



**HAL**  
open science

## Free amino acids quantification in cloud water at the puy de Dôme station (France)

Pascal Renard, Maxence Brissy, Florent Rossi, Martin Lereboure, Saly Jaber, Jean-Luc Baray, Angelica Bianco, Anne-Marie Delort, Laurent Deguillaume

► **To cite this version:**

Pascal Renard, Maxence Brissy, Florent Rossi, Martin Lereboure, Saly Jaber, et al.. Free amino acids quantification in cloud water at the puy de Dôme station (France). 2021. hal-03409641

**HAL Id: hal-03409641**

**<https://uca.hal.science/hal-03409641v1>**

Preprint submitted on 29 Oct 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



## 1 Free amino acids quantification in cloud water at the puy de 2 Dôme station (France)

3 Pascal Renard<sup>1</sup>, Maxence Brissy<sup>2</sup>, Florent Rossi<sup>2</sup>, Martin Lereboure<sup>2</sup>, Saly Jaber<sup>2</sup>, Jean-Luc  
4 Baray<sup>1,3</sup>, Angelica Bianco<sup>1</sup>, Anne-Marie Delort<sup>2,\*</sup> and Laurent Deguillaume<sup>1,3,\*</sup>

5 <sup>1</sup> Université Clermont Auvergne, Laboratoire de Météorologie Physique, OPGC/CNRS UMR 6016, Clermont-Ferrand,  
6 France.

7 <sup>2</sup> Université Clermont Auvergne, CNRS, SIGMA Clermont, Institut de Chimie de Clermont-Ferrand (ICCF), Clermont-  
8 Ferrand, France.

9 <sup>3</sup> Université Clermont Auvergne, Observatoire de Physique du Globe de Clermont-Ferrand, UAR 833, Clermont-Ferrand,  
10 France.

11 *Correspondence to:* Anne-Marie Delort (a-marie.delort@uca.fr) and Laurent Deguillaume (laurent.deguillaume@uca.fr)

### 12 **Abstract.**

13 Eighteen free amino acids (FAAs) were quantified in cloud water sampled at the puy de Dôme station (PUY - France)  
14 during 13 cloud events. This quantification has been performed without concentration neither derivatization, using LC-  
15 MS and the standard addition method to avoid matrix effects. Total concentrations of FAAs (TCAAs) vary from 1.2  $\mu\text{M}$   
16 to 7.7  $\mu\text{M}$ , Ser (Serine) being the most abundant AA (23.7 % in average) but with elevated standard deviation, followed  
17 by Glycine (Gly) (20.5 %), Alanine (Ala) (11.9 %), Asparagine (Asn) (8.7 %), and Leucine/Isoleucine (Leu/I) (6.4 %).  
18 The distribution of AAs among the cloud events reveals high variability. TCAA constitutes between 0.5 and 4.4 % of the  
19 dissolved organic carbon measured in the cloud samples. AAs quantification in cloud water is scarce but the results agree  
20 with the few studies that investigated AAs in this aqueous medium. The environmental variability is assessed through a  
21 statistical analysis. This work shows that AAs are correlated with the time spent by the air masses in the boundary layer,  
22 especially over the sea surface before reaching the PUY. The cloud microphysical properties fluctuation does not explain  
23 the AAs variability in our samples confirming previous studies at PUY. We finally assessed the sources and the  
24 atmospheric processes that potentially explain the prevailing presence of certain AAs in the cloud samples. The initial  
25 relative distribution of AAs in biological matrices (proteins extracted from bacterial cells or mammalian cells, for example)  
26 could explain the dominance of Ala, Gly and Leu/I. AA composition of aquatic organisms (*i.e.*, diatoms species) could  
27 also explain the high concentrations of Ser in our samples. The analysis of the AAs hydrophobicity also indicates a higher  
28 contribution of AAs (80 % in average) that are hydrophilic or neutral revealing the fact that other AAs (hydrophobic) are  
29 less favorably incorporated into cloud droplets. Finally, the atmospheric aging of AAs has been evaluated by calculating  
30 atmospheric lifetimes considering their potential transformation in the cloud medium by biotic or abiotic (mainly  
31 oxidation) processes. The most concentrated AAs encountered in our samples present the longest atmospheric lifetimes  
32 and the less dominant are clearly efficiently transformed in the atmosphere, potentially explaining their low concentrations.  
33 However, this cannot fully explain the relative contribution of several AAs in the cloud samples. This reveals the high  
34 complexity of the bio-physico-chemical processes occurring in the multiphase atmospheric environment.

### 35 **1. Introduction**

36 Free or combined amino acids (AAs) that make up proteins and cell walls in living organisms are ubiquitous chemical  
37 compounds found in various environments. In the atmosphere, they are commonly detected in the condensed phases due



38 to their low vapor pressures. They have been studied and characterized in atmospheric particles (Barbaro et al., 2020;  
39 Matos et al., 2016), rain water (Mace et al., 2003a; Mace et al., 2003b; Xu et al., 2019; Yan et al., 2015), fog water (Zhang  
40 and Anastasio, 2003b) and more recently in cloud water (Bianco et al., 2016b; Triesch et al., 2021). Many efforts have  
41 been made in the past to assess their sources, their role in the atmospheric chemical and physical processes and their fate  
42 (Cape et al., 2011). However, despite those investigations, their exact role in the atmosphere is still poorly understood.  
43 They have been studied for their hygroscopic properties since they can modify the ability of the particle to act as cloud  
44 condensation nuclei (CCN) (Chan et al., 2005; Kristensson et al., 2010; Li et al., 2013) or ice nuclei (IN) (Pummer et al.,  
45 2015; Szyrmer and Zawadzki, 1997). More recently, the role of AAs in new particle formation has been also discussed  
46 (Ge et al., 2018). This raises the question of their role in aerosol and cloud formation and hence in the radiative forcing  
47 of the Earth's surface. In atmospheric aqueous phases, some AAs have been found to potentially influence atmospheric  
48 chemistry by reacting with atmospheric oxidants (Bianco et al., 2016b; McGregor and Anastasio, 2001; Zhang and  
49 Anastasio, 2003a); the study from De Hann et al. even showed that AAs can react with glyoxal to form secondary aerosol  
50 mass (De Haan et al., 2011). AAs are part of the proteinaceous fraction of aerosol particle that significantly contribute to  
51 the organic carbon and organic nitrogen fraction of aerosol particles. Their presence in aerosol particles can modify their  
52 chemical properties such as acidity/basicity and buffering ability (Cape et al., 2011; Zhang and Anastasio, 2003b). Finally,  
53 AAs are also transferred by atmospheric deposition to other ecosystems such as aquatic surfaces where they act as  
54 nutrients since they are particularly bioavailable (Wedyan and Preston, 2008). Atmospheric AAs can therefore contribute  
55 to the nutrient cycling at global scale as well as the global carbon and nitrogen cycles.

56 AAs have been detected in the atmosphere under various contrasted environmental scenarios such as urban area (Barbaro  
57 et al., 2011; Di Filippo et al., 2014; Ren et al., 2018; Zhu et al., 2020), background/rural sites (Bianco et al., 2016b; Helin  
58 et al., 2017; Samy et al., 2011; Song et al., 2017), marine environment (Mandalakis et al., 2011; Matsumoto and Uematsu,  
59 2005; Triesch et al., 2021; Violaki and Mihalopoulos, 2010) and polar regions (Barbaro et al., 2015; Feltracco et al., 2019;  
60 Mashayekhy Rad et al., 2019; Scalabrin et al., 2012). The quantity and type of AAs detected in all the compartments  
61 (aerosol particles, cloud water, rainwater) vary over a wide range. Indeed, their emissions, residence times and spatial and  
62 temporal distributions are driven by complex bio-physico-chemical processes occurring in the atmosphere (transport,  
63 chemical and biological transformations, deposition, *etc.*). Proteinaceous materials detected in the atmosphere are in  
64 majority linked with emissions of primary biological aerosol particles that notably include viruses, bacteria, fungi, algae,  
65 spores, pollens and fragments of plants and insects (Després et al., 2012; Fröhlich-Nowoisky et al., 2016). The main  
66 source is consequently from biogenic origin, but several anthropogenic sources can also contribute (industry, agricultural  
67 practices, wastewater treatment). It is suggested that AAs are directly emitted into the atmosphere or result from the  
68 transformations of proteins by enzymatic activity, decomposition by the temperature or the photochemistry (Mopper and  
69 Zika, 1987). There are some studies highlighting other possible sources such as emissions by volcanoes (Scalabrin et al.,  
70 2012), biomass burning emissions (Chan et al., 2005) and marine emissions by sea bubble bursting (Barbaro et al., 2015;  
71 Matsumoto and Uematsu, 2005). Due to the wide variety of AAs sources in the atmosphere, it is rather difficult to correlate  
72 AA concentration and speciation with specific sources: Abe et al. (2016) recently proposed to use AAs as markers for  
73 biological sources in urban aerosols (Abe et al., 2016); Matsumoto and Uematsu (2005) suggested that the major source  
74 of free amino acids (FAAs) in aerosols over the remote North Pacific are related to long-range transport from continental  
75 areas; Scalabrin et al. (2012) used AAs ratio to evaluate aerosol aging in the atmosphere.

76 The analysis of AAs in the atmosphere is essential and has been widely conducted to document the concentrations of  
77 aerosol particles, their environmental variability, and their effects on atmospheric physico-chemical processes. AAs can



78 also be transferred into the atmospheric aqueous media after activation of aerosol particles into cloud droplets. They  
79 consequently contribute to the complex dissolved organic matter measured in clouds that is composed by a significant  
80 fraction of biological-derived material (lipids, peptides, carbohydrates...) (Bianco et al., 2018; Cook et al., 2017; Zhao et  
81 al., 2013). However, only few studies focus on the detection of AAs in cloud water (Bianco et al., 2016b; Triesch et al.,  
82 2021) mainly because of the inherent difficulty to sample clouds. AA concentration in cloud water results from the  
83 dissolution of the soluble fraction of the aerosol particles acting as CCN and IN; some very recent studies also argue that  
84 AAs could be processed in the cloud medium by the biological activity (Bianco et al., 2019). For instance, the  
85 biodegradation of AAs was demonstrated to occur in rainwater (Xu et al., 2020) and in microcosms mimicking the cloud  
86 environment (Jaber et al., 2021). The presence of transcripts of genes coding for AAs and proteins biosynthesis and  
87 biodegradation has been also shown directly in cloud water samples (Amato et al., 2019). AAs can also be photo-  
88 transformed by abiotic processes mainly implying oxidants (Jaber et al., 2021). They can produce other compounds such  
89 as carboxylic acids, nitrate, and ammonia (Berger et al., 1999; Berto et al., 2016; Bianco et al., 2016a; Marion et al., 2018;  
90 Pattison et al., 2012), thus potentially contributing to the formation in aqueous phase of secondary organic aerosol  
91 (“aqSOA”). It is therefore crucial to document AA concentration levels and speciation in clouds.

92 This aim of this work is devoted to the quantification of FAAs in cloud waters. This is quite a challenge due to the  
93 chemical complexity of the cloud medium and the low concentration of FAAs ( $\approx \mu\text{M}$ ). In atmospheric waters, namely  
94 fog (Zhang and Anastasio, 2003b), rain (Gorzelska et al., 1992; Mopper and Zika, 1987; Xu et al., 2019; Yan et al., 2015)  
95 and clouds (Bianco et al., 2016b), the main technique that has been commonly used is liquid chromatography coupled  
96 with fluorescence detection. This approach is based on pre- or post-column derivatization of the AAs to increase the  
97 sensitivity and simplify the separation by chromatography, but it is time consuming. More recently, Triesch et al. (2021)  
98 have used liquid chromatography hyphenated to mass spectrometry (LC-MS) to detect derivatized AAs after concentration  
99 of cloud water samples. The use of LC-MS represents a significant improvement as it allows a unique identification. We  
100 propose here to go further using LC-MS without pre-concentration and derivatization of the sample. In addition, to avoid  
101 matrix effect, we propose to quantify the AAs by the standard addition method (Hewavitharana et al., 2018). Cloud  
102 sampling is performed at the puy de Dôme station (PUY) in France offering possibility to collect 13 samples for various  
103 environmental conditions. Variability of cloud AA concentrations together with cloud bio-physico-chemical properties  
104 and air mass history is thus discussed in this work.

## 105 2. Methods/Materials

### 106 2.1 Site and cloud sampling

107 13 clouds were sampled from 2014 to 2020 at the puy de Dôme station (PUY) in France (45.77 °N, 2.96 °E; 1465m a.s.l.).  
108 This mountain observatory is part of the multi-site platform CO-PDD combining *in situ* and remote sensing observations  
109 at different altitudes (Baray et al., 2020). PUY belongs to international atmospheric survey networks: ACTRIS (Aerosols,  
110 Clouds, and Trace Gases Research Infrastructure), EMEP (the European Monitoring and Evaluation Program) and GAW  
111 (Global Atmosphere Watch) as example. Meteorological parameters, atmospheric gases, aerosols, and clouds are  
112 monitored over long-term period, to investigate the bio-physico-chemical processes linking those elements and to evaluate  
113 the anthropogenic forcing on climate.

114 The sampling is performed using cloud water collectors previously described (Deguillaume et al., 2014), under non-  
115 precipitating and non-freezing conditions. Before cloud collection, the cloud impactors are cleaned and sterilized by



116 autoclaving. After sampling, the cloud waters are stored in sterilized bottles; a fraction of the aqueous volume is filtered  
117 using a 0.2  $\mu\text{m}$  nylon filter within 10 min after sampling to eliminate non soluble particles and microorganisms. The  
118 samples are then kept in the dark and stored at 4°C or kept at -80°C (depending on the targeted compounds) before the  
119 analyses.

