ASTROBIOLOGY
Volume 23, Number 1, 2023
Mary Ann Liebert, Inc.
DOI: 10.1089/ast.2022.0023

Research Articles

Open camera or QR reader and scan code to access this article and other resources online.

A Deep Ultraviolet Raman and Fluorescence Spectral Library of 51 Organic Compounds for the SHERLOC Instrument Onboard Mars 2020

Joseph Razzell Hollis,[1,2] Sunanda Sharma,[1] William Abbey,[1] Rohit Bhartia,[3] Luther Beegle,[1] Marc Fries,[4] Jeffrey D. Hein,[1] Brian Monacelli,[1] and Austin D. Nordman[1]

Abstract

We report deep ultraviolet (DUV) Raman and Fluorescence spectra obtained on a SHERLOC (Scanning Habitable Environments with Raman and Luminescence for Organics and Chemicals) analog instrument for 51 pure organic compounds, including 5 carboxylic acids, 10 polycyclic aromatic hydrocarbons, 24 amino acids, 6 nucleobases, and 6 different grades of macromolecular carbon from humic acid to graphite. Organic mixtures were not investigated. We discuss how the DUV fluorescence and Raman spectra exhibited by different organic compounds allow for detection, classification, and identification of organics by SHERLOC. We find that 1- and 2-ring aromatic compounds produce detectable fluorescence within SHERLOC's spectral range (250–355 nm), but fluorescence spectra are not unique enough to enable easy identification of particular compounds. However, both aromatic and aliphatic compounds can be identified by their Raman spectra, with the number of Raman peaks and their positions being highly specific to chemical structure, within SHERLOC's reported spectral uncertainty of $\pm 5\,cm^{-1}$. For compounds that are not in the Library, classification is possible by comparing the general number and position of dominant Raman peaks with trends for different kinds of organic compounds. Key Words: SHERLOC—Perseverance rover—Organic library—Mars—DUV Raman. Astrobiology 23, 1–23.

1. Introduction

On February 18, 2021, NASA's Perseverance rover landed on Mars and began surface science operations with the goal of exploring and understanding the geological and astrobiological history of Jezero crater (Farley *et al.,* 2020; Williford *et al.,* 2018). To this end, Perseverance is equipped with seven scientific instruments, including SHERLOC (Scanning Habitable Environments with Raman and Luminescence for Organics and Chemicals). SHERLOC is a compact deep ultraviolet (DUV) Raman and fluorescence spectrometer mounted on the turret at the end of the rover's robotic arm. It is designed to detect and characterize organic material and minerals *in situ*, providing a first assessment of any potential biosignatures that may be found preserved in the rocks of Jezero, as well as essential documentation for cached samples that will be eventually returned to Earth (Bhartia *et al.,* 2021).

SHERLOC operates by scanning the target surface with a 248.6 nm laser, collecting any photons of Raman scattering or fluorescence emission from illuminated material in the near-surface, and producing a spectral map indicative of organic and mineral composition at submillimeter spatial resolutions. The principles behind combined fluorescence and Raman spectroscopy relating to SHERLOC have been described in detail elsewhere (Bhartia *et al.,* 2021; Razzell Hollis *et al.,* 2021). SHERLOC will contribute to the search for organic matter and potential biosignatures in Jezero

[1]NASA Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA.
[2]Department of Life Sciences, The Natural History Museum, London, United Kingdom.
[3]Photon Systems, Covina, California, USA.
[4]NASA Johnson Space Center, Houston, Texas, USA.

crater by detecting, classifying, and identifying organic compounds based on their DUV Raman and fluorescence spectra, and correlating their spatial distribution to mineralogy and visible morphology (Bhartia *et al.,* 2021).

The use of DUV excitation provides SHERLOC with significantly increased sensitivity to aromatic organic compounds, meaning that these molecules can produce detectable fluorescence and/or Raman signals even at very low concentrations. The wavelength of fluorescence and frequency of any dominant Raman peaks provide the first insights into the class of organic compounds detected; for example, whether it is an aromatic carboxylic acid versus aliphatic amino acid. Identification of specific organic compounds relies on comparing the observed Raman spectrum with the spectra of known molecules, as the patterns of vibrational modes (the number of Raman peaks and their positions) are highly specific to chemical structure.

Although Raman peak positions and shapes are generally independent of the excitation wavelength used, fluorescence background variability and Raman resonance effects can lead to meaningful variations in measured spectra. Therefore, it is crucial that reference spectra are acquired using the appropriate DUV excitation wavelength rather than relying on pre-existing Raman databases that are typically acquired using visible excitation. Furthermore, SHERLOC's mass and volume requirements lead to a unique optical design, which means previous studies done on other DUV Raman instruments may not be entirely representative of SHERLOC in terms of detection thresholds or observed peak intensities and widths (Bhartia *et al.,* 2021).

To help accurately classify/identify individual organics and minerals on Mars, we have assembled a DUV Spectral Library of reference samples that are relevant to Jezero crater and astrobiology. We have used the SHERLOC Brassboard, a custom DUV spectrometer designed to be as optically similar to SHERLOC as possible while being able to operate under ambient temperature and pressure on Earth, to record spectra similar to what SHERLOC would observe.

DUV spectra for 62 different minerals have already been published (Razzell Hollis *et al.,* 2021), and we now describe the spectra for 51 distinct organic compounds from several different groups/classes that are considered astrobiologically

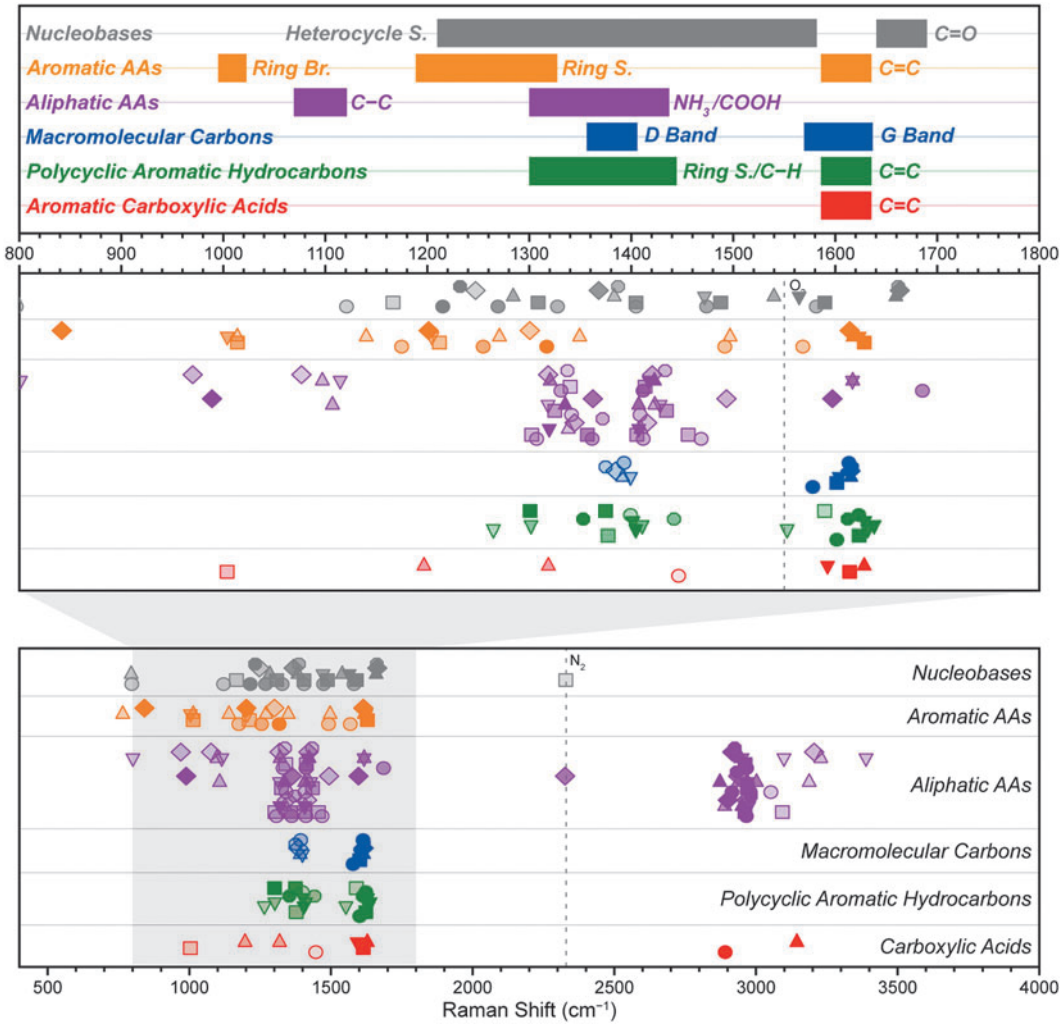

FIG. 1. Summary of Raman peak positions for 51 organic compounds, grouped and colored according to class, showing how different organics can be classified and identified by the pattern of major Raman peak positions. Within each group, peaks from the same spectrum are aligned horizontally and share a common symbol, and their relative intensities are indicated by opacity. AA = amino acid; *Br.* = breathing mode; *S.* = stretching mode.

relevant, including carboxylic acids, polycyclic aromatic hydrocarbons, α-amino acids, nucleobases, and macromolecular carbon. These groups cover a broad range of possible organic signatures that SHERLOC might detect: purely abiotic organic molecules, the building blocks of proteins and nucleic acids in terrestrial organisms, and carbonaceous material typical of geothermally degraded biological matter present in fossilized organisms.

