, The Astrophysical journal letters, 887, L31
Saïagh K., Cloix M., Fray N., Cottin H., 2014, Planetary and Space Science, 90, 90
Saïagh K., Cottin H., Aleian A., Fray N., 2015, Astrobiology, 15, 268
Saladino R., Ciambecchini U., Crestini C., Costanzo G., Negri R., Di Mauro E., 2003, ChemBioChem, 4, 514
Saladino R., Crestini C., Ciambecchini U., Ciciriello F., Costanzo G., Di Mauro E., 2004, ChemBioChem, 5, 1558
Saladino R., Crestini C., Costanzo G., Negri R., Di Mauro E., 2001, Bioorganic & medicinal chemistry, 9, 1249

Saladino R., Crestini C., Neri V., Brucato J.R., Colangeli L., Ciciriello F., Di Mauro E., Costanzo G., 2005, ChemBioChem, 6, 1368

Saladino R., Crestini C., Neri V., Ciciriello F., Costanzo G., Di Mauro E., 2006, ChemBioChem, 7, 1707

Saladino R., Crestini C., Pino S., Costanzo G., Di Mauro E., 2012, Physics of life reviews, 9, 84

Sandford S.A., Nuevo M., Bera P.P., Lee T.J., 2020, Chemical reviews, 120, 4616

Schwartz A.W., Chittenden G., 1977, Biosystems, 9, 87

Sephton M.A., 2002, Natural product reports, 19, 292

Shock E.L., Schulte M.D., 1990, Geochimica et Cosmochimica Acta, 54, 3159

Simon M.N., Simon M., 1973, The Astrophysical Journal, 184, 757

Singh J., 2008, Journal of Molecular Structure, 876, 127

Stoks P.G., Schwartz A.W., 1979, Nature, 282, 709

Susi H., Ard J.S., 1971, Spectrochimica Acta Part A: Molecular Spectroscopy, 27, 1549

Szczepaniak K., Szczesniak M.M., Person W.B., 2000, The Journal of Physical Chemistry A, 104, 3852

Szczesniak M, Nowak M, Szczepaniak K, et al., 1985, Spectrochimica Acta Part A: Molecular Spectroscopy, 41, 237

Szczesniak M., Szczepaniak K., Kwiatkowski J., KuBulat K., Person W., 1988, Journal of the American Chemical Society, 110, 8319

Turner B.E., 1971, Astrophysical Journal, vol. 163, p. L35, 163, L35

Vasyunin, A. I., Caselli, P., Dulieu, F., & Jiménez-Serra, I. 2017, ApJ, 842, 33

Vignoli Muniz G.S., Agnihotri A.N., Augé B., Mejía C., Martinez R., Rothard H., Domaracka A., Boduch P., 2022, ACS Earth and Space Chemistry, 6, 2149

Zhang Z.E., Yang Y.-l., Zhang Y., Cox E.G., Zeng S., Murillo N.M., Ohashi S., Sakai N., 2023, The Astrophysical Journal, 946, 113