## 120 2.2 Physical, chemical, and microbiological characterization of clouds

121 A systematic characterization is performed on cloud samples allowing to document the available PUYCLOUD database  
122 ([http://opgc.fr/vobs/so\\_interface.php?so=puycloud](http://opgc.fr/vobs/so_interface.php?so=puycloud)) of the cloud water chemical and biological composition (Renard et  
123 al., 2020). These data are reported in Table S1 for the studied cloud events.

124 Chemical composition analyses are performed on cloud samples: pH, total organic carbon (TOC) concentration, and  
125 concentrations of main inorganic ionic species. TOC analyses are performed with a TOC analyzer (Shimadzu). The  
126 spectrofluorimetric method based on the reactivity of p-hydroxyphenilacetic acid with horseradish peroxidase is used to  
127 measure the concentration of  $\text{H}_2\text{O}_2$  in cloud water (Wirgot et al., 2017). Ionic inorganic species ( $\text{Ca}^{2+}$ ,  $\text{K}^+$ ,  $\text{Mg}^{2+}$ ,  $\text{Na}^+$ ,  
128  $\text{NH}_4^+$ ,  $\text{Cl}^-$ ,  $\text{SO}_4^{2-}$  and  $\text{NO}_3^-$ ) are measured by ion chromatography (Deguillaume et al., 2014).

129 Cloud microphysical properties are determined with the Gerber particle volume monitor-100 (PVM-100) providing liquid  
130 water content (LWC) and effective droplet radius ( $r_e$ ) parameters.

131 The biology of cloud water is also assessed by quantification of bacteria density (CFU  $\text{mL}^{-1}$ ) at 17°C (Vaïtilingom et al.,  
132 2012) and adenosine triphosphate (ATP) concentration is measured using the BioThema<sup>®</sup> ATP Biomass kit HS (Koutny  
133 et al., 2006).

## 134 2.3 Quantification of Amino Acids (AAs)

### 135 2.3.1 Sample preparation

136 Before analysis by LC-MS, in order to apply the standard addition method to quantify AAs (Hewavitharana et al., 2018),  
137 standard solutions are used to spike cloud water samples. Standard solutions are prepared in ultra-pure water and  
138 contained alanine (Ala, SIGMA-ALDRICH), arginine (Arg, SIMAFEX), asparagine (Asn, SIGMA-ALDRICH),  
139 aspartate (Asp, SIGMA-ALDRICH), glutamine (Gln, SIGMA-ALDRICH) glutamic acid (Glu, SIGMA-ALDRICH),  
140 glycine (Gly, MERCK), histidine (His, SIGMA-ALDRICH), leucine/isoleucine (Leu/I, SIGMA-ALDRICH), lysine (Lys,  
141 SIGMA-ALDRICH), methionine (Met, SIGMA-ALDRICH), phenylalanine (Phe, ACROS organics), proline (Pro,  
142 SIGMA-ALDRICH), serine (Ser, SIGMA-ALDRICH), threonine (Thr, SIGMA-ALDRICH), tryptophan (Trp, SIGMA-  
143 ALDRICH), tyrosine (Tyr, SIGMA-ALDRICH), valine (Val, SIGMA-ALDRICH), cysteine (Cys, SIGMA-ALDRICH).  
144 The ratio between the sample volume and the standard solution volume is respectively (9:1). The mixture is then vortex  
145 mixed for 1 min.

146 Ten samples ready for LC-MS analysis are prepared containing the original cloud water added with 20 AAs at final  
147 concentrations set to 1.0, 5.0, 10, 25, 50, 100, 150, 500  $\mu\text{g L}^{-1}$ . This range of concentrations is appropriate considering  
148 previous quantification of AAs in cloud waters sampled at PUY (Bianco et al., 2016b). This also allows to cover large  
149 range of AA concentrations that can be highly variable depending on the cloud events. L-Lys isotope ( $^{13}\text{C}_6$ , 99 %;  $^{15}\text{N}_2$ ,  
150 99 %) is also added to each sample at the concentration of 15  $\mu\text{g L}^{-1}$  for mass calibration ( $m/z = 155.1273$ ).



### 151 2.3.2 LC-MS analysis

152 LC-MS analyzes are performed using an UltiMate™ 3000 (Thermo Scientific™) LC equipped with a Q-Exactive™  
153 Hybrid Quadrupole-Orbitrap™ Mass Spectrometer (Thermo Scientific™) ionization chamber. Chromatographic  
154 separation of the analytes is performed on BEH Amide/HILIC (1.7 μm, 100 mm × 2.1 mm) column with column  
155 temperature of 30°C. The mobile phases consist of 0.1 % formic acid and water (A) and 0.1% formic acid and acetonitrile  
156 (B) with a 0.4 mL min<sup>-1</sup> flow rate. A four-step linear gradient is applied during the analysis: 10 % A and 90 % B in 8 min,  
157 42 % A and 58% B in 0.1 min, 50 % A and 50 % B for 0.9 min and 10 % A and 90 % B for 3 min.

158 The Q-Exactive ion source is equipped with electrospray ionization (ESI) and the Q-Orbitrap™. The volume of injection  
159 is 5 μL, and the flow injection analyses are performed for individual AA solutions to obtain the mass spectral data, from  
160 which ions are carefully chosen for analysis in the selected ion monitoring (SIM) mode, using the above-mentioned  
161 parameter conditions. The mass resolution is set to 35000 FWHM (full width at half maximum), and the instrument is  
162 tuned for maximum ion throughput. AGC (automatic gain control) target or the number of ions to fill C-Trap is set to 10<sup>5</sup>  
163 with injection time of 100 ms. Tests with standard solution and cloud water samples show a better sensitivity in positive  
164 mode of ionization for all AAs and with a preference for [M+H]<sup>+</sup> ionization (ESI+). Other Q-Exactive™ generic  
165 parameters are: N<sub>2</sub> flow rate set at 13 a.u, sheath gas (N<sub>2</sub>) flow rate set at 50 a.u, sweep gas flow rate set at 2 a.u, spray  
166 voltage set at 3.2 kV in positive mode, capillary temperature set at 320°C, and heater temperature set at 425°C.

167 Analysis and visualization of the data set are performed using Xcalibur™ 2.2 software; it allows controlling and  
168 processing data from Thermo Scientific™ LC-MS systems and associated instruments. Examples of chromatograms and  
169 MS spectra for three AAs (Ser, Val and Trp) are presented in Figures S1 (a, b, c). For quality control, one cloud sample  
170 has been analyzed in MS<sup>2</sup> to check the presence of isobaric molecules. The peak with a retention time of 2.89 min and  
171 m/z of 118,0867 [M+H]<sup>+</sup> has been found to correspond to the mixture of 2 isobaric molecules: valine and betaine (Figure  
172 S2). Therefore, Val cannot be quantified. Leu and Ile could not be also distinguished as there are isobaric with the same  
173 retention time (hereafter, Leu/I). Cys is not quantifiable as it forms S-S bonds. Consequently, 18 AAs can be quantified  
174 in this study: Ala, Arg, Asn, Asp, Gln, Glu, Gly, His, Leu/I, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr. The retention times  
175 and exact masses measured by LC-MS of all the AAs are summarized in Table S2.

### 176 2.3.3 Standard addition

177 Cloud water is a complex mixture, conducive to disturbance in the LC-MS analytical signal. To restrain this matrix effect,  
178 the AAs quantification is performed with the method of the standard addition, which consists of the addition of a series  
179 of small volumes of concentrated standard to an existing unknown. For each AA, this method provides a calibration curve.  
180 Figure S3 presents, as an example, how the concentration of Gly is measured for a particular cloud event (11 Mar cloud)  
181 using the standard addition method. The magnitude of the intercept on the x-axis of the trendline is the original  
182 concentration of Gly.

183 Table S3 displays calibration curve data measured for the 13 different cloud samples for each AA. The linearity of the  
184 calibration curves is attested by the high R<sup>2</sup> values (> 0.95). The AA concentrations and their standard deviation (STD)  
185 are calculated according to the equation from Broekaert and Daniel (Broekaert, 2015). More details can be found in SI  
186 (Figure S3 and attached explanations of the calculations).



## 187 2.4 Evaluation of air mass history

188 The CAT model (Computing Atmospheric Trajectory Tool) is a three-dimensional (3D) forward/backward kinematic  
189 trajectory code which has been recently developed and used to characterize the atmospheric transport of air masses  
190 reaching PUY station (Renard et al., 2020). Backward trajectories clusters have been calculated for all clouds included in  
191 the PUYCLOUD database. The temporal resolution of the backtrajectories is 15 min and the total duration is 72 h. The  
192 model is initialized with wind fields from the ECMWF ERA-5 reanalyzed with a horizontal resolution of 0.5° and 23  
193 vertical pressure levels between 200 and 1000 hPa. On the basis of the atmospheric boundary layer height (ABLH) and  
194 altitude of topography interpolated for each trajectory point, this numerical tool allows to calculate the percentage of  
195 points above the sea and the continental surfaces (Sea surface vs Continental surface), hereafter named “zone”. A “zone  
196 matrix” is thus constructed from CAT model outputs and used for a statistical classification of each cloud event. All the  
197 data relative to the 13 clouds of this study are reported in Table S1.

198 This classification proposed by Renard et al. (2020) is based on a statistical analysis considering the cloud chemical  
199 concentrations of the PUYCLOUD dataset. This allows clustering clouds in four categories: “Highly marine”, “Marine”,  
200 “Continental” and “Polluted”. The “Marine” clouds come predominantly from western sectors and have the lowest ion  
201 concentrations. Marine category is predominant and the most “homogeneous” in terms of concentrations. The “Highly  
202 marine” category with a similar air mass history, gathers the clouds with the highest sea-salts concentrations (Cl<sup>-</sup>, Mg<sup>2+</sup>  
203 and Na<sup>+</sup>). The Continental category corresponds mainly to air masses arriving from the northeast sector with high  
204 concentrations of potentially anthropogenic ions (NH<sub>4</sub><sup>+</sup>, NO<sub>3</sub><sup>-</sup> and SO<sub>4</sub><sup>2-</sup>). Finally, the “Polluted” category gathers some  
205 cloud samples with the highest anthropogenic ion concentrations. All the data relative to the clouds studied in the present  
206 work are reported in Table S1.

## 207 2.5 Statistical analysis

208 With the objective to categorize cloud samples, we performed agglomerative hierarchical clustering (AHC), an iterative  
209 classification, based on AA concentrations. The AHC dendrogram shows the progressive grouping of the data. The  
210 dissimilarity between samples is calculated with the Ward’s agglomeration method using Euclidean distance. The number  
211 of categories to retain is automatically defined on the base of the entropy (Addinsoft, 2020).

212 A large variability of the AA concentrations and relative proportions in the 13 cloud samples from PUY is observed. In  
213 order to better understand this variability, a PLS regression is performed to analyze the correlations between the  
214 explanatory (X) and dependent (Y) variables. The Xs variables gather the biological (ATP and bacteria density), physical  
215 (temperature and pH) and chemical (TOC, Ca<sup>2+</sup>, K<sup>+</sup>, Mg<sup>2+</sup>, Na<sup>+</sup>, NH<sub>4</sub><sup>+</sup>, Cl<sup>-</sup>, NO<sub>3</sub><sup>-</sup> and SO<sub>4</sub><sup>2-</sup> concentrations) parameters, the  
216 “zone” matrix (Sea/Continental surface </> ABLH), as well as the seasons. The Ys variables are the 18 AA concentrations  
217 (Ala, Arg, Asn, Asp, Gln, Glu, Gly, His, Leu/I, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr). The Mann–Whitney nonparametric  
218 tests are carried out to validate significant differences (p-value < 0.05) between two groups (Renard et al., 2020).

## 219 3. Results

### 220 3.1 Evaluation of LC-MS technique for a direct measurement of AAs in cloud

221 The analytical method used in this study allows assaying AAs directly in cloud samples. MS hyphenated to LC allows  
222 the analysis of the underivatized and non-concentrated analytes, avoiding potential biases and time-consuming processes.  
223 The standard addition method also prevents matrix effects which are very commonly encountered with environmental



224 matrices (Hewavitharana et al., 2018). 18 AAs in cloud water sampled at PUY have been identified and their  
225 concentrations quantified (Table S1). Concentrations and standard deviation (STD) values obtained for all AAs and cloud  
226 samples are reported in Table S3 and detailed in Figure S3. The median STD is 12 nM (ranging from 6 nM for Trp to 44  
227 nM for Ser). The relative standard deviation (RSD) ranges from 8 % for Ala to 119 % for Arg (median: 23 %).

228 The STD values, as calculated in this work in the context of a standard addition (equation detailed in Figure S3), could  
229 be compared to the limit of quantification (LOQ) established in works using internal standard method (Broekaert, 2015).  
230 Both equations are similar and provide comparable results. However, the precision depends on the number of standard  
231 points added in the method, and not on the number of replicates. The values in this work are globally low and consistent  
232 with those reported in previous works on cloud waters and aerosol particles (Table S4). A recent study performed by  
233 Triesch et al. (2021) was able to quantify Val in cloud water samples, but they could not measure Arg, Asn, His, Lys, Cys  
234 and Tyr concentrations. Triesch et al. (2021) is also based on LC-MS (Orbitrap™), but with samples concentrated (factor  
235 44) and derivatized with a pre-column. They reported LOQ values ranging from 0.2 to 1.0  $\mu\text{g L}^{-1}$ , vs STD from 1.1 to 4.6  
236  $\mu\text{g L}^{-1}$  in this study. STDs are also within the same range of magnitude than those reported on aerosol particles by Helin  
237 et al. (2017) using direct injection of extracted AAs in LC-MS (triple-quadrupole technology), with values varying from  
238 4 to 160 nM, vs STD values from 8 to 44 nM in this work.