In this study, we consider how the DUV fluorescence and Raman spectra of different organics may be used to (1) detect different kinds of organic material, (2) classify the type of material, and (3) identify specific compounds. Figure 1 shows how the major peak positions of different classes tend to appear at certain Raman shifts, with skeletal modes (*e.g.*, C–C/C = C) appearing between 1000 and 1800 cm^{-1}, and hydrogen modes (C–H/O–H, etc.) appearing at 2900–3400 cm^{-1}, demonstrating how organics may still be classified by the pattern of Raman peaks even if the specific spectrum is not in the library.

The addition of organic references to the DUV Spectral Library will enable the Mars 2020 science team to rapidly interpret the organic aspects of SHERLOC spectra obtained during surface operations, and support tactical decision making regarding the caching of samples for eventual return to Earth (Beaty *et al.,* 2019; MEPAG E2E-iSAG, 2012).

2. Methods

Table 1 lists the individual organic compounds used in this study, along with their sources and minimum purities where applicable. Coal samples with source IDs including ''DECS'' or ''PSOC'' were obtained from the Pennsylvania State University Coal Sample Bank.

2.1. Sample preparation

With the exception of highly oriented pyrolytic graphite (HOPG), which was a single square crystal mounted face-up on a metal surface, all samples consisted of fine powders, and were used as received without further preparation or purification. A small quantity of powdered material was placed on a clean aluminum wafer using a cleaned spatula and patted flat to provide a more even surface for scanning. Spatulas and wafers were cleaned with distilled water and ethyl alcohol between uses.

2.2. DUV spectroscopy

All samples were measured on the SHERLOC Brassboard instrument at the NASA Jet Propulsion Laboratory, which is described in detail in Razzell Hollis *et al.* (2021). The Brassboard is an optical analogue of SHERLOC, adapted to work under terrestrial ambient conditions, which also uses a pulsed NeCu laser (Photon Systems, Inc.) with an excitation wavelength of 248.5794 nm. Spectral calibration was based on fitting a quadratic function to the secondary laser line at 252.93 nm and the reported peak positions of three Raman standards: acetonitrile, powdered calcite, and HOPG, which occur between 919 and 2942 cm^{-1} (254.37–268.17 nm).

The mean absolute error of the calibrated Raman peak positions with respect to their literature values was $2.7 \pm 1.6\,cm^{-1}$ ($\approx 0.018 \pm 0.010$ nm). The spectral calibration in the fluorescence region (275–353 nm) was validated using the emission lines of an Hg arc lamp, which showed that the mean absolute error of the calibration increases considerably with wavelength beyond the calibrated range, averaging 0.38 ± 0.22 nm between 275–310 nm, and 1.24 ± 0.27 nm between 310 and 355 nm (Razzell Hollis *et al.,* 2021).

The off-axis optical design of both instruments causes the projection of the spectrum to curve across the charge-coupled device (CCD), referred to as a ''spectral smile'' (Uckert *et al.,* 2021). To avoid disruptive levels of dark noise in the Raman region, which would result from integrating over the full height of the CCD, both instruments must read out data for the Raman and fluorescence regions separately using different vertical binning ranges. The SHERLOC onboard flight software is capable of reading each region independently from a single acquisition (Bhartia *et al.,* 2021). The Brassboard instrument emulates this by reading the Raman region and the fluorescence region of the CCD in two separate scans, which are then trimmed and combined to produce a full spectrum (Razzell Hollis *et al.,* 2021).

To encompass the full range of exposure durations and energies that may be used in SHERLOC activities, each sample was scanned multiple times with the Brassboard instrument at different laser settings similar to those used by SHERLOC: 5–800 pulses per point (approximately equivalent to a total laser energy of 21–3300 μJ, or 0.28–44 J/cm^2). Each sample was measured using an automated sequence of six scans at different laser settings. Each scan recorded 25 individual spectra in a 5×5 grid with 150 μm spacing (750×750 μm), and was done in a different location to minimize any photochemical damage caused by cumulative exposure to the DUV laser.

The sequence of laser settings used was 5/25/100/400/800 laser pulses per point with the laser power supply providing 20 Å of current at 375 V, followed by a final scan of 800 pulses per point at 15 Å and 375 V. The output energy of each laser pulse during each scan was recorded by a photodiode at the laser output, calibrated to the energy received at the sample. The average pulse energy was 2.84 ± 0.37 μJ/pulse at 15 Å and 4.15 ± 0.38 μJ/pulse at 20 Å. Images were taken before and after each scan, and no visible signs of laser-induced photodamage were observed for any sample.

2.3. DUV data processing

All spectra were processed using in-house Python scripts utilizing Numpy (van der Walt *et al.,* 2011), SciPy (Jones *et al.,* 2001), and LMFIT (Newville *et al.,* 2014) packages. Each scan was processed in the following steps, which are visualized in Supplementary Fig. S1: (1) individual spectra in the scan were rescaled to account for intensity changes resulting from variations in laser output, by multiplying intensity of each spectrum by a normalization factor equal to the maximum laser energy recorded during the scan divided by the laser energy recorded during that spectrum. (2) Cosmic rays were automatically identified as outliers in the distribution of intensity values in each CCD channel and replaced by the average value of adjacent spectral points (Uckert *et al.,* 2019).

(3) Because the Spectral Library required measurement of nominally homogeneous samples, outlier spectra were removed automatically, defined as >10% of the spectrum

Table 1. Summary of the Organic Compounds Characterized in This Study, with Their Minimum Purities and Sources, and Product Numbers and Lots Where Available

Sample	Min. purity, %	Source
Carboxylic acids		
Palmitic acid	99	Sigma P0500 (lot SLBN0554V)
Benzoic acid	99.50	Sigma-Aldrich 242381 (lot MKBW7354V)
4-Ethylbenzoic acid	99	Aldrich 191280
Phthalic acid[a]	99.50	Sigma-Aldrich 402915 (lot BCBK8052V)
Mellitic acid[a]	99	Aldrich M2705 (lot 08205PO)
Polycyclic aromatic hydrocarbons		
Biphenyl	99.50	Sigma-Aldrich B34656 (lot MKBV9808V)
Fluorene	98	Aldrich 128333
Naphthalene[a]	99	Sigma-Aldrich 147141 (lot 06004TH)
Anthracene	99	Sigma-Aldrich 31581 (lot SZBD155XV
Anthracene	99	Fluka 10580
1-Methylanthracene	97	Aldrich 333573
Phenanthrene[a]	96	Sigma P2528 (lot 052K1449)
Tetracene	98	Aldrich B2403
Fluoranthene	98	Aldrich 423947
Pyrene[a]	99	Sigma 82648 (lot BCBF4410B)
Perylene[a]	99.50	Sigma-Aldrich 394475 (lot 11524CB)
Amino acids		
Glycine	99.60	J.T. Baker 4059-00 (lot Y25588)
D-Alanine	99	Sigma-Aldrich 162655 (lot 04330HI)
L-Alanine	98	Sigma-Aldrich A7627 (lot 29H0949)
L-Arginine HCl	98	Sigma-Aldrich A5131 (lot 29H0695)
L-Asparagine	98	Sigma-Aldrich A0884 (lot 29H5419)
L-Aspartic acid	98	Sigma-Aldrich A9256 (lot 29H0070)
L-Cystine	98	Sigma-Aldrich C8755 (lot 78H0880)
L-Glutamine	98	Sigma-Aldrich G3126 (lot 19H0754)
L-Glutamic acid	98	Sigma-Aldrich G1251 (lot 118H0989)
L-Histidine	98	Sigma-Aldrich H8125 (lot 107H1452)
L-Hydroxyproline	98	Sigma-Aldrich H5534 (lot 108H1225)
L-Isoleucine	98	Sigma-Aldrich I2752 (lot 28H0046)
L-Leucine	98	Sigma-Aldrich L8000 (lot 128H1155)
L-Lysine HCl	98	Sigma-Aldrich L5626 (lot 48H1031)
L-Methionine	98	Sigma-Aldrich M9625 (lot 88H1334)
D-Phenylalanine		Unknown
L-Phenylalanine	98	Sigma-Aldrich P2126
L-Proline	98	Sigma-Aldrich P0380 (lot 128H1156)
D-Serine	98	Sigma-Aldrich S4250 (lot 107H1251)
L-Serine	98	Sigma-Aldrich S4500
L-Threonine	98	Sigma-Aldrich T8625 (lot 108H0660)
L-Tryptophan[a]	98	Sigma-Aldrich T0254 (lot 102K0370)
L-Tyrosine	98	Sigma-Aldrich T3754 (lot 107H1188)
L-Valine	98	Sigma-Aldrich V0500 (lot 38H0564)
Nucleobases		
Purine	98	Sigma-Aldrich P55805 (lot STBG1731V)
Adenine	99	Sigma-Aldrich A8626 (lot WXBC1564V)
Cytosine	99	Sigma-Aldrich C3506 (lot SLBF0582V)
Guanine	98	Sigma-Aldrich G11950 (lot BCBQ3784V)
Thymine	99	Sigma-Aldrich T0376
Uracil	99	Sigma-Aldrich U0750
Macromolecular carbons		
Humic acid		Fluka 53680 (lot 1102067)
Lignite		PSOC-1533
Sub-bituminous B		DECS-26 (Freestone County, TX)
Bituminous hvC		DECS-5 (Sevier County, UT)
Anthracite		PSOC-1468
HOPG		SPI Supplies

Samples that have been previously characterized on other instruments are identified as follows: [a]Abbey et al (2017).

