From Microorganisms to Biosignatures: Subcritical Water Extraction as a Sample Preparation Technique for Future Life Detection Missions

Zuzana Cieslarova¹, Aaron C. Noell¹, Peter A. Willis¹, and Maria F. Mora¹

¹Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA

Abstract  Amino acids are important targets in the search for life beyond Earth. The types of amino acids, their relative abundances, and enantiomeric excess all serve as biosignatures. To detect such biosignatures, identification and quantification of individual amino acids are required. Regardless of the analytical workflow used, it is critical to release amino acids from the sample without altering their molecular distributions. Subcritical water extraction (SCWE) is a promising technique for the release of organics from a variety of samples that has been under development for flight implementation for over 20 years. However, there is a risk of racemization and/or degradation of amino acids during extraction that could alter their distributions. This work demonstrates for the first time that SCWE (200°C for 30 min) can release amino acids from bacterial cells and spores, while maintaining their native distributions and enantiomeric excess.

Plain Language Summary  Chemical analysis can help us find signs of life on other planets, if they are present. Biological samples contain chemical fingerprints that are not present in samples without life. Here, we demonstrate that by using a simple method similar to a pressure cooker, those chemical fingerprints can be released from microorganisms and detected without alteration.

1. Introduction

The discovery of potentially habitable conditions on Europa and Enceladus has led to an increased interest in searching for extant life in these locations (Chyba & Phillips, 2002; Porco et al., 2017). There is a consensus that multiple lines of evidence will be required to provide the extraordinary proof required to support any claims of the detection of extraterrestrial life (Hand et al., 2017). Among the possible lines of evidence, chemical biosignatures are considered key because of the way that life is expected to leave an imprint on chemical distributions, or by the presence of chemical complexity, that is difficult to explain via abiotic processes (Lovelock, 1965; McKay, 2004).

Amino acids are among the targets that could be used to determine the presence of life beyond Earth (Neveu et al., 2018). Amino acids are the core of terrestrial life as the building blocks of proteins (Aerts et al., 2020), but are also found in meteorites, which indicates they do not have an Earth-specific origin (Glavin et al., 2010; Pizzarello et al., 2012). For this reason, the detection of one or more of the typical “meteoritic/abiotic” amino acids is not indicative of life, although detecting multiple more complex members of this compound class would suggest biotic processes. Fortunately, instead of relying only on the detection of certain amino acids, we can use their relative distributions to establish their biotic or abiotic origin. The first relative distribution is related to chiral forms of amino acids and expressed by enantiomeric excess. Life on Earth uses mostly left-handed amino acids, while abiotic samples contain mixtures of both chiral forms. The second distribution is related to the overall complexity of amino acids, as expressed by the ratio of each amino acid to the simplest one, glycine. While abiotic processes tend to produce simpler, easier to synthesize amino acids like glycine, amino acids produced biotically tend to be more complex as a requirement for protein functionality. Therefore, the determination of both enantiomeric excesses and relative abundances of amino acids could provide strong evidence of biogenicity (Creamer et al., 2017; McKay, 2004).

A key challenge for amino acid analysis is their release from cells or proteins present in the sample. The gold standard method to achieve this in terrestrial laboratories is acid hydrolysis. While some alteration of amino acids can occur during this treatment (Kaiser & Benner, 2005; Pickering & Newton, 1990), it is overall a robust method for assessing amino acid distributions. However, implementing an extraction that requires 6 N HCl heated up to 110°C for 24 hr involves considerable developmental and operational risks which would significantly complicate...
its flight implementation. Other challenges beyond the storage and handling of 6 N HCl, include the additional sample preparation steps required to clean up the sample following extraction.

An alternative to acid hydrolysis is subcritical water extraction (SCWE). SCWE has proven to be efficient in extracting a variety of compounds from biological and environmental samples (Gbashi et al., 2017). SCWE takes advantage of increased hydrolysis at high temperature and pressure (100–374°C). Under these conditions, the dissociation constant of water increases by 3 orders of magnitude (vs. ambient conditions), allowing it to act as an acid or base catalyst for hydrolysis, for example, the natural pH reaches as low as ∼4 before the supercritical point (Savage, 1999).