239 Looking more carefully at the median of the AA concentrations RSDs (calculated from data displayed in Table S3), it  
240 appears that some AAs (Ala, Gly, Leu/I, Pro, Ser and Thr) have low RSDs (from 8 to 13 %) while others (Tyr, Lys, Trp,  
241 Gln, Met and Arg) present higher RSDs (from 44 to 119 %). The RSD values obtained in this work are within the same  
242 range of order than those reported by Helin et al. (2017). To conclude, these uncertainties do not change the final range  
243 of magnitude of the AA concentrations.

### 244 3.2 Cloud physico-chemical characteristics

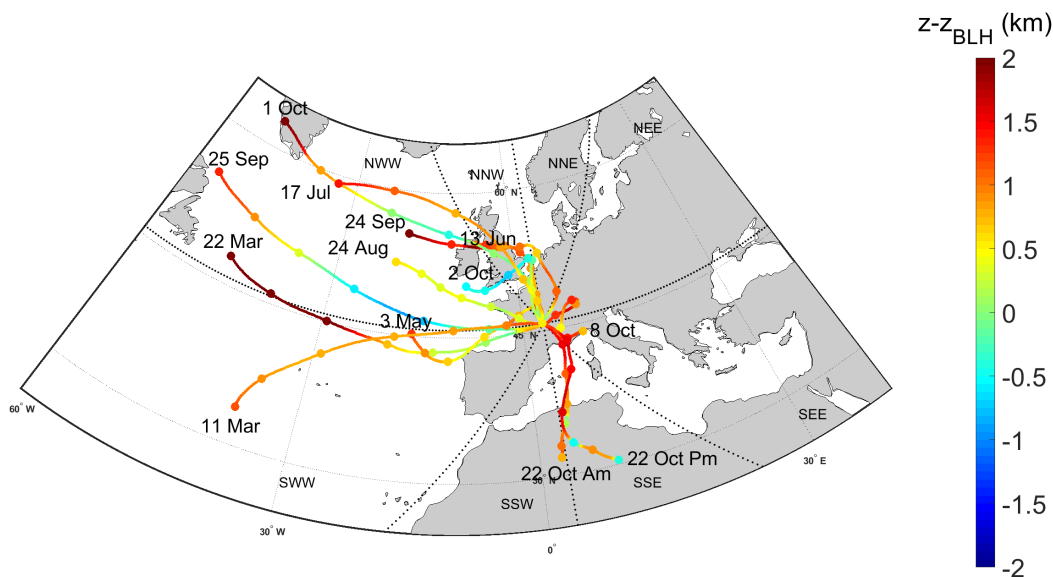
245 Table S1 presents data characterizing properties of cloud samples (chemical composition, microphysical properties, air  
246 mass history). Among the 13 studied clouds, 12 clouds are classified as “Marine” according to their ion concentrations  
247 (Renard et al. 2020). 17 Jul cloud from the North-East is classified as “Continental” due to its  $\text{NH}_4^+$ ,  $\text{NO}_3^-$  and  $\text{SO}_4^{2-}$   
248 concentrations that are significantly higher than the other studied clouds (Table S1).

249





250 Similarly to the work performed by Renard et al. (2020), the CAT model is used to characterize the air mass history of  
251 the cloud samples. Figure 1 represents the mean backtrajectories calculated over the sampling period of the 13 cloud  
252 samples; Figure S4 presents the backtrajectory calculations, every hour, for individual cloud event over the sampling  
253 period. The CAT model provides a “zone matrix” (Table S1) gathering the percentage of time spent by the air masses  
254 over the Sea surface and over the Continental surface with the discrepancy between the presence in the boundary layer (<  
255 ABLH) or in the free troposphere (> ABLH). During the 72h backtrajectories, the air masses, in average, spent a  
256 significant time in free troposphere ( $\approx 80\%$ ), and above the Sea surface (56%) (Figure 1 and Table S1). This is consistent  
257 with the conclusion from Renard et al. (2020) arguing that the marine category is the most encountered one at PUY; a  
258 category characterized by a low ionic content. However, even if the sampled clouds belong mainly to one category  
259 (“Marine”), they present chemical compositions that vary significantly from one sample to the other. This is discussed in  
260 the following section where AAs content is presented, and its variability analyzed.



261

262 **Figure 1.** Backtrajectory plots of air masses reaching PUY. Colors correspond to the air mass height minus the atmospheric  
263 boundary layer height (ABLH). Positive values (> ABLH, red) indicate the air mass is in the free troposphere. Negative values  
264 (< ABLH, blue) indicate the air mass is in the boundary layer. Each trajectory plot is the mean value of a cluster of 45 CAT  
265 trajectories calculated over 72 h, every hour from the begin to the end of the cloud sampling period. Trajectory points are  
266 calculated every 15 min and dots on the figure indicate 12 h intervals. All the trajectory clusters (without averaging) for each  
267 of the 13 events are given in Figure S4.

### 268 3.3 Quantification of AAs in cloud waters

#### 269 3.3.1 Concentration and distribution of AAs at PUY

270 AA concentrations (in nM) measured in the 13 cloud samples are reported in Table S1. Figure 2 represents the distribution  
271 of AA concentrations; minimum, maximum, mean, STD, and RSD of concentrations of those compounds are reported in  
272 Table 1.

273

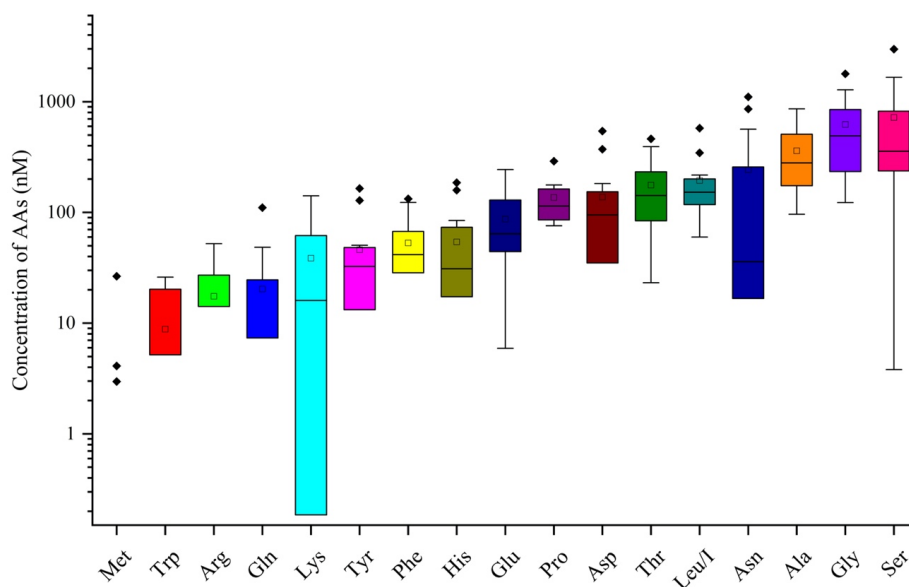
274



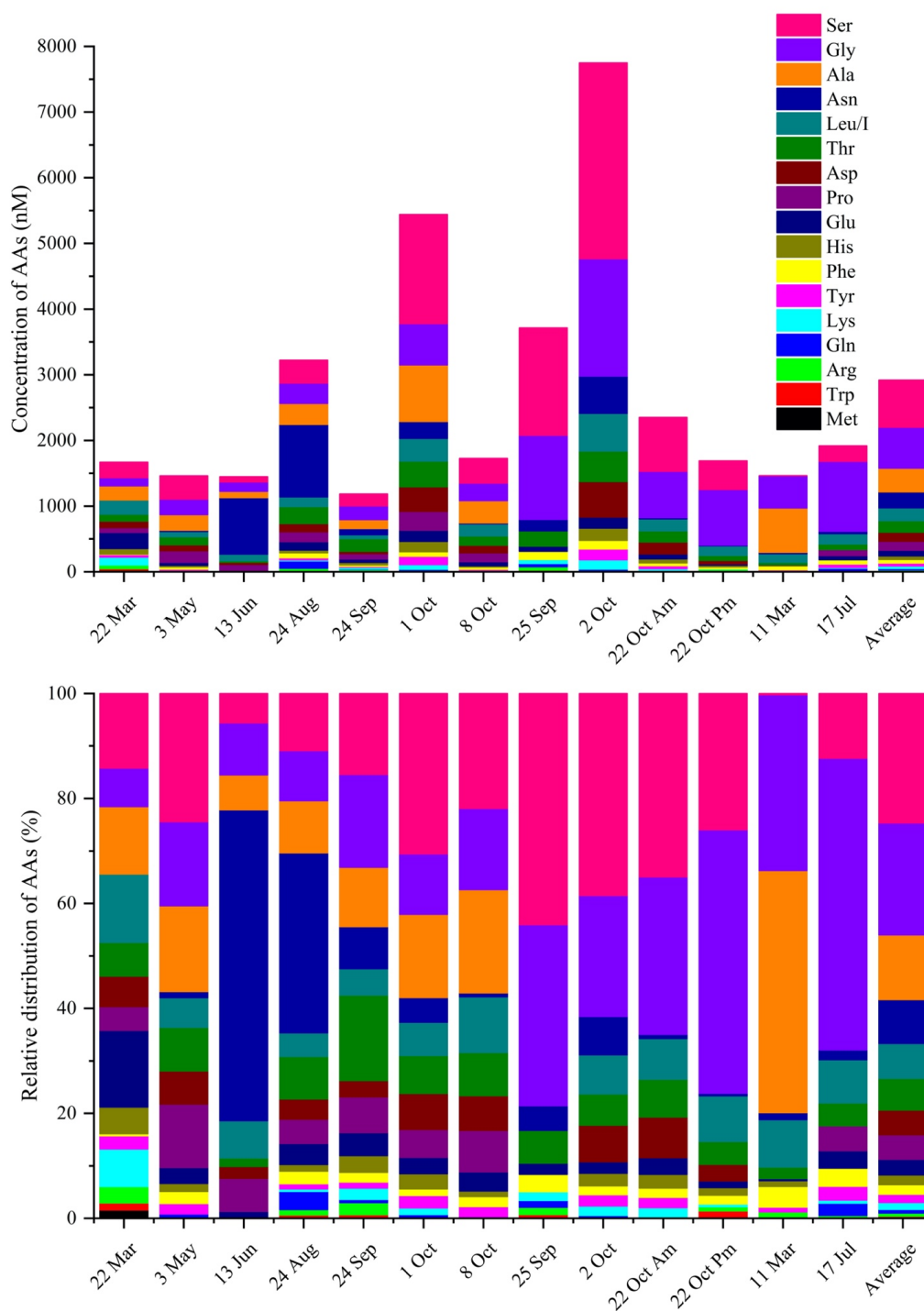
275 **Table 1. Distribution of AA concentrations measured in the 13 clouds sampled at PUY: minimum, maximum, mean, standard**  
 276 **deviation (STD), and relative standard deviation (RSD).**

Label	Minimum (nM)	Maximum (nM)	Mean (nM)	$\sigma$ (nM)	RSD
Ser	4	2983	721	866	120 %
Gly	123	1787	622	507	81 %
Ala	96	862	360	270	75 %
Asn	8	1105	264	375	142 %
Leu/I	60	577	194	141	72 %
Thr	23	462	176	133	75 %
Asp	33	543	165	166	100 %
Pro	76	290	137	72	53 %
Glu	6	244	87	70	81 %
His	16	185	65	61	93 %
Phe	6	133	57	39	68 %
Tyr	13	165	55	50	91 %
Lys	0	141	50	48	96 %
Gln	2	111	33	36	108 %
Arg	4	52	25	17	69 %
Trp	3	26	14	9	66 %
Met	3	27	11	13	119 %
<b>TCAA</b>	<b>1187</b>	<b>7749</b>	<b>2696</b>	<b>1936</b>	<b>72 %</b>

277 The total concentrations of free amino acids (TCAAs) vary significantly between cloud samples: the lowest concentration  
 278 is 1.2  $\mu\text{M}$  (24 Sep cloud), and the highest one is 7.7  $\mu\text{M}$  (2 Oct cloud) while the mean value is equal to 2.7  $\mu\text{M}$  (Table 1  
 279 and Figure 3a). In detail, Ser is the most abundant AA in the 13 cloud samples, with the highest STD (from 4 to 2983  
 280 nM), followed by Gly (from 123 to 1787 nM), Ala (from 96 to 862 nM), and Asn (from 8 to 1105 nM) (Figure 2). This  
 281 ranking seems common and ubiquitous, from polar to urban sites, in clouds as in rainwaters or aerosols (Table S4). Ser  
 282 is also preponderant in marine clouds at Cabo Verde Islands (Triesch et al., 2021) and rural fogs in Northern California  
 283 (Zhang and Anastasio, 2003b).



284 **Figure 2. Distribution of each AA for the 13 cloud samples. AA concentrations are in logarithmic scale. The bottom and top**  
 285 **lines of the box correspond to the 25<sup>th</sup> and 75<sup>th</sup> percentiles, respectively. The middle line represents the median value and the**  
 286 **square the mean value. The ends of the whiskers are the 10<sup>th</sup> and 90<sup>th</sup> percentiles, and the filled diamonds are outliers**  
 287 **(concentrations above the 90<sup>th</sup> percentile).**  
 288



290

291  
292

Figure 3. a. Distribution(nM) and b. relative contributions (% nM) of AAs molar concentrations in each cloud event sampled at PUY.