DECS = Department of Energy Coal Sample; HOPG = highly oriented pyrolytic graphite; PSOC = Penn. State/Office of Coal Research.

being >2 standard deviations from the average, followed by recalculation of the average and standard deviation using remaining spectra. (4) For duplex scans that recorded both the Raman and fluorescence regions of the CCD, the average Raman and average fluorescence values were added together, after trimming to remove dark noise from nonread regions. (5) Background intensity was removed from the Raman region of the averaged spectrum by subtraction of a polynomial baseline between 0 and 4000 cm^{-1}. The baseline was fitted to a set of points at Raman shifts that were free of Raman peaks, specific to each mineral or its class, and the fit was optimized by linear least-squares regression.

For some samples, a second baseline fit and subtraction was required to ensure a clean background removal. (6) The baselined spectrum was then normalized, dividing intensity values by the maximum found between 400 and 4000 cm^{-1}, ignoring the 600–800 cm^{-1} region obscured by a secondary

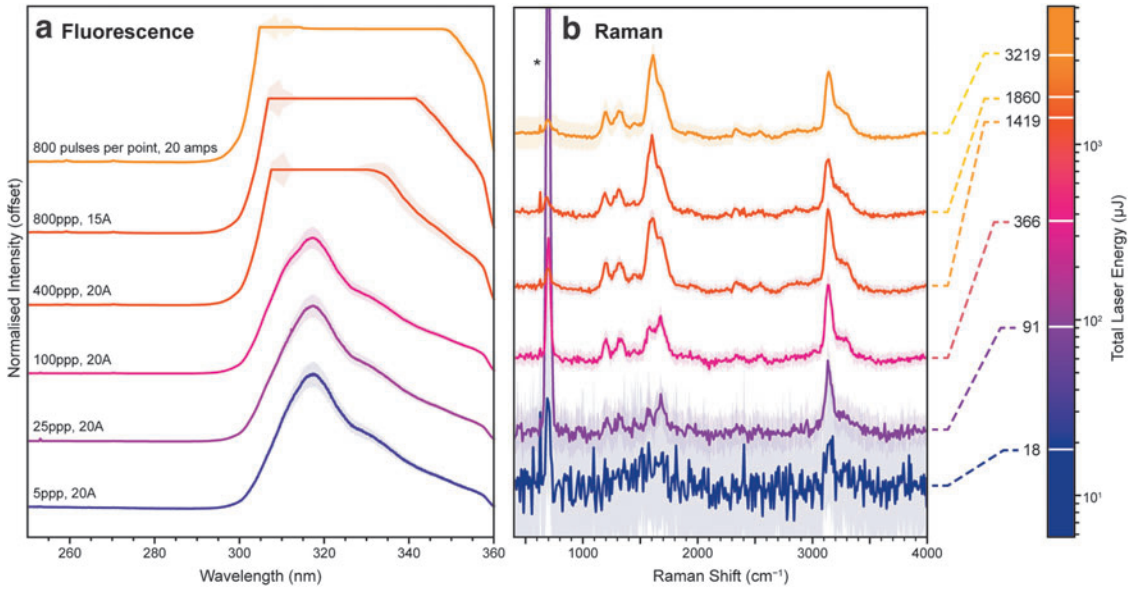

FIG. 2. Variation in the DUV fluorescence **(a;** left**)** and Raman **(b;** right**)** spectra of mellitic acid under Earth ambient conditions, measured using different laser settings (in pulses per point, ppp, and amps, Å) to explore a range of total exposure energies comparable with different SHERLOC activities. Each spectrum was averaged over 25 points, then normalized to the maximum intensity in either the 255–355 nm region, or the 400–4000 cm^{-1} region after baseline subtraction. Colored areas indicate ±1 standard deviation. *Indicates the NeCu laser emission line at 253.93 nm. DUV = deep ultraviolet; SHERLOC = Scanning Habitable Environments with Raman and Luminescence for Organics and Chemicals.

NeCu laser emission line at 252.93 nm (~695 cm^{-1}). (7) Raman peak positions and intensities were determined by dividing the spectrum into sections and fitting each peak with a Gaussian function, with multiple functions used for overlapping peaks.

3. Results

During surface operations, SHERLOC activities typically consist of (1) survey/recon scans that rapidly cover a large area using fewer laser pulses per point, mainly detecting high-yield fluorescence signal, and (2) detail scans that focus on a smaller area of interest with more laser pulses per point to maximize signal from lower yield Raman scattering.

The parameters of each SHERLOC scan can be tailored to the target to ensure the best possible dynamic range for a given operating window, with nominal values of 4–10 pulses per point for a survey scan (~36–90 µJ of total incident laser energy) and 200–800 pulses per point for a detail scan (1800–8100 µJ). To replicate these operating regimes with the Brassboard instrument, each sample was scanned multiple times at different laser settings between 5 and 800 pulses per point, and either 15 or 20 Å of driving current, to cover a range of laser energies from ~20 to ~3300 µJ. The full set of spectra for each sample at different exposures are presented in Supplementary Figs. S2–S11.

The Brassboard instrument is not identical to SHERLOC, due to the necessity of operating under terrestrial ambient conditions rather than martian ambient, which leads to some discrepancies in optical output and measured spectra. These differences are described in detail in our previous paper (Razzell Hollis *et al.,* 2021), but are generally as follows: (1) a lower laser energy output (3–4 µJ per pulse vs. ~9 µJ per pulse for SHERLOC), leading to lower Raman and fluorescence signal for a given number of pulses; (2) less effective rejection of secondary laser emission lines, namely the 252.93 nm line that appears in most Brassboard spectra and effectively obscures the 650–750 cm^{-1} region; (3) a dip in fluorescence at ~298 nm, which has not been observed in SHERLOC spectra, attributed to an instrumental artifact. A

TABLE 2. RAMAN AND FLUORESCENCE SIGNAL VALUES OBTAINED FOR MELLITIC ACID WITH VARYING LASER SETTINGS

Pulses per Point	*Current, Å*	*Dose, µJ*	*Fluor. signal, counts*	*Raman signal, counts*	*Noise, counts*	*Raman SNR*
5	20	18.3	3530	n.d.	1.1	
25	20	91.4	14,437	4.2	1.2	3.4
100	20	366	36,568	37.2	3.7	10
400	20	1420	64,134	272	6.0	45
800	15	1860	63,312	390	10.4	37
800	20	2980	65,899	664	8.9	74

Fluorescence signal was defined as the average intensity at 316 nm, Raman signal was the average intensity at 1600 cm^{-1} after baseline subtraction, and noise was estimated as the standard deviation of the baselined spectrum between 2000 and 2100 cm^{-1}.
SNR = signal:noise ratio.

second artifact is observed at ~355 nm, where measured signal drops precipitously as it approaches the edge of the detector.

Figure 2 shows the DUV Raman and fluorescence spectrum of powdered mellitic acid, an aromatic carboxylic acid that is considered representative of highly oxidized aromatic molecules that remain after prolonged exposure of organic material to the martian surface environment (Benner *et al.,* 2000; Moores and Schuerger, 2012). Multiple spectra are shown to highlight how the spectrum varied when measured using different laser settings and total laser energies, reflecting different operating regimes for SHERLOC.

As an aromatic organic compound, mellitic acid has a strong absorption band in the ultraviolet (UV), leading to intense fluorescence and fulfilling the conditions for resonant Raman, which significantly enhances scattering from vibrational modes associated with the chromophore, in this case the ring-stretching modes of the benzene unit. As a result of DUV resonant enhancement, the Raman spectrum of mellitic acid is only $\sim 10^3$ times weaker than its fluorescence (Table 2), and both contributions can be observed simultaneously due to the effective spectral separation of Raman scattering and fluorescence in different wavelength ranges (Bhartia *et al.,* 2008, 2021).

Mellitic acid exhibited a single strong fluorescence peak with a maximum at ~316 nm, attributable to the electronic π^*-π transition of the benzene ring. The fluorescence peak was visible even with the shortest exposure of 5 laser pulses (~18 μJ), increasing in intensity in proportion to the total laser energy until it reaches the detector's saturation limit of ~64,000 counts between 100 and 400 laser pulses (~370 and ~1420 μJ). The detection of strong fluorescence using just 5 laser pulses supports the expectation that SHERLOC will be readily able to detect trace amounts of aromatic organic material during survey activities, particularly once the higher energy output of the SHERLOC laser is accounted for (Bhartia *et al.,* 2021; Razzell Hollis *et al.,* 2021).