SCWE is promising for implementation on flight missions because of its ability to target a wide range of organics using only water as a solvent. Water is compatible with multiple materials used for flight hardware and it is preferred over concentrated acids or organic solvents. Previous studies on SCWE focused on astrobiology have described extraction conditions to liberate amino acids from soil (Amashukeli et al., 2007; Kehl et al., 2019; Noell et al., 2018). More recently, the use of SCWE with liquid samples was demonstrated by analyzing samples from Mono Lake, a hypersaline lake considered as an analog to ocean worlds (Ferreira Santos et al., 2018; Jaramillo et al., 2021). However, a dedicated study of the use of SCWE to break down microorganisms into their amino acid building blocks, and the effects of this treatment upon their distributions, has never been performed. There is a concern that during SCWE there may be a significant conversion of larger amino acids into simpler ones as well as racemization. Both of these processes would confound our ability to recognize biosignatures, should they be present. Therefore, here we report for the first time how SCWE affects the distribution and enantiomeric excess of amino acids extracted from cultures of *Escherichia coli* cells and *Bacillus subtilis* spores.

2. Materials and Methods

2.1. Chemicals and Cell Culture

β-Alanine, γ-aminobutyric acid (GABA), glycine, L- and D-alanine, -histidine, -leucine, -serine, -valine, sodium tetraborate, sodium taurocholate hydrate, γ-cyclodextrin, acetonitrile, dimethylformamide (DMF), methanol, and hydrochloric acid (HCl) were obtained from Sigma-Aldrich. 5-carboxyfluorescein succinimidyl ester (CFSE) and sodium hydroxide (NaOH) were purchased from Thermo Fisher Scientific. Amino acid solutions were prepared individually at 20 mM in ultrapure water and stored at 4°C. These solutions were diluted accordingly to prepare calibration curves to quantify the amino acids present in the extracts. CFSE was dissolved in DMF as 20 mM solution and kept at −20°C.

*E. coli* culture was grown from frozen glycerol stocks by reseeding them in Luria broth (LB) growth media at 37°C until a concentration of 10⁸ cells/mL was obtained. Before extraction, 5 mL of *E. coli* stock was centrifuged for 5 min at 5,000 rcf, the supernatant was removed, and 5 mL of water was added to the remaining pellet and vortexed to homogenize it. This procedure was repeated three times to remove soluble compounds from the growth media. The last supernatant obtained by this process was used as blank.

*B. subtilis* spores (ATCC 27370) were prepared by growing vegetative cells on tryptic soy agar plates at 30°C for 2 days. Vegetative cells were then transferred to plates with a sporulation medium composed of 1.6% nutrient broth, 1.7% agar, 0.2% KCl, 0.05% MgSO₄, 1 mM Cu(NO₃)₂, 100 mM MnCl₂·4H₂O, 1 mM FeSO₄, and 1% glucose, and incubated at 37°C for a total of 4 days. Sporulation was deemed complete when >90% of cells were spores according to phase-contrast microscopy. Spores were separated from residual vegetative cells and debris by cycles of suspension, sonicating, and centrifugation in the following solutions: biowater (Molecular Biology Grade, Fisher BioReagents), 1 M NaCl, 0.5 M NaCl + 1 M KCl, 0.2 mg/mL lysozyme and 10 mM Tris-HCl solution under constant shaking for 2 hr at 37°C, and another seven times with biowater until no more cell debris was observed. The final spore suspension had a concentration of 3 × 10⁹ spores/mL and was stored in the dark at 4°C until use. The suspension was diluted to a final concentration of 1 × 10⁷ spores/mL before extraction.

2.2. Subcritical Water Extraction

All extractions were performed in a microwave reactor (Monowave 450, Anton Paar). Sample vials made of borosilicate glass were sealed with PEEK caps over silicone gaskets. Extractions were performed at 200°C for 30 min (∼10 bar), in 4 mL of water or 10 mM HCl.
2.3. Capillary Electrophoresis With Laser-Induced Fluorescence

A SCIEX P/ACE MDQ capillary electrophoresis instrument with laser-induced fluorescence detection was used to analyze the extracts. Fused silica capillaries from Polymicro Technologies with a length of 50 cm (50 μm I.D. × 360 μm O.D.) and the detection located 10 cm from the end of the capillary were used. New capillaries were rinsed with methanol, water, 0.5 M HCl, water, and 0.1 M NaOH for 4 min, followed by 5 min rinse with water and 10 min rinse with the background electrolyte (BGE). BGE was composed of 80 mM sodium tetraborate, 30 mM sodium taurocholate, 30 mM γ-cyclodextrin and 6% v/v acetonitrile. Between runs, the capillary was rinsed with HCl, NaOH, water, and BGE for 3 min. Samples were injected by applying 0.5 psi for 4 s. All separations were performed using +25 kV (500 V/m).

Derivatization was performed as previously described (Creamer et al., 2017). Briefly, samples were diluted with sodium tetraborate (final concentration 50 mM) and 2 μL of 2 mM CFSE in DMF was added. The samples were left to react for 2 hr in the dark before analysis. Concentrations of amino acids were determined by using external calibration (peak area vs. concentration).