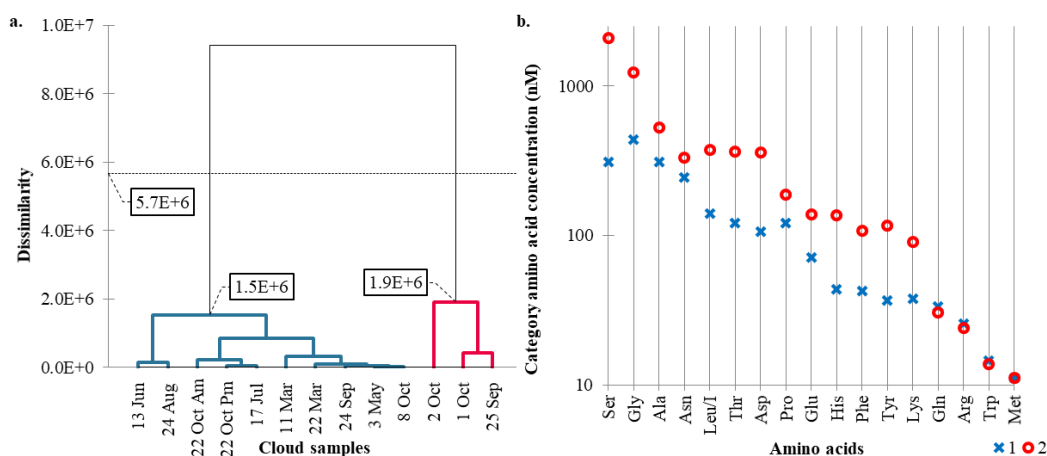


293 Figure 3 illustrates for each cloud event the relative and absolute molar concentrations of AAs. As discussed above, the  
294 TCAAs strongly vary between the different cloud events (Figure 3a). Their relative concentrations (Figure 3b) also vary  
295 among the cloud samples. For example, Ser contribution exceeds 50 % in 25 Sep cloud, while Ser is almost absent in 11  
296 Mar cloud sample, and vice versa for Ala. Asn prevails in 13 Jun and 24 Aug clouds. Nevertheless, the relative  
297 concentrations are quite similar, and the highest TCAAs do not seem to be explained by the mere presence, in excess, of  
298 a single AA.

299 Agglomerative hierarchical clustering (AHC) confirms these observations. AHC, used to categorize cloud samples based  
300 on the AA concentration, successfully groups the 13 observations, with a satisfactory cophenetic correlation (correlation  
301 coefficient between the dissimilarity and the Euclidean distance matrices) of 0.79 (Figure 4a).

302 The dotted line in Figure 4a represents the degree of truncation (dissimilarity =  $5.7 \cdot 10^6$ ) of the dendrogram used for  
303 creating categories. This truncation is automatically chosen, based on the entropy level. The AHC profile plot (Figure 4b)  
304 details the average composition of these two categories determined from the 18 AAs. In detail, the blue category gathers  
305 10 cloud samples with lower AA concentrations. This blue category is the most homogeneous (within-class variance =  
306  $3.7 \cdot 10^5$ ), compared to the red category (within-class variance =  $1.2 \cdot 10^6$ ). Conversely, the red one, more heterogeneous,  
307 gathers 3 cloud samples with higher AA concentrations except for Met (absent in most of the 13 samples). Note that the  
308 13 Jun and 24 Aug cloud samples are isolated in the blue category due to their high Asn concentration.

309 These two AHC categories reflect the variability of some AAs in the 13 cloud samples. Because the computed p-value in  
310 the Mann-Whitney test (Not shown) is lower than the significance level  $\alpha = 0.05$ , the distribution of AA (Asp, Gly,  
311 His, Leu/I, Lys, Phe, Ser, Thr and Tyr) concentrations can be accepted as significantly different between both AHC  
312 categories.



313

314 **Figure 4. a. Dendrogram representing the agglomerative hierarchical clustering (AHC) based on dissimilarities using the**  
315 **Ward's method on concentrations of the 18 AAs. The 13 cloud samples are assigned to one of two established categories by**  
316 **entropy (i.e., dissimilarity <  $5.7 \cdot 10^6$ ).** b. Profile plot established by the AHC from the 18 main AAs. The Y axis, in logarithmic  
317 scale, displays the average AA concentrations of the category.

318



### 319 3.4 Comparison with previous studies on clouds, fogs, and rain

320 To our knowledge, only two studies refer to the AA characterization in cloud water (Table 2). A first one has been  
321 performed at PUY on 25 cloud samples; 16 AAs have been quantified by a different analytical procedure, using high-  
322 performance liquid chromatography connected to a fluorescence detection after derivatization of the AAs (Bianco et al.,  
323 2016b). They report a mean TCAA concentration of 2.67  $\mu\text{M}$  with values ranging from 1.30 to 6.25  $\mu\text{M}$ . These reported  
324 concentrations are within the same range of magnitude those of the present study (from 1.2 to 7.7  $\mu\text{M}$ ). However, the  
325 main difference between the present study and Bianco's study lies in the relative concentrations of the various AAs. Trp,  
326 Leu/I, Phe and Ser were the four most concentrated AAs (mean concentrations of 563, 548, 337 and 281 nM, respectively)  
327 while we found Ser, Gly Ala, and Asn as the most abundant AAs (mean concentrations of 721, 622, 360, and 264 nM,  
328 respectively). This discrepancy could result from the differences in the analytic tools, or on sampling characteristics, *i.e.*,  
329 cloud waters in the Bianco's study have been sampled during two short periods (March/April and November 2014),  
330 whereas in the present work, cloud waters have been collected over 6 years and covering different seasons.

331 The second one is a recent study reporting the characterization of AAs in cloud waters sampled at a marine site, the Cape  
332 Verde Islands (Triesch et al., 2021). Results also indicate variability of AA concentrations in cloud samples with values  
333 varying from 11.2 to 489.9  $\text{ng m}^{-3}$ . These TCAAs are within the same range of magnitude as observed in this study (from  
334 39 to 244  $\text{ng m}^{-3}$ ). In both studies (Cape Verde Islands and PUY), Ser, Ala and Gly are amongst the major AAs, but Asp  
335 is found to be highly concentrated in the Cape Verde's study. They also find that the relative distributions of these four  
336 AAs greatly change during the campaign period. Gly and Ser are found to be the dominant AAs in the first seven cloud  
337 samples, while Ala and Asp are also highly present together with Gly and Ser during the last part of the campaign (3  
338 samples). They conclude that these differences are due to the different types of clouds sampled during this campaign.  
339 Triesch et al. (2021) show that some clouds present low TCAAs (less than 65  $\text{ng m}^{-3}$ ) with a dominance of Gly and Ser  
340 and a second group with elevated TCAAs (more than 250  $\text{ng m}^{-3}$ ) and Ser as major AA, followed by Ala and Gly. This  
341 enrichment of cloud waters in AAs could be due to oceanic sources or may be the result of *in situ* formation of AAs in  
342 cloud water by for example enzymatic degradation of proteins, as reported by the authors. These hypotheses are also  
343 supported by elevated concentrations of Asp at the end of the campaign that is a biologically produced AA. Globally, the  
344 concentrations and major groups of AAs reported by Triesch et al. (2021) agree with the present work. This can be  
345 explained by the remoteness of both locations and also the relevant marine influence encountered at PUY (Renard et al.,  
346 2020).

347 In fog waters, at Davis in Northern California, Zhang et al. (2003b) measured elevated concentration of TCAAs with a  
348 mean concentration of 20  $\mu\text{M}$ . This is probably due to the proximity of the sampling site collection with local emission  
349 of aerosol particles in this rural environment; however, dominant AAs are the same (Ser, Gly, Ala, Asn and Leu/I). Two  
350 other studies in rainwater display similar AA concentrations and concentration ranking (Yan et al., 2015; Xu et al., 2019).  
351 The study in Korea measured lower AA concentrations (Free and combined AAs) at Seoul (an inland urban area) than  
352 those at Uljin (a coastal rural area) attributed to differences in contributing sources (Yan et al., 2015). A similar work has  
353 been performed at a suburban site in Guiyang (China) over one year and have shown a seasonal effect with a maximum  
354 level of AAs (Free and combined AAs) at spring and a minimal one at winter (Xue et al., 2019).

355 To conclude, few studies presented above report concentrations of AAs in cloud and fog waters. This is a challenging  
356 issue to compare those 3 studies that have been performed for contrasted environmental conditions and for a limited  
357 number of samples.



358 **Table 2. FAA concentrations in atmospheric aqueous samples: cloud, fog, and rain (n is relative to the number of the samples).**

Localization	Environment/ medium	Period / Samples	Separation/ Detection Method	Concentrations of FAAs (Range and mean values)	Distribution Major FAAs	Reference
<b>Puy de Dôme Mountain, France (1465 m)</b>	Rural + marine influence (Cloud)	03/2014 05-10/2018 09-10/2019 03-07/2020 13 samples	HPLC-MS/MS  Standard addition	Range: 39 - 244 ng m <sup>-3</sup>	Ser > Gly > Ala > Asn > Leu/I	(This work)
<b>Puy de Dôme Mountain, France (1465 m)</b>	Rural + marine influence (Cloud)	03-04/2014 (spring) 11/2014 (winter) 25 samples	HPLC- Fluorescence  OPA- Derivatization	Mean: 118.6 ± 97.6 ng m <sup>-3</sup>	Trp > Leu/I > Phe > Ser	(Bianco et al., 2016b)
<b>Cabo Verde islands (744 m)</b>	Marine (Cloud)	09-10/2017 (winter) 10 samples	HPLC-MS  OPA- Derivatization	Range: 11.2 - 489.9 ng m <sup>-3</sup>	Ser > Asp > Ala > Gly > Thr	(Triesch et al., 2021)
<b>Northern California Davis, US</b>	Rural (Fog)	1997 - 1999 (winter) 11 samples	HPLC- Fluorescence  OPA- Derivatization	Mean: 40.8 ± 38.0 ng m <sup>-3</sup> (FAAs, protein type)	Ser > Gly > Leu > Ala > Val	(Zhang and Anastasio, 2003b)
<b>Atlantic Ocean, Golf Mexico (Cruise)</b>	Marine (Rain)	09-10/1985 (n = 3) 02, 06, 09/1986 (n = 4) 7 samples	HPLC  OPA/NAC- Derivatization	Mean: 604 ± 585 µg L <sup>-1</sup>	Gly > Ser > Ala > acidic AAs	(Mopper and Zika, 1987)
<b>Seoul, South Korea (17 m)</b>	Urban (Rain)	03/2012 - 04/2014 36 samples	HPLC  OPA- Derivatization	Mean: 21.0 ± 17.9 µg L <sup>-1</sup>	THAA:  Gly > Glu, Ala, Asp, Ser	(Yan et al., 2015)
<b>Uljin, South Korea (30 m)</b>	Marine (Rain)	02/2011 - 01/2012 31 samples	HPLC  OPA- Derivatization	Mean: 100.9 ± 110.2 µg L <sup>-1</sup>		
<b>Guiyang, China (1300 m)</b>	Suburban (Rain)	05/2017 – 04/2018 Summer (n = 29) Autumn (n = 9) Winter (n = 14) Spring (n = 13) 65 samples	HPLC  OPA- Derivatization	Total range: 1.1 - 10.1 µM Mean: 3.7 µM Summer range: 1.3 - 6.6 µM Mean Summer: 2.9 µM Autumn range: 1.1 - 8.8 µM Mean Autumn: 4.4 µM Winter range: 1.5 - 9.9 µM Mean Winter: 3.4 µM Spring range: 2.6 - 10.1 µM Mean Spring: 5.2 µM	THAA:  Glu + Gln, Gly, Pro > Asp, Ala	(Xu et al., 2019)

- 359 • HPLC: High performance liquid chromatography  
 360 • OPA: Ortho-phthalaldehyde  
 361 • THAA: Total hydrolyzable amino acids

### 362 3.5 Analysis of the environmental variability

363 A large variability of the AA concentrations and relative proportions in the 13 cloud samples from PUY is observed  
 364 (Table S1). To better understand this variability, data are analyzed in parallel with various environmental factors such as  
 365 the air masses history and quantitative physical, chemical, and biological measurements. During their atmospheric  
 366 transports, the air masses received chemical species under various forms and from various sources, and could also undergo



367 multiphasic chemical transformations, as well as deposition. This section is devoted to the correlation between the AA  
 368 concentrations and the air mass history. To this end, PLS regressions are performed, and the results are validated with  
 369 nonparametric tests (Mann–Whitney tests).

370 The PLS matrix of the explanatory variables (the “Xs”) is composed of the “zone matrix” (Sea/Continental surface  
 371 </> ABLH) from the CAT model, to which are added the temperature, the pH, the inorganic ion concentrations, the  
 372 bacteria density, the ATP concentration, and the seasons (Table S1). The matrix of the dependent variables (the “Ys”) is  
 373 composed of the AA concentrations.