In the Raman region, the vibrational modes of mellitic acid appeared in two spectral ranges, 1200–1700 and 3100–3300 cm^{-1}. Peaks in the 1200–1700 cm^{-1} range are assigned to stretching modes of the aromatic ring and carboxylic acid units, while peaks in the 3100–3300 cm^{-1} range are assigned to the stretching modes of the carboxylic O–H bonds (Osterrothová and Jehlička, 2010; Socrates, 2001). None of the Raman peaks could be resolved above background noise at 5 pulses (~18 μJ) but were just visible at 25 pulses (91 μJ). At longer exposures, measured Raman intensity and signal:noise ratio (SNR) increased approximately in proportion to the total laser energy as expected (Table 2), with SNR increasing from 3.4 at 91 μJ to 74 at 2980 μJ.

We selected seven compounds to represent different classes of organic material relevant to SHERLOC's mission of searching for potential biosignatures on Mars: L-alanine (an aliphatic α-amino acid), L-phenylalanine (an aromatic α-amino acid), naphthalene (a polycyclic aromatic hydrocarbon, hereafter referred to as a PAH), adenine (a nucleobase of ribonucleic acid [RNA] and deoxyribonucleic acid [DNA]), palmitic acid (an aliphatic carboxylic acid), mellitic acid (an aromatic carboxylic acid), and humic acid (a mixture of organic macromolecules).

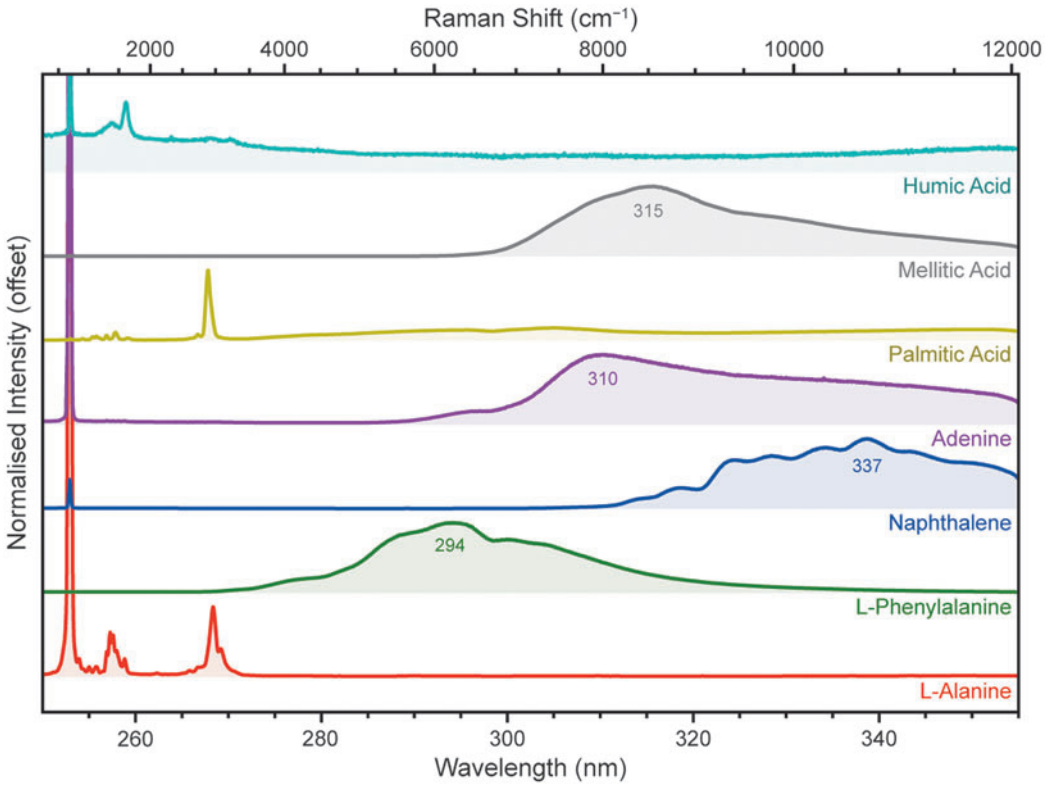

FIG. 3. Normalized DUV fluorescence spectra for seven compounds representing different classes of organic material, measured under Earth ambient conditions. Spectra were averaged over 25 points at 100 laser pulses per point (250–350 μJ) unless stated otherwise in the text.

With the exception of humic acid, all of these compounds have been detected in abiotic organic material extracted from chondritic meteorites, including Murchison (Burton *et al.,* 2012; Pizzarello *et al.,* 2006; Schmitt-Kopplin *et al.,* 2010). L-Alanine, L-phenylalanine, and adenine comprise some of the building blocks of life as we know it, while naphthalene is a common organic marker used to assess organic preservation. Humic acid is not a specific compound but a mixture of various organic macromolecules, rich in phenols and carboxylic acids, and is a major component of decaying plant matter—which makes it a potential analogue for a partially degraded biosignature (McDonald *et al.,* 1998).

Figure 3 shows the average DUV spectrum for each of these compounds, averaged over 25 points at 100 laser pulses per point (a total laser energy of 250–350 μJ). The exceptions were L-phenylalanine and naphthalene, which were measured using lower laser energies of 39 and 22 μJ, respectively, to avoid fluorescence signal saturation.

The spectra in Fig. 3 vary between different classes of organic compounds, which can be put in three categories: (1) aliphatic organics (*e.g.*, L-alanine and palmitic acid) that exhibit very little innate fluorescence in the 250–355 nm range, and have spectra that are instead dominated by their Raman peaks ~255–270 nm (1000–3500 cm^{-1}). (2) Monocyclic and bicyclic aromatic organics (L-phenylalanine, naphthalene, adenine, mellitic acid) that exhibit well-defined fluorescence peaks in the 250–355 nm range. (3) Tricyclic and larger aromatics (*e.g.*, humic acid), which are known to fluoresce at longer wavelengths that fall entirely outside than SHERLOC's detection range (Cloutis *et al.,* 2016; Dartnell *et al.,* 2012; Eshelman *et al.,* 2018), leading to a DUV spectrum that exhibits only Raman peaks.

The shift of aromatic fluorescence to longer wavelengths with increasing size of the aromatic unit is the result of greater delocalization of π-electrons, which reduces the energy gap for optical transitions such as absorption and fluorescence (Lakowicz, 2006). Some spectra exhibit multiple local maxima, for example, naphthalene, which we attribute to strong vibronic coupling between radiative relaxation and individual vibrational energy levels, usually requiring molecules that are organized into a highly ordered, crystalline structure. We again observe a regular dip in signal at 298 nm, most obvious in the adenine spectrum, which cannot be attributed to actual emission but is rather a drop in signal received by the detector, as it was consistently observed across several disparate spectra in this study.

Figure 4 shows the normalized Raman region of each compound's DUV spectrum, after subtraction of a polynomial baseline. The chosen spectra were acquired using 800 pulses per point (1800–2700 μJ) to maximize the measured Raman signal, with the exception of palmitic acid, which was acquired with 100 pulses (369 μJ) to avoid signal saturation. Each compound in Fig. 4 exhibited a very distinct spectrum, with Raman peaks generally occurring in two distinct spectral regions, 800–1800 cm^{-1} and between 2800 and 3400 cm^{-1}.

Assignment of Raman peaks to specific vibrational modes is highly dependent on structure, but may be generalized to some extent: organic Raman peaks around 800–1800 cm^{-1} tend to be