3. Results

To study the effect of SCWE on the biosignatures of homochirality and relative ratios to glycine (Creamer et al., 2017; Davila & McKay, 2014; Dorn et al., 2011), two organisms were selected: cells of Gram-negative E. coli and spores of B. subtilis. There is a large body of literature on SCWE of peptides and proteins (Esteban et al., 2008; Hawthorne et al., 1994). In the context of astrobiology, SCWE has been used for the extraction of amino acids from Mars analog samples (Amashukeli et al., 2007; Kehl et al., 2019; Mora et al., 2020; Noell et al., 2018). These studies focused on optimizing temperature and time, while maximizing yield. Optimum amino acid yields were reported at 200°C in the presence of acid. In addition, Amashukeli et al. (2007) and Noell et al. (2018) compared the yields of amino acids extracted via SCWE to the standard HCl vapor-phase and liquid-phase hydrolysis yields. Extraction efficiencies ranged from ~20% to 60% depending on temperature, time, and the specific amino acid.

Based on these previous studies, 30 min and 200°C were selected as extraction time and temperature for all the experiments described here. Ex extractions were performed in water and 10 mM HCl, and the amino acid content was determined by CE-LIF. Overall, it was observed that the addition of acid increases the amount of amino acids released, as shown in Figure 1.
In addition, enantiomeric excesses (ee) of amino acids were calculated (Table 1) according to Equation 1:

\[
\% \text{ ee} = \left( \frac{L - D}{L + D} \right) \times 100
\]

where \( L \) and \( D \) are the concentrations of L-amino and D-amino acid, respectively. It is worth noting that nature uses mostly L-amino acids, therefore, we would expect to obtain 100% enantiomeric excesses for all measured amino acids, with the exception of Ala and Ser, due to the fact that their D-forms are present in bacteria either as components of cell walls (D-Ala) or during stationary phase (D-Ser). Other deviations from 100% ee are likely due to racemization during SCWE.

Finally, the relative abundances of amino acids to glycine were determined. Figure 2a shows relative abundances of amino acids found in \( E. \) coli cells and \( B. \) subtilis spores extracts next to values reported in the literature for \( E. \) coli, as a biotic reference, and the CM2 meteorite, as an abiotic reference (Akira et al., 1985; Booth, 1987; Cronin & Pizzarello, 1983; Nishikawa & Ooi, 1982; Okayasu et al., 1997).

Because the concentration of glycine obtained here would be higher than the native value, due to the degradation of amino acids into glycine during SCWE (Abdelmoez et al., 2010), the relative abundances with respect to total amino acid content excluding glycine (sum of His, Leu, Val, Ser, and Ala) were also calculated and plotted next to the biotic and abiotic references.

To quantify how much a measured biosignature overlaps with the biotic and abiotic references, these distributions were treated as vectors, and cosine similarities between them were calculated according to Equation 2:

\[
\text{Cosine similarity} = \text{Sc}(A, B) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2} \sqrt{\sum_{i=1}^{n} B_i^2}}
\]

where \( A \) is the vector corresponding to the measured abundances ([AA]/[Gly] or [AA]/[Total measured AA−Gly]) and \( B \) is the vector corresponding to the reference (biotic or abiotic). Cosine similarity determines the angle between each vector; the closer the value is to 1, the more similar the vectors are. While not a formal distance metric, it is a useful way to identify similar “vectors” where the absolute magnitude is not informative. Figure 3 shows the cosine similarities for relative abundances for each organism and SCWE condition in relation to the biotic and abiotic references.

### 4. Discussion

Two organisms were selected for this study: \( E. \) coli cells and \( B. \) subtilis spores. \( E. \) coli was chosen because it is the most widely studied prokaryote and \( B. \) subtilis spores as a representative organism able to survive extreme conditions and long periods of time with little or no nutrients. Also, bacterial spores are highly resistant to radiation, heat, and harsh chemicals (Setlow, 2006), so it was speculated that their inherent durability would provide a challenge to hydrolysis via SCWE. Although lower cell densities are predicted for samples on ocean worlds (Hand et al., 2017), \( \sim 10^8 \) cells/mL and \( \sim 10^7 \) spores/mL were used for this study to ensure that a sufficient amount of amino acids could be extracted to observe the different distributions. While amino acid analysis was performed using CE-LIF, the goal of this work was to understand any disturbance of the molecular distributions by SCWE upstream of any potential analytical system. As such, work on cellular limits of detection for an explicit end-to-end analysis scenario is outside of the scope of this work and will be reported separately.