374 **Table 3. PLS correlation matrix between AA concentrations and “zone matrix” (Sea/Continental surface </> ABLH) from the**  
 375 **CAT model, and temperature, pH, cation and anion concentrations, TOC and H<sub>2</sub>O<sub>2</sub> concentrations, bacteria density (CFU/mL)**  
 376 **and ATP concentration, and the seasons (Fall/Winter and Spring/Summer) determined from 13 cloud sampled at PUY. Highest**  
 377 **correlations are displayed in dark red and highest anti-correlations in dark blue.**

Variables	Ser	Gly	Ala	Asn	Leu/I	Thr	Asp	Pro	Glu	His	Phe	Tyr	Lys	Gln	Arg	Trp	Met
Sea surface (< ABLH)	0.88	0.68	0.18	0.38	0.76	0.82	0.74	0.31	0.53	0.70	0.84	0.71	0.58	0.21	0.20	0.08	0.08
Sea surface (> ABLH)	-0.57	-0.45	0.03	-0.14	-0.50	-0.52	-0.42	-0.28	-0.04	-0.35	-0.37	-0.58	-0.11	0.16	0.31	0.38	0.25
Continental surface (< ABLH)	0.54	0.33	0.45	0.26	0.57	0.63	0.55	0.63	0.33	0.56	0.22	0.68	0.18	0.17	-0.48	-0.06	0.00
Continental surface (> ABLH)	-0.31	-0.21	-0.26	-0.23	-0.27	-0.32	-0.33	-0.11	-0.48	-0.36	-0.40	-0.18	-0.43	-0.36	-0.38	-0.41	-0.31
Temperature (°C)	-0.08	0.29	-0.14	0.29	-0.10	-0.15	-0.07	-0.25	-0.41	-0.28	0.40	-0.13	-0.36	0.51	-0.13	0.09	-0.20
pH	-0.10	0.00	-0.13	-0.19	0.02	-0.25	-0.11	-0.23	0.40	0.06	-0.32	-0.03	0.22	0.09	0.34	0.61	0.67
Na <sup>+</sup> (μM)	0.77	0.93	0.17	-0.07	0.62	0.44	0.65	0.03	0.07	0.46	0.80	0.54	0.43	0.02	0.03	0.12	-0.01
NH <sub>4</sub> <sup>+</sup> (μM)	-0.38	-0.03	-0.13	-0.10	-0.21	-0.43	-0.17	-0.36	-0.13	-0.10	-0.11	-0.16	-0.21	0.03	-0.22	-0.05	0.13
Mg <sup>2+</sup> (μM)	0.06	-0.11	0.09	0.10	0.17	0.11	0.15	0.23	0.46	0.41	-0.02	0.31	0.18	-0.11	-0.08	-0.08	0.30
K <sup>+</sup> (μM)	0.75	0.85	0.18	0.08	0.72	0.55	0.75	0.08	0.08	0.48	0.73	0.61	0.40	0.00	-0.24	-0.16	-0.16
Ca <sup>2+</sup> (μM)	-0.08	-0.34	0.04	0.30	-0.08	0.20	0.00	0.38	-0.05	0.09	-0.04	0.11	-0.28	-0.08	-0.39	-0.53	-0.41
SO <sub>4</sub> <sup>2-</sup> (μM)	-0.11	0.12	0.37	-0.29	0.13	-0.11	0.12	-0.01	0.41	0.18	-0.15	0.08	0.13	0.10	-0.13	0.21	0.58
NO <sub>3</sub> <sup>-</sup> (μM)	-0.17	0.00	-0.05	0.02	-0.01	-0.03	0.02	-0.12	0.02	-0.07	0.01	-0.05	-0.21	0.07	-0.26	-0.39	-0.05
Cl <sup>-</sup> (μM)	0.67	0.76	0.01	-0.05	0.60	0.34	0.50	-0.18	0.34	0.44	0.63	0.44	0.65	0.03	0.40	0.28	0.38
TOC (mgC L <sup>-1</sup> )	0.03	0.09	-0.09	0.25	0.00	-0.10	-0.10	0.15	-0.35	-0.09	0.16	0.19	-0.26	0.04	-0.27	0.06	-0.16
H <sub>2</sub> O <sub>2</sub> (μM)	-0.19	-0.01	-0.44	-0.10	-0.26	-0.46	-0.40	-0.32	-0.26	-0.25	-0.24	-0.15	-0.16	0.00	0.21	0.43	0.20
Bacteria (17°C; CFU/mL)	0.28	0.07	0.50	0.20	0.41	0.51	0.41	0.33	0.43	0.44	0.21	0.36	0.36	0.11	-0.19	-0.39	0.19
ATP (nM)	0.13	0.58	0.00	-0.14	0.19	-0.04	0.22	-0.32	0.05	0.19	0.33	0.17	0.01	0.19	-0.16	0.14	0.16
Fall / Winter	0.44	0.40	0.46	-0.44	0.32	0.36	0.36	0.29	-0.15	0.15	0.21	0.25	0.21	-0.39	-0.23	-0.31	-0.25
Spring / Summer	-0.44	-0.40	-0.46	0.44	-0.32	-0.36	-0.36	-0.29	0.15	-0.15	-0.21	-0.25	-0.21	0.39	0.23	0.31	0.25

378

379 The correlation matrix of this PLS (Table 3) displays significant (anti-) correlations. First, 9 of the 18 AAs (Gly, His, Tyr,  
 380 Asp, Leu/I, Thr, Phe and Ser) are robustly correlated with Sea surface under the atmospheric boundary layer height  
 381 (< ABLH), with correlation coefficients (r) ranging from 0.68 to 0.88. These 9 AAs are also significantly anticorrelated  
 382 with Sea surface in free atmosphere (> ABLH) (r ranging from -0.35 to -0.58), confirming direct influences from the  
 383 boundary layer. These 9 AAs coherently correlate with Na<sup>+</sup>, Cl<sup>-</sup> and K<sup>+</sup> concentrations, confirming a marine influence for  
 384 those AAs, similar to the observations of Triesch et al. (2021).

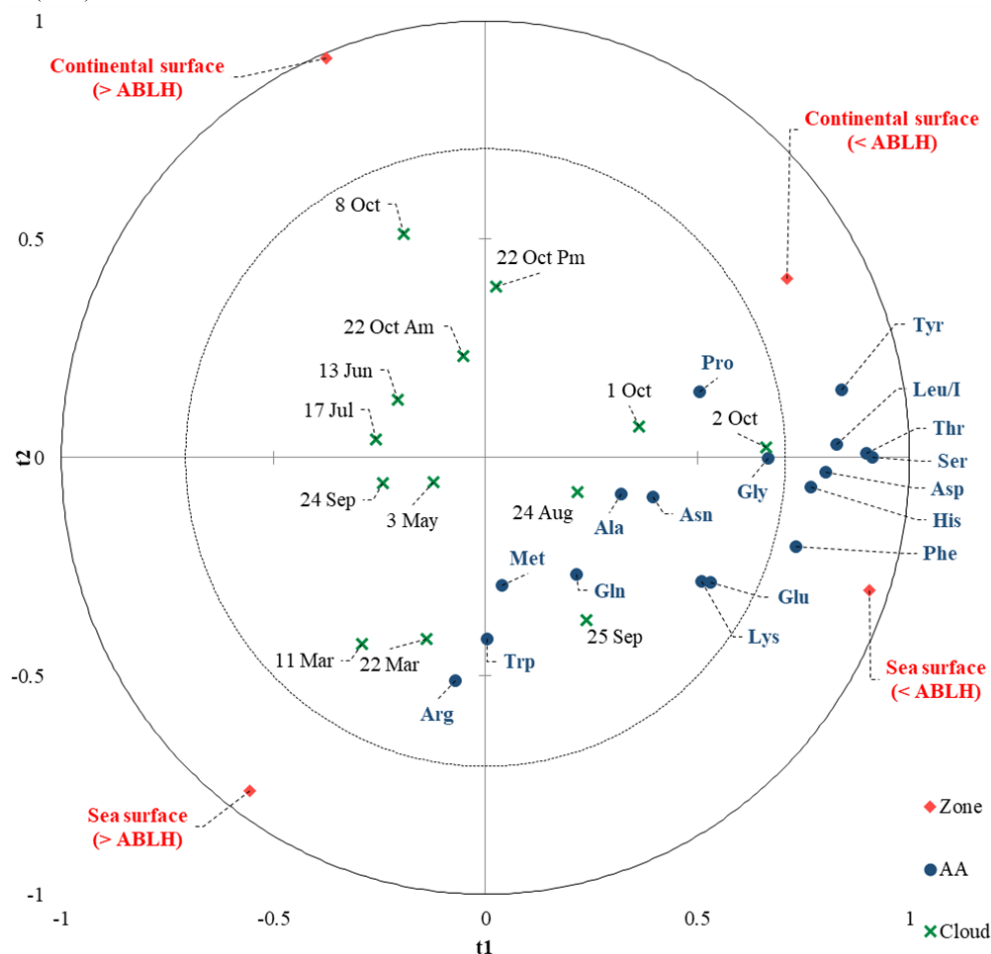
385 To a lesser extent, the same tendency (correlation / anticorrelation) is observed with Continental surface  
 386 (< ABLH / > ABLH). PUY is a remoteness site and the presence of anthropic ions, such as NO<sub>3</sub><sup>-</sup> and NH<sub>4</sub><sup>+</sup>, are correlated  
 387 with Continental surface (> ABLH) (Renard et al. 2020). Thereby, the AAs and, in particular, the 9 aforementioned AAs  
 388 are slightly anticorrelated with these anthropic ions.

389 No correlation appears between TOC concentration and the most abundant AAs, confirming the boundary layer influence,  
 390 as well as the variability of AA proportion in organic carbon. These 9 AAs are also slightly anti-correlated with H<sub>2</sub>O<sub>2</sub>  
 391 concentration, suggesting a potential influence of the photochemistry. The biological parameters, in particular the bacteria



392 density, are overall significantly correlated with the AA concentrations. The 9 AAs most correlated with the Sea surface  
393 (< ABLH) are, to a lesser extent, also correlated with Fall/Winter.

394 To go further in modelling the environmental variability of the AA concentrations in our cloud samples, we performed a  
395 simplified PLS restricting the Xs to the most correlated parameters (the PLS variable importance in the projection), *i.e.*,  
396 the zone matrix. The index of the predictive quality of the models obtained with this PLS ( $Q^2 = 0.19$  with one component)  
397 is satisfactory given the complexity of the cloud composition. Figure 5 displays PLS correlation chart with t component  
398 on axes t1 and t2. The main axis (t1) is linked to the ABLH and most of the AAs are correlated with < ABLH. The t2 axis  
399 is linked to the zone (Sea / Continental surfaces) and it reveals a preponderance of marine influence, which is consistent  
400 with the dominant western oceanic influence at PUY. Cloud water is a complex matrix resulting from the interaction of  
401 many factors; cloud samples from more continental zones (northeast) could influence this model. Nevertheless, it appears  
402 that the air mass history remains the prevailing parameter, after considering more cloud events, as observed in Renard et  
403 al. (2020).



404

405 **Figure 5.** Partial least squares (PLS) chart with t component on axes t1 and t2. The correlations map superimposes the  
406 dependent variables from the chemical matrix (blue circles), the explanatory variables (red diamonds) and the cloud events  
407 (green cross).





408 The following section is devoted to the analysis of the processes occurring in the atmosphere that could potentially explain  
409 the AA levels and distributions in the clouds sampled at PUY. These processes are linked to their sources and to their  
410 potential biotic and/or abiotic transformations in the atmosphere.

#### 411 **4. Discussion**

412 Results reveal high variability in the relative concentrations of FAAs among cloud samples; however, some major FAAs  
413 could be detected following this relative concentration ranking: Ser > Gly > Ala > Asn > Leu/I. On the contrary Trp and  
414 Met present very low concentrations.

##### 415 **4.1 Potential influence of the initial AAs distribution in biological matrices**

416 As free AAs are mostly from biological origin, we first compared the AA composition of various biological  
417 macromolecules (proteins, peptidoglycans, *etc.*) that can be the source of AAs after hydrolysis with the relative  
418 concentrations of AAs measured in the studied cloud samples.

419 A study reports the relative distributions of AAs in proteins extracted from bacterial cells (*S. aureus* and *E. coli*) and from  
420 various mammalian cells (Okayasu et al., 1997). Although there are some differences between mammalian and bacterial  
421 protein composition, some AAs are clearly dominant. As a first approximation, we can share AAs in two groups: Ala,  
422 Gly, Asp+Asn, Glu+Gln, Val, Leu/I, Lys are more abundant compared to others (Ser, His, Arg, Pro, Tyr, Met, Cys, Phe).  
423 Globally, this relative abundance of AAs initially constituting proteins presents similarities with the relative  
424 concentrations present in our samples. In particular, Ala, Gly, Leu/I and Asn/Asp are the most abundant in our samples  
425 as in the proteins. It is also true for Met which abundance is low. However, Ser proportion is clearly less dominant in  
426 proteins in comparison with the proportion encountered in the studied cloud samples. Ser is also a dominant AA in other  
427 atmospheric samples (Table 3), revealing another sources or processes.

428 We looked at the composition of peptidoglycan that form the cell wall of microorganisms that can be an atmospheric  
429 source of AAs. Peptidoglycan monomers consist in two joint amino sugars (Nag, Nam) connected with a pentapeptide  
430 containing the L-Ala, D-Gly L-Lys, D-Ala, D-Ala sequence. Peptidoglycans can thus represent a major source of Ala and  
431 Gly, that are the major AAs detected in our samples.

432 Finally, we searched for the potential origin of Ser. Hecky et al. (1973) reports the AA composition of cell walls from six  
433 different diatom species, selected on the basis of taxonomy and habitat diversity. Three are from estuarine origin, the  
434 other from fresh waters. The protein template of these cell walls is composed of the following AA sequences: Asp-Ser-  
435 Ser-Gly-Thr-Ser-Ser-Asp-Ser-Gly. Ser is thus highly abundant in these aquatic organisms and plays an important role in  
436 the complexation of the silicon ( $\text{Si}^{4+}$ ). AAs in sea water during phytoplankton blooms were also investigated (Ittekkot,  
437 1982): Glu concentration is maximum in the early stages of the bloom, while Asp, Gly, Ala and Lys concentrations  
438 increase at the end of the bloom. In parallel, Ser was one of the most abundant AA and its concentration remains high all  
439 along the bloom period. Ser could come from the cell walls of some phytoplankton species which are diatoms. Hashioka  
440 et al. showed that diatoms could contribute up to 80% of the total phytoplankton in the ocean to during bloom events  
441 (Hashioka et al., 2013). The high Ser concentration measured in our cloud samples could thus originate from diatoms and  
442 could be a marker of their oceanic origin.



443 In the following, we aim at discussing more the variability of the AAs distributions and concentrations among the samples  
444 looking at the air mass history (*i.e.*, sources) and their atmospheric transformations.