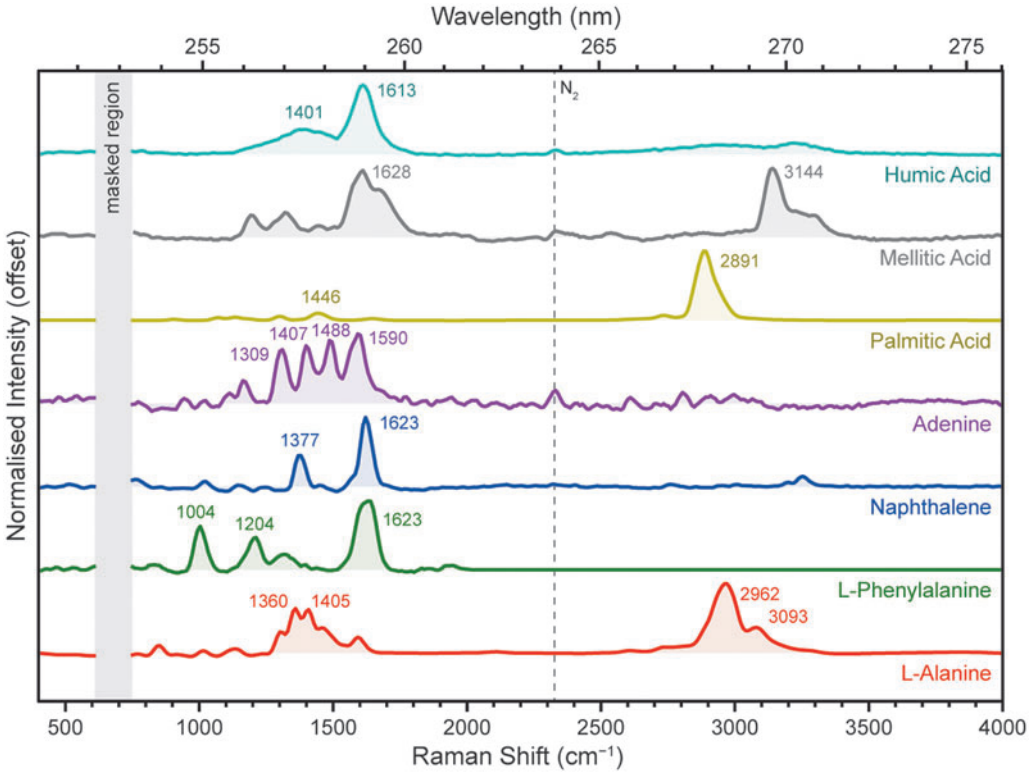

FIG. 4. Normalized DUV Raman spectra for seven compounds representing different classes of organic material, measured under Earth ambient conditions. Spectra were averaged over 25 points at 800 pulses per point (1800–2600 μJ), and normalized after baseline subtraction. The vertical gray bar indicates a region obscured by laser emission; numbers indicate the average position (in cm^{-1}) of key peaks.

assigned to stretching modes of skeletal bonds, for example, C–C, C=C, C–O, and C=O, and bending/rocking modes of hydrogen-bearing groups such as NH_2 and CH_3, while organic Raman peaks around 2800–3400 cm^{-1} tend to be assigned to stretching modes of C–H, N–H, and O–H bonds.

The selective resonant enhancement of aromatics over the nonresonant aliphatics leads to distinct differences in their overall spectra, in particular the relative intensity of the skeletal mode region (800–1800 cm^{-1}) versus the C–H/N–H/O–H stretching mode region (2800–3400 cm^{-1}). This is a direct consequence of molecular resonance enhancing Raman scattering from those vibrational modes coupled to the absorption transition, which in the case of aromatic organics tend to be ring-breathing/stretching and C=C stretching modes (Asher et al, 1986). Thus, for an aromatic hydrocarbon such as naphthalene, Raman scattering from the C=C stretching vibration at 1600 cm^{-1} is enhanced while the C–H stretching vibrations at ~3100 cm^{-1} are unchanged.

3.1. Carboxylic acids. In addition to palmitic acid and mellitic acid, we investigated three other aromatic carboxylic acids: benzoic acid, 4-ethylbenzoic acid, and phthalic acid. Figure 5a shows the full DUV fluorescence spectra for all four carboxylic acids, as expected we observed strong fluorescence between 250 and 355 nm for all four aromatic acids. The number and position of the carboxylic acid units around the benzene ring profoundly affect the wavelength of its fluorescence, with λ_{max} varying from 304 nm for benzoic acid, to 339 nm for phthalic acid, to 315 nm for mellitic acid. Even the presence of an alkyl group on the ring can have a marked impact on λ_{max}, as shown by the >35 nm shift between benzoic acid 4-ethylbenzoic acid.

In the Raman region, the palmitic acid spectrum was dominated by a single C–H stretching mode at 2891 cm^{-1}, though some very minor peaks can be observed at 1000–1500 cm^{-1}, including a C–C stretching mode at 1446 cm^{-1} (Fig. 5b and Table 3). The aromatic amino acids all

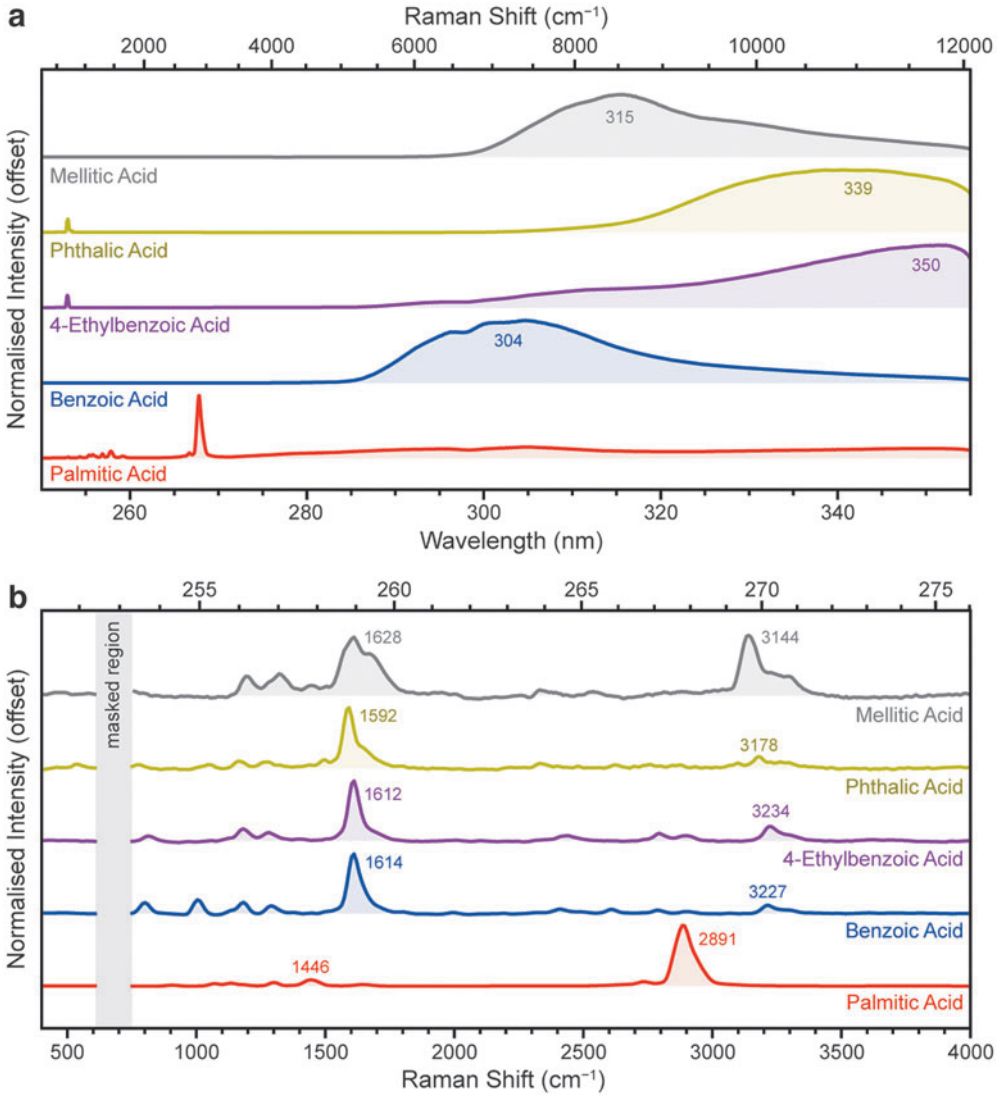

FIG. 5. Normalized DUV fluorescence spectra **(a)** and the baselined Raman region **(b)** for five carboxylic acids, as measured under Earth ambient conditions. Fluorescence spectra were obtained with 100 laser pulses per point (~360 μJ), with the exception of benzoic acid (5 ppp, 18 μJ) and phthalic acid (25 ppp, 92 μJ). Raman spectra were obtained with 800 ppp (1800–2600 μJ), with the exception of palmitic acid (100 ppp, 369 μJ).

Table 3. Spectral Position of the Fluorescence Maximum and Major Raman Peaks of Each Carboxylic Acid, with Assignments to Generalized Vibrational Modes

Sample	*Fluor, nm*	*Raman, cm^{-1}*	*Assignment*	*References*
Palmitic acid	n.d.	1446	C – C (*S.*)	Socrates (2001)
		2891	C – H (*S.*)	Socrates (2001)
Benzoic acid	304	1614	C = C (*S.*)	Socrates (2001)
		3227	Overtone	
4-Ethylbenzoic acid	350	1612	C = C (*S.*)	Socrates (2001)
		3234	Overtone	
Phthalic acid	339	1592	C = C (*S.*)	Osterrothová and Jehlička (2010)
		3178	Overtone	
Mellitic acid	315	1628	C = C (*S.*)	Osterrothová and Jehlička (2010)
		3144	O–H (*S.*)	Socrates (2001)

All Raman peak positions have an estimated uncertainty of $\pm 5\,cm^{-1}$.
n.d. = not detected; *S.* = stretching mode.

appeared quite similar, dominated by a single major peak, the C = C stretching mode of the benzene ring, which appeared at $1614\,cm^{-1}$ in benzoic acid, $1612\,cm^{-1}$ in 4-ethylbenzoic acid, $1592\,cm^{-1}$ in phthalic acid, and $1629\,cm^{-1}$ in mellitic acid.