We demonstrated that the addition of acid leads to a greater extraction efficiency for bacterial cells and spores, similarly to previous reports with soil. This demonstrates that by reducing the pH, acid-catalyzed hydrolysis
occurs leading to higher yield of extraction of both whole cells as well as free proteins and polypeptides in a sample.

While higher extraction efficiency is desired, the preservation of molecular distributions is equally important in the context of life detection. For amino acids, we focus on two distributions: enantiomeric excess and relative abundances. Enantiomeric excess is important because biological amino acids are mostly left-handed, while racemic mixtures have mostly been found in meteorites (Glavin et al., 2010). Because racemization can occur at high temperatures and extreme pH values, we determined enantiomeric excesses when using water and diluted HCl as extraction solvents (Table 1). Because the largest reported %ee of an abiotic amino acid is 60% (Pizzarello et al., 2012), we used this value as a threshold to determine if a sample has a biotic component. Therefore, we interpret a %ee > 60% to be indicative of a clear biotic origin. Although minor racemization occurred during SCWE, enantiomeric excess as a biosignature was preserved. Histidine was well preserved in water, but racemization appears to occur in acid, moving the ee below the 60% threshold. Serine shows a behavior that is organism-dependent, with a clear biotic signal for B. subtilis spores but not E. coli cells. Because D-serine is
known to be present in *E. coli* during its stationary phase of growth (as a mechanism to signal a reduction in metabolism (Lam et al., 2009)), it is possible that after removal from the growth media the population entered the stationary phase. While racemization during extraction could also play a role, that is less consistent with the large %ee measured with the *B. subtilis* spores. In contrast, a D-alanine excess was found for both *B. subtilis* spores and *E. coli* cells when extracted in water, but an L-alanine excess was observed when using acid. This can be explained by the presence of D-alanine in bacterial cell membranes as part of peptidoglycan (Vollmer et al., 2008) in combination with incomplete protein hydrolysis during extraction. Peptidoglycans are linear glycan strands cross-linked by short peptides; for most bacteria the peptide contains five amino acid residues, with one L-alanine and two D-alanine units. Because most cellular proteins contain hundreds to thousands of amino acids, it is reasonable to expect their complete hydrolysis would take longer than the hydrolysis of peptidoglycan. Thus, in water, when the hydrolysis of peptides is less efficient, we observe the D-alanine excess derived from the hydrolysis of peptidoglycan. However, when the efficiency of peptide hydrolysis is increased by the addition of acid we see a shift to an excess of L-alanine, as expected. However, the clear imprint of the more easily released D-alanine from the peptidoglycan remains, further demonstrating that even with the pH adjustment, this method does not fully hydrolyze all proteins in the sample. Overall, this data shows that these extraction conditions do not consistently racemize amino acids, with the exception of histidine at low pH. Additionally, although an excess of L-amino acids is expected for terrestrial life, the measurement of an excess of D-amino acids in extraterrestrial samples would also be scientifically significant and could indicate biological origin as an excess of D-amino acids has never been found in meteorites. The data also shows the importance of looking for multiple chiral pairs, because even when dealing with pure cell strains, significant biological sources of D-amino acids can be present, leading to an ambiguous measurement when taken alone.

It is known that amino acids can decompose under SCWE conditions, potentially disrupting the second distribution of interest, their relative abundance. Although the main decomposition products are carbon monoxide, water, and ammonia, production of glycine, alanine, valine, and proline has been reported (Abdelmoez et al., 2007, 2010). However, our results indicate that glycine is the main degradation product for both cells and spores. In addition, it has been shown that amino acids are more labile in acidic and neutral solutions (Abdelmoez et al., 2007), the conditions preferred here to increase overall yield. Consequently, it is possible a biotic sample could appear as abiotic because of the decomposition of more complex amino acids to glycine, thus skewing the molecular distribution. As shown in Figure 2, both *E. coli* cells and *B. subtilis* spores extracts show relatively low abundances of amino acids compared to glycine, making it difficult to assess whether the distribution appears more like the biotic reference or the abiotic reference. Because of this, we calculated the relative abundances of amino acids with respect to the total amino acid content excluding glycine, that is, [AA]/[His] + [Leu] + [Val] + [Ser] + [Ala]). As shown in Figure 2b, we can clearly see that the distribution patterns are more similar to the biotic reference, whereas it is visually much more ambiguous when looking at the ratios to glycine in Figure 2a.