#### 445 **4.2 Potential influence of the air mass origin on the AA concentrations and their relative distribution**

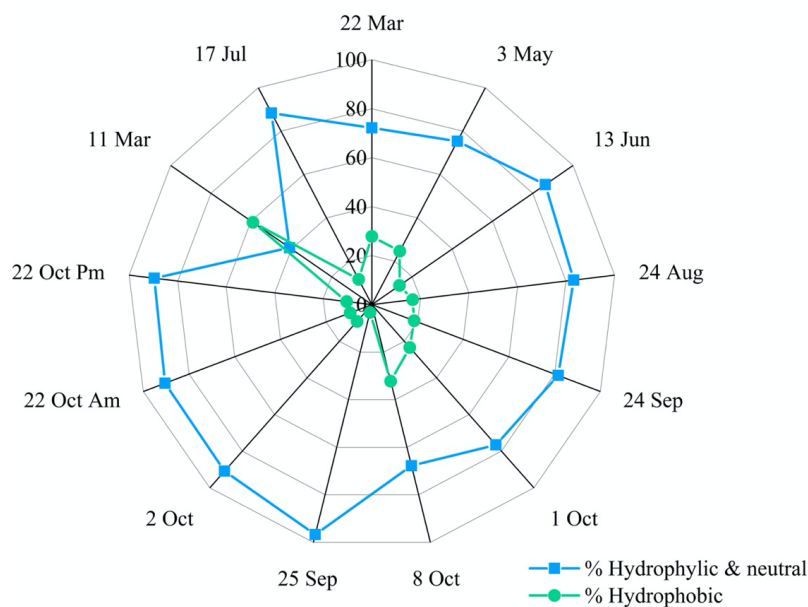
446 Table S4 summarizes the studies that analyze AAs quantity and distribution in various atmospheric media. Interestingly,  
447 the systematic presence of Ser, Ala and Gly is observed in the various atmospheric waters including clouds (Triesch et  
448 al., 2021), fogs (Zhang and Anastasio, 2003b) and rains (Mopper and Zika, 1987; Yan et al., 2015). These 3 AAs are also  
449 significantly present in aerosols over contrasted regions over the world: rural sites (Zhang and Anastasio, 2003b) marine  
450 sites (Matsumoto and Uematsu, 2005; Triesch et al., 2021; Violaki and Mihalopoulos, 2010; Wedyan and Preston, 2008),  
451 urban or suburban sites (Barbaro et al., 2011; Samy et al., 2013) and polar sites (Scalabrin et al., 2012).

452 Looking more specifically at only two AAs (Gly and Ala), this list of studies can be extended to other works: in rain (Xu  
453 et al., 2019), in marine aerosols (Mace et al., 2003b; Mandalakis et al., 2011), in rural aerosols (Ruiz-Jimenez et al., 2021;  
454 Samy et al., 2011), in polar and remote sites (Barbaro et al., 2020; Barbaro et al., 2015; Feltracco et al., 2019). We can  
455 notice that Gly is globally one of the major FAAs in all the reported studies (see Table S4 and joint explanations). In the  
456 present study, we detect significant concentrations of Leu/I in agreement with only 3 other studies (Bianco et al., 2016b;  
457 Mashayekhy Rad et al., 2019; Wedyan and Preston, 2008).

458 We overall found the same major groups of AAs that are commonly detected in marine clouds and aerosols. However,  
459 one of the main differences is the high concentration of Asn in two of our samples, instead of the more common Asp,  
460 suggesting potential conversion Asp/Asn (Jaber et al., 2021), and indicating that the origin of the clouds and aerosols is  
461 not the only main driving factor explaining the final observed FAAs relative proportion in the clouds sampled at PUY.  
462 Moreover, the presence of similar trends of AA composition in aerosols sampled under different sites (rural, marine,  
463 urban, and polar) and in our cloud samples shows various influences from both continental and marine sources.

464 In agreement with the results of the PLS analysis (Figure 5), a significant correlation ( $r = 0.78$ ) is observed between the  
465 TCAA and the time spent by the air mass over the sea and under the boundary layer (Sea surface < ABLH). The correlation  
466 between the TCAA and Sea and the Continental surfaces (< ABLH) is even higher ( $r = 0.86$ ) (Figure S5), indicating that  
467 the boundary layer influences the total amount of AAs rather than their relative concentration. When the air mass is  
468 transported in the free troposphere, the TCAA is lower, possibly because of the remoteness of the direct sources and  
469 because of chemical transformations that might be more intense in this upper part of the atmosphere.

470 To go further, Triesch et al. (2021) compared the AA composition of samples collected at Cabo Verde (marine  
471 environment) in both aerosol and cloud phases. They show that FAAs are partitioned according to their hydrophobic  
472 properties. They show that the hydrophobic AAs (Ala, Val, Phe, Leu/I) represent a much lower proportion (about 25 %)  
473 of the total AAs present in cloud water, compared to the neutral (Ser, Gly, Thr, Pro, Tyr) plus the hydrophilic AAs (Glu,  
474 Asp, Gaba). Figure 6 shows the distribution of the AAs in our samples collected at the PUY station according to their  
475 hydrophobic *versus* hydrophobic + neutral properties. Clearly the concentrations of hydrophilic (Glu, Asp, Gln, Asn, His,  
476 Lys, Arg) and neutral (Trp, Tyr, Gly, Thr, Ser, Pro) AAs are much higher (average value of 80 %) than that of hydrophobic  
477 (Leu/I, Phe, Met, Ala) ones in all the samples, except in the 11 Mar sample where the hydrophilic + neutral AAs represent  
478 only 40.8 % of the total FAAs. Our results are consistent with those measured in cloud samples at Cabo Verde; this  
479 suggests that the hydrophobic nature of AAs is less favorable for their incorporation in cloud droplets due to their low  
480 solubility.



481

482 **Figure 6. Relative composition of AAs grouped by hydrophathy (hydrophilic + neutral versus hydrophobic AAs) observed in**  
483 **each cloud sample.**

484 Although the initial AA composition of the emitted aerosols can greatly impact the type of FAAs, the aging of the samples  
485 due to biotic and abiotic processes must be considered to explain the presence of major or minor groups of AAs.

#### 486 4.3 Potential influence of the atmospheric aging of AAs

487 Table 4 reports calculated, and experimental lifetimes of the different AAs targeted in this work considering different  
488 biotic and abiotic processes.

489 First, AA theoretical lifetimes are calculated considering the reactivity constants of AAs with HO<sup>•</sup> radicals, O<sub>3</sub> and <sup>1</sup>O<sub>2</sub><sup>•</sup>.  
490 The values issued from the work of Jaber et al. (2021), Triesch et al. (2021) and Mc Gregor and Anastasio (2001) are  
491 reported in columns A, B and C, respectively. At first glance it can be noticed that the lifetimes depend on the AAs and  
492 can vary from a few hours or even minutes to a few days. Globally, reported values from the three studies are rather  
493 consistent although they were calculated using a different set of reactivity constants and different oxidant concentrations  
494 (see foot notes of Table 4). These AA theoretical lifetimes could explain the very low Met and Trp concentrations  
495 measured in our cloud samples which are very reactive, and the large one measured for Gly and Asn and in some extent  
496 for Ser and Ala which are very slowly transformed. However, they do not fit with the large amounts of Leu/I, except for  
497 the values given by Mc Gregor and Anastasio (2001).



498 **Table 4.** Estimated atmospheric lifetimes of AAs degraded by atmospheric biological and chemical processes. AAs are classified  
 499 following their mean concentrations measured in the present study. A brief description of the calculations is added below this  
 500 table. More information can be found in SI for the calculations performed in this study based on the work from Jaber et al.  
 501 (2021).

AAs	Theo. lifetimes by oxidation (days)	Theo. lifetimes by oxidation (days)	Theo. lifetimes by oxidation (days)	Exp. lifetimes by oxidation processes (days)	Exp. lifetimes by oxidation processes (days)	Exp. lifetimes by biological processes (days)
Reference	This study (adapted from Jaber et al. (2021))	Triesch et al. (2021)	Mc Gregor and Anastasio (2001)	This study (adapted from Jaber et al. (2021))	Mc Gregor and Anastasio (2001)	This study (adapted from Jaber et al. (2021))
Degradation processes	Oxidants: HO <sup>•</sup> , O <sub>3</sub> and <sup>1</sup> O <sub>2</sub> <sup>*</sup>	Oxidant: HO <sup>•</sup>	Oxidants: HO <sup>•</sup> , O <sub>3</sub> and <sup>1</sup> O <sub>2</sub> <sup>*</sup>	Irradiation experiments in artificial cloud medium	Irradiation experiments in fog waters	4 microbial strains in artificial cloud medium
Additional information	(A)	(B)	(C)	(D)	(E)	(F)
<b>Ser</b>	4.47	1.64	/	17.55	> 3.67	0.63 (15.1h)
<b>Gly</b>	41.26	3.09	> 170	\$	> 3.67	4.20
<b>Ala</b>	4.16	6.83	/	22.60	> 3.67	0.31 (7.6h)
<b>Asn</b>	22.05	/	/	\$	> 3.67	0.34 (8.1h)
<b>Leu/I</b>	0.64 (15.4h)	0.29	6.67	43.34	4.2	7.09
<b>Thr</b>	2.21	1.03	/	4.67	/	1.28
<b>Asp</b>	22.47	/	/	\$	2.42	1.55
<b>Pro</b>	1.72	1.70	/	\$	/	0.31 (7.4h)
<b>Glu</b>	5.49	3.29	37.5h	17.64	2.25	0.19 (4.5h)
<b>His</b>	0.10 (2.5h)	/	0.2 (5h)	22.60	1.00-1.83	1.79
<b>Phe</b>	0.17 (4.2h)	0.08	1.75	/	> 3.67	1.80
<b>Tyr</b>	0.08 (2.0h)	0.04	0.05 (1.2h)	3.56	1.25-2.33	0.86 (20.5h)
<b>Lys</b>	3.25	/	/	\$	> 3.67	0.59 (14.3h)
<b>Gln</b>	2.13	0.97	/	\$	/	0.20 (4.8h)
<b>Arg</b>	0.32 (7.7h)	/	/	\$	> 3.67	0.37 (9h)
<b>Trp</b>	0.06 (1.4h)	/	0.01 (0.15h)	10.51	0.11-0.38	6.97
<b>Met</b>	0.01 (0.13h)	0.06	0.01 (0.24h)	6.23	0.07-0.52	2.75

502 (A): Theoretical calculations considering kinetic rate constants for the AAs oxidation by HO<sup>•</sup>. O<sub>3</sub> and <sup>1</sup>O<sub>2</sub><sup>\*</sup> following Jaber et al. (2021). Aqueous  
 503 concentrations of HO<sup>•</sup>, O<sub>3</sub> and <sup>1</sup>O<sub>2</sub><sup>\*</sup> are respectively equal to 10<sup>-14</sup>, 5.0 10<sup>-10</sup> and 1.0 10<sup>-12</sup> M.

504 (B) Theoretical calculations by Triesch et al. (2021). The mean lifetimes are estimated by considering pH-dependent rate constant of AAs with HO<sup>•</sup>.  
 505 An HO<sup>•</sup> concentration of 2.2 10<sup>-14</sup> M is considered in this study.

506 (C) Theoretical calculations by Mc Gregor and Anastasio (2001) were done under typical midday, wintertime conditions. Several oxidants were  
 507 considered: the photoproduction of HO<sup>•</sup> and <sup>1</sup>O<sub>2</sub><sup>\*</sup> in the droplets, the source of HO<sup>•</sup> and O<sub>3</sub> in the aqueous phase by mass transfer.

508 (D) Experimental irradiation of 19 AAs at a concentration of 1 μM each in an artificial cloud medium were conducted. HO<sup>•</sup> production was performed  
 509 using Fe-Ethylenediamine-N,N'-disuccinic acid (EDDS) complex solution. HO<sup>•</sup> concentration of 8.3 10<sup>-13</sup> M is estimated. \$: Lifetimes cannot be  
 510 calculated since a production is observed during the experiment.

511 (E) Irradiation experiments using simulated sunlight illumination were performed on real fog waters containing AAs.

512 (F) Biodegradation experiments of 19 AAs were performed by Jaber et al. (2021) using 4 microbial strains (*Rhodococcus enclensis* PDD-23b-28,  
 513 *Pseudomonas graminis* PDD-13b-3, *Pseudomonas syringae* PDD-32b-74 and *Sphingomonas sp.* PDD-32b-11) in artificial cloud water.

514  
 515 A second approach is to consider transformation rate measurements to further calculate experimental lifetimes.  
 516 Experimental investigations were designed to evaluate both abiotic and biotic processes. Photodegradation experiments  
 517 have been designed to assess oxidations processes, the first one was performed by Jaber et al (2021) in a microcosm  
 518 mimicking cloud production environment using artificial cloud medium (Table 4, Column D), the second one (Mc Gregor and



519 Anastasio, 2001) consisted in irradiating real fog samples (Table 4, column E). In both cases the HO<sup>•</sup> concentration is  
520 quantified. Interestingly the obtained experimental lifetimes are globally longer than the theoretical ones, many of them  
521 exceeded 3 days and only Trp and Met lifetimes in the fog experiment are less than one hour. Moreover, certain AA  
522 lifetimes could not be calculated from transformation rates measured by Jaber et al. (2021) as they observed production  
523 and not degradation of some AAs (Gly, Asn, Asp, Pro, Phe, Lys, Arg, Gln). These experimental results, which are different  
524 from theoretical ones, reflect a much higher complexity of the occurring transformations. On the one hand, irradiations  
525 are performed on complex media containing a mixture of AAs, as well as other carbon and nitrogen sources. So, the  
526 measured transformation rates are net values reflecting both synthesis and degradation processes, and even potential inter-  
527 conversion mechanisms. On the other hand, theoretical lifetimes are calculated from reactivity constants measured in pure  
528 water containing a single AA without any other C or N components and thus far from the chemical reactivity in real  
529 environmental samples. In addition, it is difficult to interpret this data in more detail. Indeed, very few studies have studied  
530 the photo-produced compounds during these oxidation processes. Some works report the formation of carboxylic acids,  
531 nitrate and ammonia from AAs photo-transformations, or the conversion of AAs in another different AAs (His to Pro,  
532 Asp and Asn, Phe to Tyr, Pro to Glu) (see Jaber et al. (2021) for review). More detailed pathways of abiotic  
533 transformations are only available for Trp, Tyr and Phe (Bianco et al 2016a; Pattison et al., 2012). In spite of this complex  
534 situation, the long lifetimes or net production of Ser, Leu/I, Gly and Asn (Table 4, columns D and E) could explain the  
535 relative high concentrations of these compounds in cloud waters collected at PUY. On the contrary the short lifetimes  
536 (< 1 day) measured in fogs could explain the low concentrations of Met and Trp. However, the lifetimes reported here  
537 cannot fully explain intermediate concentration values measured for most of the other AAs; more work is needed to better  
538 understand oxidation pathways in complex atmospheric media and measure additional transformation rates.