All four aromatic carboxylic acids exhibited a number of very weak minor modes between 800 and $1300\,cm^{-1}$. Compared with palmitic acid, where the C–H stretching mode at $\sim 2890\,cm^{-1}$ was by far the most dominant peak, the C–H stretching modes of the aromatic carboxylic acids were almost undetectable due to their lack of resonant enhancement. To avoid signal saturation, the palmitic acid spectrum was acquired with 100 laser pulses at 20 Å (369 μJ).

Although multiple peaks were apparent between 2400 and $3300\,cm^{-1}$ for the aromatic carboxylic acids, many of these peaks appear to be second-order overtones of the benzene ring's major vibrational modes, exhibiting the same pattern of relative intensities and positions at approximately twice the frequency. The appearance of overtones is not unusual for resonant Raman spectra, as coupling to electronic excitation permits Raman scattering from higher order vibrational modes, not just fundamental modes (Long, 1977).

Second-order overtones appear at roughly double the Raman shift of the fundamental peaks, and have a similar shape and pattern of intensities, but are usually 10^1–10^2 times less intense. The exception was mellitic acid, which exhibited an unusually strong peak at $3144\,cm^{-1}$, more consistent with a fundamental high-frequency vibration than an overtone of the main C = C peak. Based on its frequency and the number of carboxylic acid groups in the molecule, we have tentatively assigned this peak to the O–H stretching mode of carboxylic acid.

3.2. Polycyclic aromatic hydrocarbons. We measured 10 PAHs of varying size and geometry, including 3 bicyclic molecules (biphenyl, fluorene, and naphthalene), 3 tricyclic molecules (anthracene, 1-methylanthracene, and phenanthrene), 3 tetracyclic molecules (tetracene, pyrene, and fluoranthene), and 1 pentacyclic molecule (perylene). All of these compounds are known to fluoresce under UV illumination, with the wavelength of emission generally increasing with the number of coplanar aromatic rings due to greater delocalization of π-electrons, which results in lower energy absorption/emission transitions (Lakowicz, 2006). PAHs can be formed by biological degradation and are commonly proposed components in prebiotic chemistry.

Combined with their widespread presence in interstellar medium, comets, and meteorites, including those of martian origin, PAHs are of high relevance to astrobiology (Dartnell *et al.,* 2012). As Fig. 6 shows, we readily observed fluorescence from the mono- and bicyclic PAHs, while the larger PAHs tended to fluoresce at wavelengths beyond SHERLOC's detectable range. To avoid signal saturation, the fluorescence spectra of biphenyl, fluorene, and naphthalene were acquired using shorter exposures of 5 laser pulses (18–22 μJ).

The bicyclic PAHs (biphenyl, fluorene, and naphthalene) exhibited broad fluorescence features with λ_{max} of 318, 338, and 337 nm, respectively, and all three showed multiple peaks consistent with vibronic progression as described earlier in Section 3.2. The difference in fluorescence wavelength between biphenyl, fluorene, and naphthalene results from the planarity of their respective structures: the two aromatic rings of biphenyl can freely rotate around the connecting C–C bond, resulting in a lower degree of delocalization, a larger optical band gap, and a shorter wavelength of fluorescence. By comparison, the structures of fluorene and naphthalene are rigid (due to the methylene bridge in fluorene and the fused structure of naphthalene), and therefore the aromatic rings are coplanar, resulting in greater delocalization and a longer wavelength of fluorescence.

Of the tricyclic PAHs (anthracene, 1-methylanthracene, and phenanthrene), only phenanthrene produced any detectable fluorescence within SHERLOC's spectral range (250–355 nm), exhibiting a relatively narrow peak with an apparent λ_{max} of 350 nm, before the fall-off in signal at the edge of the detector (which was apparent in all spectra). This is consistent with the first (0-0) vibronic transition reported in the literature for phenanthrene, at 350 nm, while the λ_{max} is reportedly ~ 370 nm (Azumi and McGlynn, 1962). Anthracene and 1-methylanthracene fluoresce at longer wavelengths than phenanthrene due to their more linear geometry, with a reported λ_{max} at 402 nm and an onset at 370 nm, putting it entirely beyond the range we can observe with SHERLOC (Dartnell *et al.,* 2012).

However, we do observe what appears to be two anomalous fluorescence peaks at ~ 260 and ~ 270 nm for both

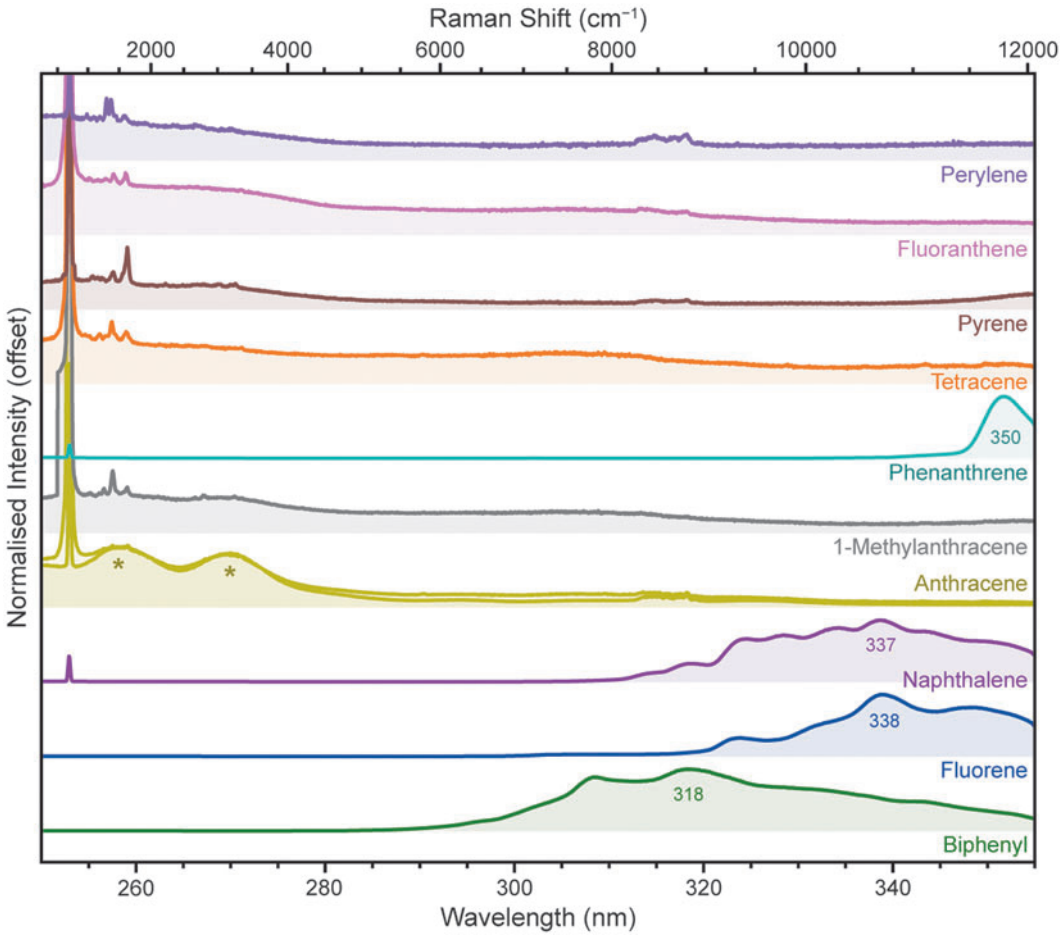

FIG. 6. Normalized DUV fluorescence spectra for 10 PAHs, measured under Earth ambient conditions. The anthracene peaks marked with * are due to instrumental artifacts. PAH = polycyclic aromatic hydrocarbon.

anthracene and 1-methylanthracene, although they were much weaker in 1-methylanthracene. We confirmed that this was not a sample-specific phenomenon such as contamination or degradation by testing a second, independent sample of anthracene from Fluka (#10580), and reported the same peaks. apparent in two independent samples of anthracene and, although more weakly, in one sample of 1-methylanthracene. It is unlikely that we are observing a higher energy $S_2 \rightarrow S_0$ transition, as no such emission has been reported in the literature, only the conventional $S_1 \rightarrow S_0$ transition at $\sim$400 nm (Clairemidi *et al.,* 2008).

We suspected that this may be a false signal produced by instrumental artifacts, and conducted a detailed analysis recording the 2D response of the CCD produced by injecting monochromatic light between 250 and 410 nm into the spectrometer. We discovered that there is a secondary optical path for light between 388 and 402 nm that impinges on the CCD at a similar position to the anomalous peaks (equivalent to 250–260 nm). This secondary optical path is likely a reflection of one of the optical elements within the spectrometer.

This indicates that the anomalous peaks we observed for anthracene and 1-methylanthracene are in fact ''echoes'' of their primary fluorescence at $\sim$400 nm, produced by a fluke of instrument design. Because there are several small differences between the Brassboard and the SHERLOC flight instrument, we are unsure whether the same phenomenon would be observable on SHERLOC, but we consider the likelihood of false signals to be small, considering that it requires a strong fluorescence emission between 388 and 402 nm and the correct geometry. We note that it was only observed in 2 out of a total of 190 organic and mineral samples we have characterized for the Spectral Library so far (Razzell Hollis *et al.,* 2021).