To quantify how the relative abundances overlap with each reference, we calculated cosine similarities according to Equation 2 (Figure 3). First, we compared the biotic and abiotic references to each other, obtaining values of
0.77 and 0.74 for ratios to glycine and total amino acid content, respectively. Then, we calculated the similarity between our measured values and both the biotic and abiotic references. As shown in Figure 3a, there is no clear distinction when comparing relative abundances to glycine to the biotic and abiotic references, with all values ranging between 0.74 and 1. In contrast, cosine similarities calculated using relative abundances to the total amino acid content (Figure 3b) show two clear trends. A high similarity was obtained compared to the biotic reference, with all values grouped above 0.92. On the other hand, cosine similarities to the abiotic reference are quite low (between 0.48 and 0.77), indicating the two vectors are indeed different. These results are consistent with the fact that the measured glycine concentration after SCWE is higher than the native amount present in the sample, so the calculated abundances do not reflect the biotic nature of the sample clearly. Cosine similarities calculated using ratios relative to Ala (data not shown) display similar trends as using total amino acid content minus Gly, supporting our conclusion that degradation during SCWE leads mostly to the formation of Gly. However, despite the observed degradation to glycine, the pattern obtained can be distinguished from an abiotic one when using abundances relative to the total amino acid content. These results further demonstrate the challenge of making smoking gun measurements, and the need for multiple lines of evidence, as unknown biases could be introduced during sample processing and analysis.

These findings highlight the fact that SCWE can successfully release amino acids from microorganisms while preserving amino acid distributions that would indicate a biotic origin. However, they also demonstrate that histidine racemization can occur, and that conversion of amino acids to glycine is happening, so it is important to factor these effects into assessments of biogenicity of unknown samples. While this work showed that amino acid biosignature distributions are preserved after being released from whole cells and spores via SCWE, future work will continue to quantify the lower limit of cellular detection as well as mitigation strategies to further enhance extraction yields.

5. Conclusions

This work is the first demonstration of the release of amino acid biosignatures from whole microorganisms using SCWE. Despite the known potential for amino acid degradation under SCWE conditions, this work shows that a good extraction yield and minimal disturbance of amino acid distributions are achieved with this approach. Besides the ability of SCWE to release amino acids from vegetative cells, we were also able to demonstrate that SCWE can lyse and break down bacterial spores that can survive under extreme conditions. Thus, this work highlights the utility of SCWE for future life detection missions, regardless of the analytical method used to detect organics. The combination of SCWE for sample treatment with a separation method for analysis would greatly increase the chances of detecting organic biosignatures in situ.

Data Availability Statement

All data included here are available at https://doi.org/10.5281/zenodo.5998110.

References

Abdelmoez, W., Nakahasi, T., & Yoshida, H. (2007). Amino acid transformation and decomposition in saturated subcritical water conditions. Industrial & Engineering Chemistry Research, 46(16), 5286–5294. https://doi.org/10.1021/ie070151b
Abdelmoez, W., Yoshida, H., & Nakahasi, T. (2010). Pathways of amino acid transformation and decomposition in saturated subcritical water conditions. International Journal of Chemical Reactor Engineering, 8(1). https://doi.org/10.2202/1542-6580.1903
Aerts, J. W., Riedo, A., Melton, D. J., Martini, S., Flahaut, J., Meierhenrich, U. J., et al. (2020). Biosignature analysis of Mars soil analogs from the Atacama Desert: Challenges and implications for future missions to Mars. Astrobiology, 20(6), 766–784. https://doi.org/10.1089/ast.2019.2063
Akira, S., Kaoru, H., & Keizo, Y. (1985). Amino acids from the Yamato-791198 carbonaceous chondrite from Antarctica. Chemistry Letters, 14(8), 1183–1186
Amshukeli, X., Pelletier, C. C., Kirby, J. P., & Granthamer, F. J. (2007). Subcritical water extraction of amino acids from Atacama Desert soils. Journal of Geophysical Research, 112(G4), G04S16. https://doi.org/10.1029/2006JG000308
Booth, I. (1987). *Escherichia coli* and *Salmonella typhimurium*: Cellular and molecular Biology (pp. 493–494). American Society for Microbiology Publishers.
Chyba, C. F., & Phillips, C. B. (2002). Europa as an abode of life. Origins of Life and Evolution of the Biosphere, 32(1), 47–67. https://doi.org/10.1023/a:1013958519734
Creamer, J. S., Mora, M. F., & Willis, P. A. (2017). Enhanced resolution of chiral amino acids with capillary electrophoresis for biosignature detection in extraterrestrial samples. Analytical Chemistry, 89(2), 1329–1337. https://doi.org/10.1021/acs.analchem.6b04338