539 Potential biological transformation processes have been also evaluated in the lab. Recent work suggests that  
540 biodegradation of AAs could occur in rain and in aerosols; these hypotheses are based on Degradation Index (DI)  
541 calculations (Xu et al., 2020; Zhu et al., 2021). To calculate biotransformation lifetimes, transformation rates were  
542 measured in microcosms with 4 bacterial strains isolated from clouds and representative of this medium and incubated in  
543 artificial cloud water (Jaber et al., 2021). As in the previous case of irradiation in the same microcosm, it was shown that  
544 some AAs could be degraded but also produced depending on the bacterial strain. The resulting biodegradation rates were  
545 thus calculated considering the proportion of each type of cell in real cloud (see Table S4 for more details). From these  
546 global reaction rates, lifetimes could be calculated for individual AA (Table 4, Column F). These biological lifetimes are  
547 very different from those obtained considering oxidation processes, and globally much shorter. *Per se* they cannot explain  
548 the ranking of the larger AAs (Ser, Ala, Leu/I, Asn) and lower AAs (Met, Trp) concentrations in our cloud samples,  
549 suggesting they might not be the major contribution to the transformation of these AAs. However, when other compounds  
550 are considered with rather low concentrations such as Gln, Arg, Lys, Phe, His, experimental oxidation lifetimes are long  
551 while biodegradation lifetimes are much shorter, combination of these two processes could reflect a more realistic  
552 situation. Biosynthesis and biodegradation pathways are very complex and interconnected and are fully described in  
553 databases (see: <https://www.genome.jp/kegg/pathway.html>). The complexity comes from the behavior of the different  
554 microorganisms to use these pathways. Up to now the only biodegradation rates related to atmospheric waters are from  
555 Jaber et al. (2021) and might be incomplete; more experimental work should be conducted on real atmospheric samples.



556 **Conclusion**

557 This study reports the quantification of amino acids in cloud waters sampled at the puy de Dôme station using a new  
558 approach based on a direct *in situ* analysis of the sample. Concentration of AAs represent in average nearly 2 % of the  
559 TOC with a significant variability of TCAAs among the different samples. This heterogeneity is also observed in the AAs  
560 distribution between the samples, but certain AAs are more dominant, especially Ser, Gly, Ala, Asp and Leu/I. These AA  
561 relative proportions can be explained by the original biological matrices that emit AAs into the atmosphere, by the  
562 hydrophilicity of AAs that favours their incorporation in the cloud water and finally by their transformations during their  
563 transport into the atmosphere that modulate the total of AAs. At PUY, the residence of the air masses within the boundary  
564 layer, especially above the sea, seems also to surely increase the total amount of AAs in cloud water. Conversely, the AA  
565 concentrations seem to decrease when the photolysis conditions are more favourable (free troposphere or Spring / Summer  
566 period). In other words, the AAs concentration is modulated by the sources (mainly from the boundary layer) and the  
567 sinks associated with the photodegradation.

568 However, it is still hard to validate all the formulated hypotheses that have been proposed to explain the differences in  
569 the amounts and proportions of the various AAs found in our samples. This variability integrates many factors that are  
570 interconnected or decorrelated and that should be investigated in the future. Some future targeted works could be  
571 mentioned. First, this study is to our knowledge only the third one performed on cloud samples. More samples should be  
572 collected at different seasons and at other sites presenting contrasted environmental conditions. This is crucial to robustly  
573 evaluate atmospheric AA variability considering the effect of difference sources and atmospheric transport. Second, a  
574 major limitation encountered to interpret the impact of transformation processes on the final distribution of AAs in  
575 atmospheric samples lies on the lack of knowledge available in this field. Very few studies report the complex mechanisms  
576 of biotic and abiotic transformations of AAs under realistic atmospheric conditions. Photochemists and biologists should  
577 develop interdisciplinary work to describe these transformation pathways; this remains a challenging task.

578 **Author contributions.**

579 A.-M. D. & L. D. designed the project. A. B., M. B. & L. D. sampled the clouds at PUY. M. B., S. J. and M. L. conducted  
580 the analysis. J.-L. B. used the CAT model to calculate backward trajectories and the “matrix zone”. P. R. and F. R.  
581 performed the statistical analysis. P. R., F. R., A.-M. D. & L. D. wrote the paper.

582 **Competing interests.**

583 The authors declare that they have no conflict of interest.

584 **Acknowledgements.**

585 This work was funded by the French National Research Agency (ANR) in the framework of the ‘Investment for the  
586 Future’ program, ANR-17-MPGA-0013. S. Jaber is recipient of a grant from the Walid Joumblatt Foundation for  
587 University Studies (WJF), Beirut, Lebanon, and M. Brissy from Clermont Auvergne Metropole. CO-PDD is an  
588 instrumented site of the OPGC observatory and LaMP laboratory, supported by the Université Clermont Auvergne  
589 (UCA), by the Centre National de la Recherche Scientifique (CNRS-INSU), and by the Centre National d’Etudes Spatiales  
590 (CNES). The authors are also very grateful for the financial support from the Fédération des Recherches en



591 Environnement through the CPER funded by Region Auvergne – Rhône-Alpes, the French ministry, ACTRIS Research  
592 Infrastructure, and FEDER European Regional funds. The authors also thank the I-Site CAP 20-25.  
593



## 594 References

- 595 Abe, R.Y., Akutsu, Y., Kagemoto, H.: Protein amino acids as markers for biological sources in urban aerosols, *Environ.*  
596 *Chem. Lett.*, 14, 155-161, 2016.
- 597 Addinsoft: XLSTAT Statistical and Data Analysis Solution. New York, NY, USA. Available online:  
598 <https://www.xlstat.com> (accessed on 23 May 2020). 2020.
- 599 Amato, P., Besaury, L., Joly, M., Penaud, B., Deguillaume, L., Delort, A.-M.: Metatranscriptomic exploration of  
600 microbial functioning in clouds, *Sci. Rep.*, 9, 4383, 2019.
- 601 Baray, J.L., Deguillaume, L., Colomb, A., Sellegri, K., Freney, E., Rose, C., Van Baelen, J., Pichon, J.M., Picard, D.,  
602 Fréville, P., Bouvier, L., Ribeiro, M., Amato, P., Banson, S., Bianco, A., Borbon, A., Bourcier, L., Bras, Y., Brigante,  
603 M., Cacault, P., Chauvigné, A., Charbouillot, T., Chaumerliac, N., Delort, A.M., Delmotte, M., Dupuy, R., Farah, A.,  
604 Febvre, G., Flossmann, A., Goubeyre, C., Hervier, C., Hervo, M., Huret, N., Joly, M., Kazan, V., Lopez, M., Mailhot,  
605 G., Marinoni, A., Masson, O., Montoux, N., Parazols, M., Peyrin, F., Pointin, Y., Ramonet, M., Rocco, M., Sancelme,  
606 M., Sauvage, S., Schmidt, M., Tison, E., Vaïtilingom, M., Villani, P., Wang, M., Yver-Kwok, C., Laj, P.: Cézeaux-  
607 Aulnat-Opme-Puy De Dôme: a multi-site for the long-term survey of the tropospheric composition and climate change,  
608 *Atmos. Meas. Tech.*, 13, 3413-3445, 2020.
- 609 Barbaro, E., Morabito, E., Gregoris, E., Feltracco, M., Gabrieli, J., Vardè, M., Cairns, W.R.L., Dallo, F., De Blasi, F.,  
610 Zangrando, R., Barbante, C., Gambaro, A.: Col Margherita Observatory: A background site in the Eastern Italian Alps  
611 for investigating the chemical composition of atmospheric aerosols, *Atmos. Environ.*, 221, 117071, 2020.
- 612 Barbaro, E., Zangrando, R., Moret, I., Barbante, C., Cescon, P., Gambaro, A.: Free amino acids in atmospheric particulate  
613 matter of Venice, Italy, *Atmos. Environ.*, 45, 5050-5057, 2011.
- 614 Barbaro, E., Zangrando, R., Vecchiato, M., Piazza, R., Cairns, W.R.L., Capodaglio, G., Barbante, C., Gambaro, A.: Free  
615 amino acids in Antarctic aerosol: potential markers for the evolution and fate of marine aerosol, *Atmos. Chem. Phys.*,  
616 15, 5457-5469, 2015.
- 617 Berger, P., Karpel Vel Leitner, N., Doré, M., Legube, B.: Ozone and hydroxyl radicals induced oxidation of glycine,  
618 *Water Res.*, 33, 433-441, 1999.
- 619 Berto, S., De Laurentiis, E., Tota, T., Chiavazza, E., Daniele, P.G., Minella, M., Isaia, M., Brigante, M., Vione, D.:  
620 Properties of the humic-like material arising from the photo-transformation of l-tyrosine, *Sci. of The Total Environ.*,  
621 545-546, 434-444, 2016.
- 622 Bianco, A., Deguillaume, L., Chaumerliac, N., Vaïtilingom, M., Wang, M., Delort, A.-M., Bridoux, M.C.: Effect of  
623 endogenous microbiota on the molecular composition of cloud water: a study by Fourier-transform ion cyclotron  
624 resonance mass spectrometry (FT-ICR MS), *Sci. Rep.*, 9, 7663, 2019.
- 625 Bianco, A., Deguillaume, L., Vaïtilingom, M., Nicol, E., Baray, J.-L., Chaumerliac, N., Bridoux, M.: Molecular  
626 characterization of cloud water samples collected at the puy de Dôme (France) by Fourier Transform Ion Cyclotron  
627 Resonance Mass Spectrometry, *Environ. Sci. & Technol.*, 52, 10275-10285, 2018.
- 628 Bianco, A., Passananti, M., Deguillaume, L., Mailhot, G., Brigante, M.: Tryptophan and tryptophan-like substances in  
629 cloud water: Occurrence and photochemical fate, *Atmos. Environ.*, 137, 53-61, 2016a.
- 630 Bianco, A., Voyard, G., Deguillaume, L., Mailhot, G., Brigante, M.: Improving the characterization of dissolved organic  
631 carbon in cloud water: Amino acids and their impact on the oxidant capacity, *Sci. Rep.*, 6, 37420, 2016b.
- 632 Broekaert, J.A.C.: Daniel C. Harris: Quantitative chemical analysis, 9th ed, *Analytical and Bioanalytical Chem.*, 407,  
633 8943-8944, 2015.
- 634 Cape, J.N., Cornell, S.E., Jickells, T.D., Nemitz, E.: Organic nitrogen in the atmosphere — Where does it come from? A  
635 review of sources and methods, *Atmos. Res.*, 102, 30-48, 2011.
- 636 Chan, M.N., Choi, M.Y., Ng, N.L., Chan, C.K.: Hygroscopicity of water-soluble organic compounds in atmospheric  
637 aerosols: Amino acids and biomass burning derived organic species, *Environ. Sci. Tech.*, 39, 1555-1562, 2005.
- 638 Cook, R.D., Lin, Y.H., Peng, Z., Boone, E., Chu, R.K., Dukett, J.E., Gunsch, M.J., Zhang, W., Tolic, N., Laskin, A., Pratt,  
639 K.A.: Biogenic, urban, and wildfire influences on the molecular composition of dissolved organic compounds in cloud  
640 water, *Atmos. Chem. Phys.*, 17, 15167-15180, 2017.
- 641 De Haan, D.O., Hawkins, L.N., Kononenko, J.A., Turley, J.J., Corrigan, A.L., Tolbert, M.A., Jimenez, J.L.: Formation  
642 of nitrogen-containing oligomers by methylglyoxal and amines in simulated evaporating cloud droplets, *Environ. Sci.*  
643 *& Technol.*, 45, 984-991, 2011.
- 644 Deguillaume, L., Charbouillot, T., Joly, M., Vaïtilingom, M., Parazols, M., Marinoni, A., Amato, P., Delort, A.M.,  
645 Vinatier, V., Flossmann, A., Chaumerliac, N., Pichon, J.M., Houdier, S., Laj, P., Sellegri, K., Colomb, A., Brigante,