Fluorescence from larger PAHs could not be detected at all, as it fell completely outside SHERLOC's spectral range: tetracene reportedly fluoresces at >460 nm, pyrene at >370 nm, fluoranthene at >400 nm, and perylene at >430 nm (Dartnell *et al.,* 2012; Eshelman *et al.,* 2018). None of these samples produced anomalous peaks at short wavelengths.

In terms of their Raman spectra (Fig. 7), the PAHs all appeared fairly similar to one another, with the majority exhibiting a dominant Raman peak around 1600 cm^{-1} plus one or more peaks between 1000 and 1500 cm^{-1} (Table 4). The exceptions were 1-methylanthracene and perylene, which had a relatively weak $\sim$1600 cm^{-1} peak and were instead dominated by peaks around 1300–1375 cm^{-1}.

In all cases, the $\sim$1600 cm^{-1} peak was assigned to the aromatic C = C stretching mode, while the minor peaks tended to be collective ring-stretching modes of varying

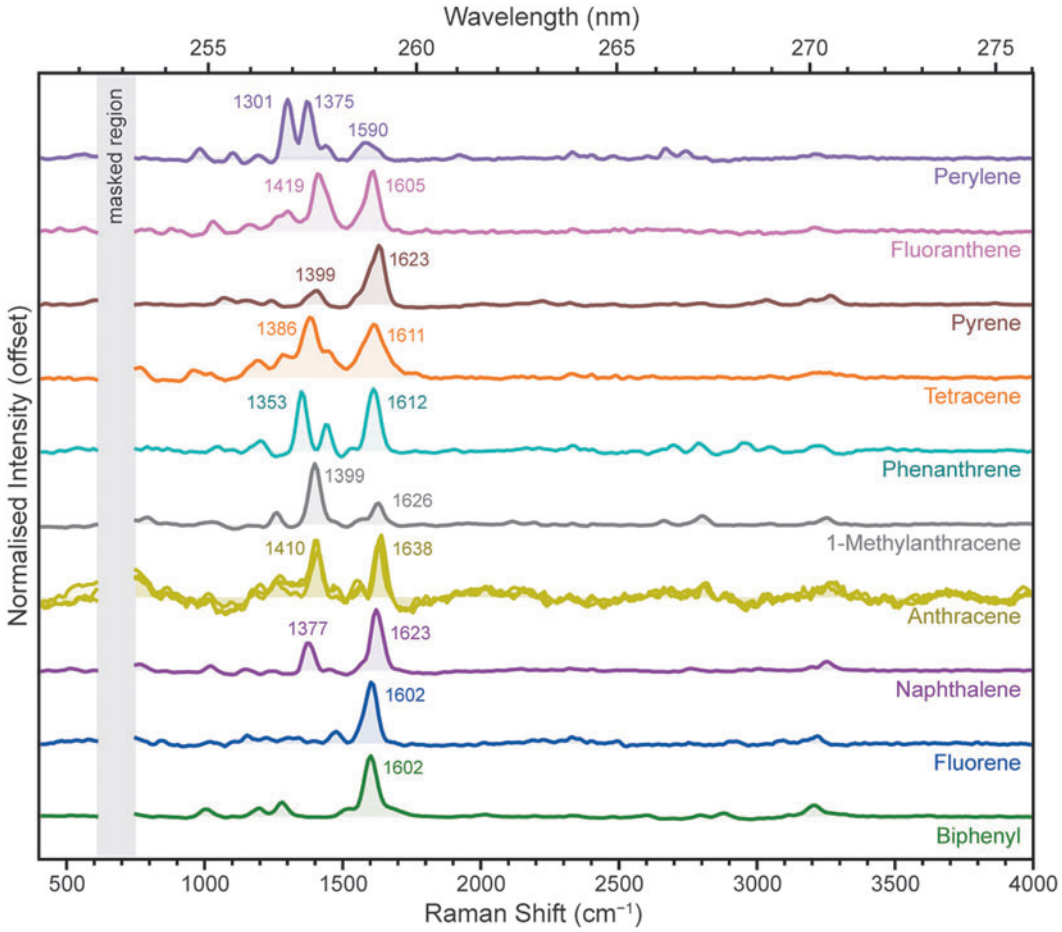

FIG. 7. Normalized DUV Raman spectra for 10 PAHs, measured under Earth ambient conditions.

TABLE 4. SPECTRAL POSITION OF THE FLUORESCENCE MAXIMUM AND MAJOR RAMAN PEAKS OF EACH POLYCYCLIC AROMATIC HYDROCARBON, WITH ASSIGNMENTS TO GENERALIZED VIBRATIONAL MODES

Sample	*Fluor, nm*	*Raman, cm^{-1}*	*Assignment*	*References*
Biphenyl	318	1602	C=C (*S.*)	
Fluorene	338	1602	C=C (*S.*)	
Naphthalene	337	1377	Ring (*S.*)	Loppnow et al (2004)
		1623	C=C (*S.*)	Asher (1984); Loppnow et al (2004)
Anthracene	n.d.	1410	Ring (*S.*)	Asher (1984); Loppnow et al (2004)
		1638	C=C (*S.*)	Asher (1984); Loppnow et al (2004)
1-Methylanthracene	n.d.	1399	Ring (*S.*)	
		1626	C=C (*S.*)	
Phenanthrene	350	1353	Ring (*S.*)	Asher (1984); Lin-Vien et al (1991)
		1612	C=C (*S.*)	Asher (1984); Lin-Vien et al (1991)
Tetracene	n.d.	1386	C=C (*S.*)	Alajtal et al (2010)
		1611	C=C (*S.*)	Alajtal et al (2010)
Pyrene	n.d.	1399	Ring (*S.*)	Asher (1984)
		1623	C=C (*S.*)	Asher (1984)
Fluoranthene	n.d.	1419	Ring (*S.*)	
		1605	C=C (*S.*)	
Perylene	n.d.	1301	Ring (*S.*)	Ong et al (1999)
		1375	Ring (*S.*)	Ong et al (1999)
		1590	C=C (*S.*)	Ong et al (1999)

All Raman peak positions have an estimated uncertainty of ±5 cm^{-1}.

symmetries. Biphenyl, naphthalene, and pyrene all exhibited one or more minor peaks around $\sim$3200 cm^{-1} that are likely second-order overtones of the dominant $\sim$1600 cm^{-1} mode. Other probable overtones can be seen at 2600–2800 cm^{-1} in perylene and phenanthrene, the two PAHs that exhibited the strongest peaks around 1300–1400 cm^{-1}.

The position of the C=C mode at $\sim$1600 cm^{-1} varied by 34 cm^{-1} between the 10 PAHs we studied, but there was no systematic trend in peak position that could be related to the size or geometry of the aromatic system; for example, the linear PAHs (naphthalene, anthracene, and tetracene) exhibited C=C peak positions of 1623, 1638, and 1611 cm^{-1}, respectively. These modes are somewhat shifted compared with values reported under longer wavelength excitation, which may be an indication of Raman dispersion effects due to resonance (Asher, 1984; Cloutis *et al.,* 2016). Despite the lack of systematic trends, the various PAHs can still be easily distinguished by the positions and relative intensities of their other modes between 1000 and 1500 cm^{-1}. For sample spectra that are clearly dominated by a single PAH, it should be possible to identify the compound from the peaks observed.

3.3. Amino acids. We measured a total of 24 α-amino acids with different side-chains, the majority of which were L-stereoisomers, including 20 of the 22 proteinogenic amino acids that form proteins found in terrestrial organisms. The amino acids were grouped based on the structure and functionalization of the side-chain as follows: aliphatic (D-alanine, L-alanine, glycine, L-isoleucine, L-leucine, L-proline, L-valine), alcoholic (L-hydroxyproline, D-serine, L-serine, L-threonine), sulfuric (L-cystine, L-methionine), carboxylic (L-aspartic acid, L-glutamic acid), amidic (L-asparagine, L-glutamine), basic (L-arginine, L-lysine), and aromatic (L-histidine, D-phenylalanine, L-phenylalanine, L-tryptophan, and L-tyrosine).

Figure 8 shows the DUV spectra for each amino acid, acquired with 100 laser pulses at 20 Å (250–350 μJ). To avoid signal saturation, the L-isoleucine spectrum was acquired with 25 pulses at 20 Å (100 μJ), L-phenylalanine and L-tryptophan with 25 pulses at 5 Å (40 μJ), D-phenylalanine with 5 pulses at 20 Å (20 μJ), and L-tyrosine with 5 pulses at 10 Å (13 μJ). As expected, the aromatic amino acids all produced strong fluorescence in this range, with a λ_{max} of 278 nm for L-histidine, 293 nm for D-phenylalanine, 294 nm for L-phenylalanine, 333 nm for L-tryptophan, and 303 nm for L-tyrosine. The fluorescence spectra of D-phenylalanine and L-phenylalanine were almost identical, showing no obvious dependence on the chirality of the α-carbon.