- 646 M., Mailhot, G.: Classification of clouds sampled at the puy de Dôme (France) based on 10 yr of monitoring of their  
647 physicochemical properties, *Atmos. Chem. Phys.*, 14, 1485-1506, 2014.
- 648 Després, V.R., Huffman, J.A., Burrows, S.M., Hoose, C., Safatov, A.S., Buryak, G., Fröhlich-Nowoisky, J., Elbert, W.,  
649 Andreae, M.O., Pöschl, U., Jaenicke, R.: Primary biological aerosol particles in the atmosphere: a review, *Tellus B*,  
650 64, 2012.
- 651 Di Filippo, P., Pomata, D., Riccardi, C., Buiarelli, F., Gallo, V., Quaranta, A.: Free and combined amino acids in size-  
652 segregated atmospheric aerosol samples, *Atmos. Environ.*, 98, 179-189, 2014.
- 653 Feltracco, M., Barbaro, E., Kirchgeorg, T., Spolaor, A., Turetta, C., Zangrando, R., Barbante, C., Gambaro, A.: Free and  
654 combined L- and D-amino acids in Arctic aerosol, *Chemosphere*, 220, 412-421, 2019.
- 655 Fröhlich-Nowoisky, J., Kampf, C.J., Weber, B., Huffman, J.A., Pöhlker, C., Andreae, M.O., Lang-Yona, N., Burrows,  
656 S.M., Gunthe, S.S., Elbert, W., Su, H., Hoor, P., Thines, E., Hoffmann, T., Després, V.R., Pöschl, U.: Bioaerosols in  
657 the Earth system: Climate, health, and ecosystem interactions, *Atmos. Res.*, 182, 346-376, 2016.
- 658 Ge, P., Luo, G., Luo, Y., Huang, W., Xie, H., Chen, J., Qu, J.: Molecular understanding of the interaction of amino acids  
659 with sulfuric acid in the presence of water and the atmospheric implication, *Chemosphere*, 210, 215-223, 2018.
- 660 Gorzelska, K., Galloway, J.N., Watterson, K., Keene, W.C.: Water-soluble primary amine compounds in rural continental  
661 precipitation, *Atmos. Environ.*, 26, 1005-1018, 1992.
- 662 Hashioka, T., Vogt, M., Yamanaka, Y., Le Quéré, C., Buitenhuis, E.T., Aita, M.N., Alvain, S., Bopp, L., Hirata, T., Lima,  
663 I., Saille, S., Doney, S.C.: Phytoplankton competition during the spring bloom in four plankton functional type  
664 models, *Biogeosciences*, 10, 6833-6850, 2013.
- 665 Helin, A., Sietiö, O.M., Heinonsalo, J., Bäck, J., Riekkola, M.L., Parshintsev, J.: Characterization of free amino acids,  
666 bacteria and fungi in size-segregated atmospheric aerosols in boreal forest: seasonal patterns, abundances and size  
667 distributions, *Atmos. Chem. Phys.*, 17, 13089-13101, 2017.
- 668 Hewavitharana, A.K., Abu Kassim, N.S., Shaw, P.N.: Standard addition with internal standardisation as an alternative to  
669 using stable isotope labelled internal standards to correct for matrix effects—Comparison and validation using liquid  
670 chromatography-tandem mass spectrometric assay of vitamin D, *J. of Chrom. A*, 1553, 101-107, 2018.
- 671 Ittekkot, V.: Variations of dissolved organic matter during a plankton bloom: qualitative aspects, based on sugar and  
672 amino acid analyses, *Mar. Chem.*, 11, 143-158, 1982.
- 673 Jaber, S., Joly, M., Brissy, M., Lereboure, M., Khaled, A., Ervens, B., Delort, A.M.: Biotic and abiotic transformation  
674 of amino acids in cloud water: Experimental studies and atmospheric implications, *Biogeosciences*, 18, 1067-1080,  
675 2021.
- 676 Koutny, M., Sancelme, M., Dabin, C., Pichon, N., Delort, A.-M., Lemaire, J.: Acquired biodegradability of polyethylenes  
677 containing pro-oxidant additives, *Polymer Degradation and Stability*, 91, 1495-1503, 2006.
- 678 Kristensson, A., Rosenørn, T., Bilde, M.: Cloud droplet activation of amino acid aerosol particles, *The J. of Phys.Chem.*  
679 *A*, 114, 379-386, 2010.
- 680 Li, X., Hede, T., Tu, Y., Leck, C., Ågren, H.: Cloud droplet activation mechanisms of amino acid aerosol particles: insight  
681 from molecular dynamics simulations, *Tellus B: Chemical and Physical Meteorology*, 65, 20476, 2013.
- 682 Mace, K.A., Duce, R.A., Tindale, N.W.: Organic nitrogen in rain and aerosol at Cape Grim, Tasmania, Australia, *J.*  
683 *Geophys. Res.: Atmospheres*, 108, 2003a.
- 684 Mace, K.A., Kubilay, N., Duce, R.A.: Organic nitrogen in rain and aerosol in the eastern Mediterranean atmosphere: An  
685 association with atmospheric dust, *J. Geophys. Res.: Atmospheres*, 108, 2003b.
- 686 Mandalakis, M., Apostolaki, M., Tziaras, T., Polymenakou, P., Stephanou, E.G.: Free and combined amino acids in  
687 marine background atmospheric aerosols over the Eastern Mediterranean, *Atmos. Environ.*, 45, 1003-1009, 2011.
- 688 Marion, A., Brigante, M., Mailhot, G.: A new source of ammonia and carboxylic acids in cloud water: The first evidence  
689 of photochemical process involving an iron-amino acid complex, *Atmos. Environ.*, 195, 179-186, 2018.
- 690 Mashayekhy Rad, F., Zurita, J., Gilles, P., Rutgeerts, L.A.J., Nilsson, U., Ilag, L.L., Leck, C.: Measurements of  
691 atmospheric proteinaceous aerosol in the Arctic using a selective UHPLC/ESI-MS/MS strategy, *Journal of The*  
692 *American Society for Mass Spectrometry*, 30, 161-173, 2019.
- 693 Matos, J.T.V., Duarte, R.M.B.O., Duarte, A.C.: Challenges in the identification and characterization of free amino acids  
694 and proteinaceous compounds in atmospheric aerosols: A critical review, *TrAC Trends in Analytical Chemistry*, 75,  
695 97-107, 2016.
- 696 Matsumoto, K., Uematsu, M.: Free amino acids in marine aerosols over the western North Pacific Ocean, *Atmos. Environ.*,  
697 39, 2163-2170, 2005.



- 698 McGregor, K.G., Anastasio, C.: Chemistry of fog waters in California's Central Valley: 2. Photochemical transformations  
699 of amino acids and alkyl amines, *Atmos. Environ.*, 35, 1091-1104, 2001.
- 700 Mopper, K., Zika, R.G.: Free amino acids in marine rains: evidence for oxidation and potential role in nitrogen cycling,  
701 *Nature*, 325, 246-249, 1987.
- 702 Okayasu, T., Ikeda, M., Akimoto, K., Sorimachi, K.: The amino acid composition of mammalian and bacterial cells,  
703 *Amino Acids*, 13, 379-391, 1997.
- 704 Pattison, D.I., Rahmanto, A.S., Davies, M.J.: Photo-oxidation of proteins, *Photochem. & Photobiol. Sci.*, 11, 38-53, 2012.
- 705 Pummer, B.G., Budke, C., Augustin-Bauditz, S., Niedermeier, D., Felgitsch, L., Kampf, C.J., Huber, R.G., Liedl, K.R.,  
706 Loerting, T., Moschen, T., Schauperl, M., Tollinger, M., Morris, C.E., Wex, H., Grothe, H., Pöschl, U., Koop, T.,  
707 Fröhlich-Nowoisky, J.: Ice nucleation by water-soluble macromolecules, *Atmos. Chem. Phys.*, 15, 4077-4091, 2015.
- 708 Ren, L., Bai, H., Yu, X., Wu, F., Yue, S., Ren, H., Li, L., Lai, S., Sun, Y., Wang, Z., Fu, P.: Molecular composition and  
709 seasonal variation of amino acids in urban aerosols from Beijing, China, *Atmos. Res.*, 203, 28-35, 2018.
- 710 Renard, P., Bianco, A., Baray, J.-L., Bridoux, M., Delort, A.-M., Deguillaume, L.: Classification of clouds sampled at the  
711 puy de Dôme station (France) based on chemical measurements and air mass history matrices, *Atmosphere*, 11, 732,  
712 2020.
- 713 Ruiz-Jimenez, J., Okuljar, M., Sietiö, O.M., Demaria, G., Liangsupree, T., Zagatti, E., Aalto, J., Hartonen, K., Heinonsalo,  
714 J., Bäck, J., Petäjä, T., Riekkola, M.L.: Determination of free amino acids, saccharides, and selected microbes in  
715 biogenic atmospheric aerosols – seasonal variations, particle size distribution, chemical and microbial relations,  
716 *Atmos. Chem. Phys.*, 21, 8775-8790, 2021.
- 717 Samy, S., Robinson, J., Hays, M.D.: An advanced LC-MS (Q-TOF) technique for the detection of amino acids in  
718 atmospheric aerosols, *Analytical and Bioanalytical Chem.*, 401, 3103-3113, 2011.
- 719 Samy, S., Robinson, J., Rumsey, I.C., Walker, J.T., Hays, M.D.: Speciation and trends of organic nitrogen in southeastern  
720 U.S. fine particulate matter (PM<sub>2.5</sub>), *J. Geophys. Res.: Atmospheres*, 118, 1996-2006, 2013.
- 721 Scalabrin, E., Zangrando, R., Barbaro, E., Kehrwald, N.M., Gabrieli, J., Barbante, C., Gambaro, A.: Amino acids in Arctic  
722 aerosols, *Atmos. Chem. Phys.*, 12, 10453-10463, 2012.
- 723 Song, T., Wang, S., Zhang, Y., Song, J., Liu, F., Fu, P., Shiraiwa, M., Xie, Z., Yue, D., Zhong, L., Zheng, J., Lai, S.:  
724 Proteins and amino acids in fine particulate matter in rural Guangzhou, Southern China: seasonal cycles, sources, and  
725 atmospheric processes, *Environ. Sci. & Technol.*, 51, 6773-6781, 2017.
- 726 Szymmer, W., Zawadzki, I.: Biogenic and anthropogenic sources of ice-forming nuclei: A review, *Bulletin of the American  
727 Meteorological Society*, 78, 209-228, 1997.
- 728 Triesch, N., van Pinxteren, M., Engel, A., Herrmann, H.: Concerted measurements of free amino acids at the Cape Verde  
729 Islands: High enrichments in submicron sea spray aerosol particles and cloud droplets, *Atmos. Chem. Phys.*, 21, 163-  
730 181, 2021.
- 731 Vařtilingom, M., Attard, E., Gaiani, N., Sancelme, M., Deguillaume, L., Flossmann, A.I., Amato, P., Delort, A.-M.: Long-  
732 term features of cloud microbiology at the puy de Dôme (France), *Atmos. Environ.*, 56, 88-100, 2012.
- 733 Violaki, K., Mihalopoulos, N.: Water-soluble organic nitrogen (WSON) in size-segregated atmospheric particles over the  
734 Eastern Mediterranean, *Atmos. Environ.*, 44, 4339-4345, 2010.
- 735 Wedyan, M.A., Preston, M.R.: The coupling of surface seawater organic nitrogen and the marine aerosol as inferred from  
736 enantiomer-specific amino acid analysis, *Atmos. Environ.*, 42, 8698-8705, 2008.
- 737 Wirgot, N., Vinatier, V., Deguillaume, L., Sancelme, M., Delort, A.M.: H<sub>2</sub>O<sub>2</sub> modulates the energetic metabolism of the  
738 cloud microbiome, *Atmos. Chem. Phys.*, 17, 14841-14851, 2017.
- 739 Xu, Y., Wu, D., Xiao, H., Zhou, J.: Dissolved hydrolyzed amino acids in precipitation in suburban Guiyang, southwestern  
740 China: Seasonal variations and potential atmospheric processes, *Atmos. Environ.*, 211, 247-255, 2019.
- 741 Xu, Y., Xiao, H., Wu, D., Long, C.: Abiotic and biological degradation of atmospheric proteinaceous matter can  
742 contribute significantly to dissolved amino acids in wet deposition, *Environ. Sci. & Technol.*, 54, 6551-6561, 2020.
- 743 Yan, G., Kim, G., Kim, J., Jeong, Y.-S., Kim, Y.I.: Dissolved total hydrolyzable enantiomeric amino acids in precipitation:  
744 Implications on bacterial contributions to atmospheric organic matter, *Geochimica et Cosmochimica Acta*, 153, 1-14,  
745 2015.
- 746 Zhang, Q., Anastasio, C.: Conversion of fogwater and aerosol organic nitrogen to ammonium, nitrate, and NO<sub>x</sub> during  
747 exposure to simulated sunlight and ozone, *Environ. Sci. & Technol.*, 37, 3522-3530, 2003a.
- 748 Zhang, Q., Anastasio, C.: Free and combined amino compounds in atmospheric fine particles (PM<sub>2.5</sub>) and fog waters  
749 from Northern California, *Atmos. Environ.*, 37, 2247-2258, 2003b.



- 750 Zhao, Y., Hallar, A.G., Mazzoleni, L.R.: Atmospheric organic matter in clouds: exact masses and molecular formula  
751 identification using ultrahigh-resolution FT-ICR mass spectrometry, *Atmos. Chem. Phys.*, 13, 12343-12362, 2013.
- 752 Zhu, R.-g., Xiao, H.-Y., Zhu, Y., Wen, Z., Fang, X., Pan, Y.: Sources and transformation processes of proteinaceous  
753 matter and free amino acids in PM<sub>2.5</sub>, *J. Geophys. Res.: Atmospheres*, 125, e2020JD032375, 2020.
- 754 Zhu, R.G., Xiao, H.Y., Luo, L., Xiao, H., Wen, Z., Zhu, Y., Fang, X., Pan, Y., Chen, Z.: Measurement report: Hydrolyzed  
755 amino acids in fine and coarse atmospheric aerosol in Nanchang, China: concentrations, compositions, sources and  
756 possible bacterial degradation state, *Atmos. Chem. Phys.*, 21, 2585-2600, 2021.
- 757