Four nonaromatic amino acids also produced detectable fluorescence despite having no functional groups capable of UV absorption/emission. L-Isoleucine and L-valine exhibited spectra with a λ_{max} $\sim$292 nm, similar to that of D/L-phenylalanine, while L-aspartic acid and L-lysine exhibited spectra with a λ_{max} of 350 nm. Given that the amino acid standards used had minimum purities of 98% and were extracted from cell cultures, we consider this unexpected fluorescence to be produced by trace biological contaminants, probably aromatic amino acids (*e.g.*, phenylalanine) and/or nucleic acids, rather than innate fluorescence from the aliphatic amino acids in question.

The Raman spectra of the amino acids (Fig. 9) were unique according to their chemical structures, but were generally dominated by either C=C stretching modes around 1600 cm^{-1} or C–H stretching modes around 2900–3000 cm^{-1}, with various minor vibrational modes between 800 and 1600 cm^{-1} (Megevand *et al.,* 2021; Rajkumar *et al.,* 1988).

The common structural motifs of all amino acids result in a number of vibrational modes present in most spectra, such as rocking vibrations of the amine group at 1290–1340 cm^{-1}, symmetric stretching vibrations of the carboxylic acid group at 1380–1440 cm^{-1}, and an amine bending mode at $\sim$1600 cm^{-1}. The only spectra that could not be readily distinguished from one another were those of stereoisomer pairs, for example, D-phenylalanine and L-phenylalanine, which appeared identical to one another. The full list of major peak positions for each amino acid is given in Table 5, along with their assignments to particular vibrational modes.

The aliphatic amino acids (D-alanine, L-alanine, glycine, L-isoleucine, L-leucine, L-proline, L-valine) were fairly typical of nonaromatic amino acid spectra, generally dominated by C–H stretching modes around 3000 cm^{-1} with amine and carboxylic acid group modes appearing as minor peaks at 1300–1330 and 1370–1420 cm^{-1}, respectively (Zhu *et al.,* 2011). These peaks were strongest in glycine, which has the fewest C–H bonds of the aliphatic amino acids, while in amino acids with large aliphatic side-chains, such as L-isoleucine and L-leucine, these peaks were obscured by overlap with C–H bending modes (Kumar *et al.,* 2006; Zhu *et al.,* 2011).

L-Leucine and L-isoleucine, which differ only in the structural isomerism of the hydrocarbon side-chain, could be distinguished by the intensity ratio of the two C–H stretching modes at $\sim$2900 and $\sim$2960 cm^{-1}. The 5-membered pyrrolidine ring of L-proline produced an additional low-frequency ring-breathing mode at 898 cm^{-1} (Deveney *et al.,* 1971), while the amine group's stretching vibration could be resolved at 3145 cm^{-1} in glycine, and 3090–3100 cm^{-1} in D-alanine and L-alanine.

The alcoholic amino acids (L-hydroxyproline, D-serine, L-serine, L-threonine) appeared spectrally similar to the aliphatic amino acids. The presence of an alcohol group was generally indicated by the appearance of a relatively weak, broad O–H stretching mode between 3000 and 3500 cm^{-1}. Like L-proline, L-hydroxyproline had a low-frequency ring-breathing mode, down-shifted to 857 cm^{-1} by the added mass of the alcohol group (Deveney *et al.,* 1971). L-Threonine also exhibited three overlapping but distinct C–H stretching modes, at 2873, 2938, and 3007 cm^{-1}.

The sulfuric amino acids (L-cystine, L-methionine) exhibited two very different spectra; the L-cystine spectrum was weak and noisy while L-methionine appeared much like the aliphatic amino acids. The characteristic C–S stretching and C–S–C bending modes of the thioether group in L-methionine could not be observed because their reported positions (650–720 cm^{-1}) are obscured by the NeCu laser's secondary emission line at $\sim$695 cm^{-1} (Zhu *et al.,* 2011). However, we were able to observe the distinctive low-frequency S–S stretching mode of L-cystine at 504 cm^{-1}, as well as a second peak at 989 cm^{-1} that we have assigned to an overtone of the S–S mode (Johnson and Asher, 1987; Zhu *et al.,* 2011). Although the monomer L-cysteine is not included in this study, its thiol group is expected to produce a characteristic S–H stretching mode at $\sim$2580 cm^{-1} (Jenkins *et al.,* 2005).

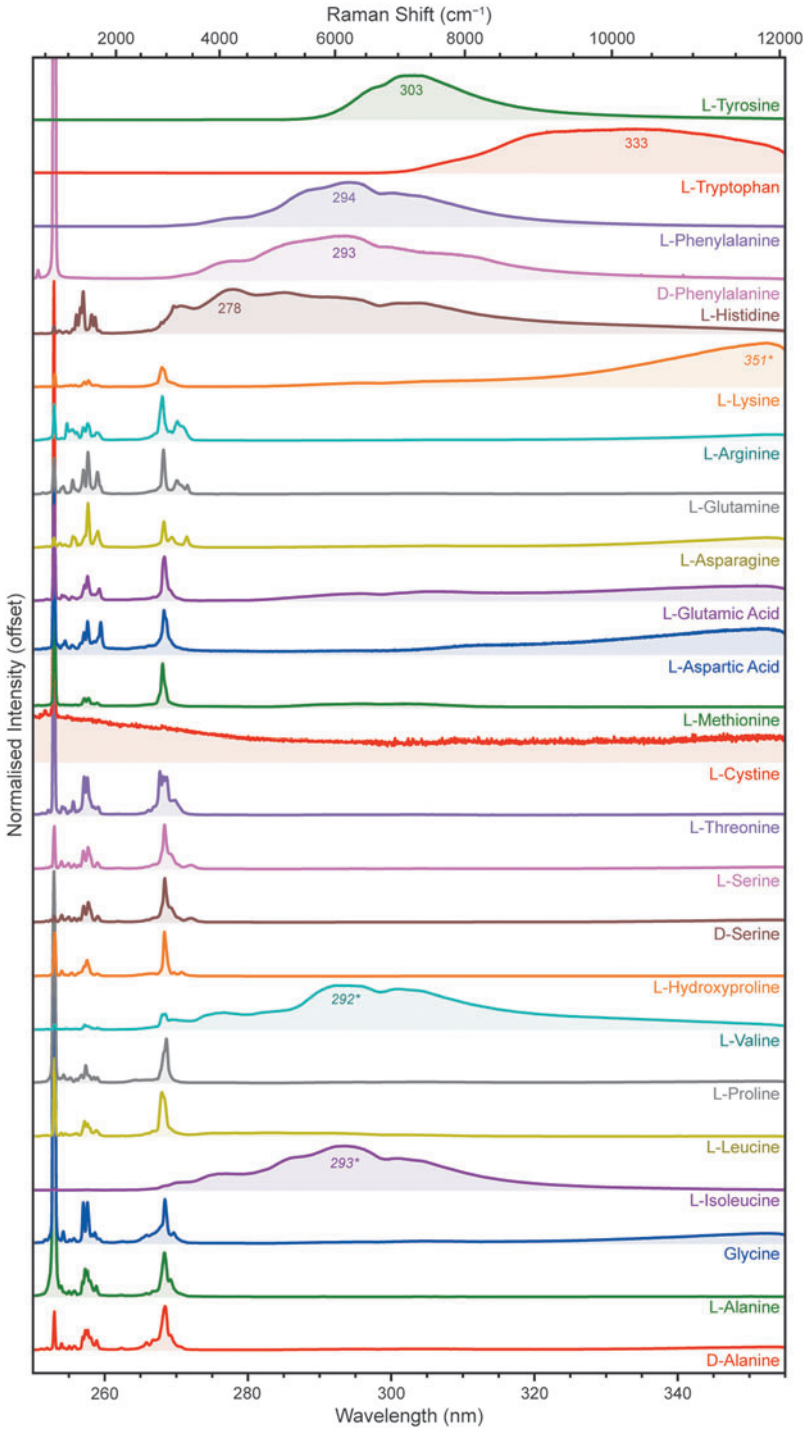

FIG. 8. Normalized DUV fluorescence spectra for 24 amino acids, measured under Earth ambient conditions.

The carboxylic amino acids (L-aspartic acid, L-glutamic acid) could be distinguished by two relatively strong peaks at 1412/1414 and 1685/1653 cm^{-1}, which were assigned to stretching modes of the carboxylic acid group and the C = O bond, respectively, although both are reported to have some contribution from vibrations of the amine group (Guicheteau *et al.,* 2006). The strength of these peaks is consistent with the presence of a second carboxylic acid group in the molecule, but their relative intensity with respect to the skeletal C–C modes at 800–1200 cm^{-1} is significantly greater than that shown in visible Raman spectra of L-aspartic acid and L-glutamic acid (Guicheteau *et al.,* 2006; Zhu *et al.,* 2011). This may indicate some preresonant enhancement for vibrations of the COOH unit.

The amidic amino acids (L-asparagine, L-glutamine) both had unusually broad, strong peaks at 1417/1423 cm^{-1} that were assigned to the C–N stretching mode of the amide group (Guicheteau *et al.,* 2006). L-Asparagine and L-glutamine both had relatively strong C = O stretching modes at 1617 cm^{-1}, consistent with the carbonyl unit of amide rather than a carboxylic acid like L-aspartic acid or L-glutamic acid. The N–H stretching mode of the amide

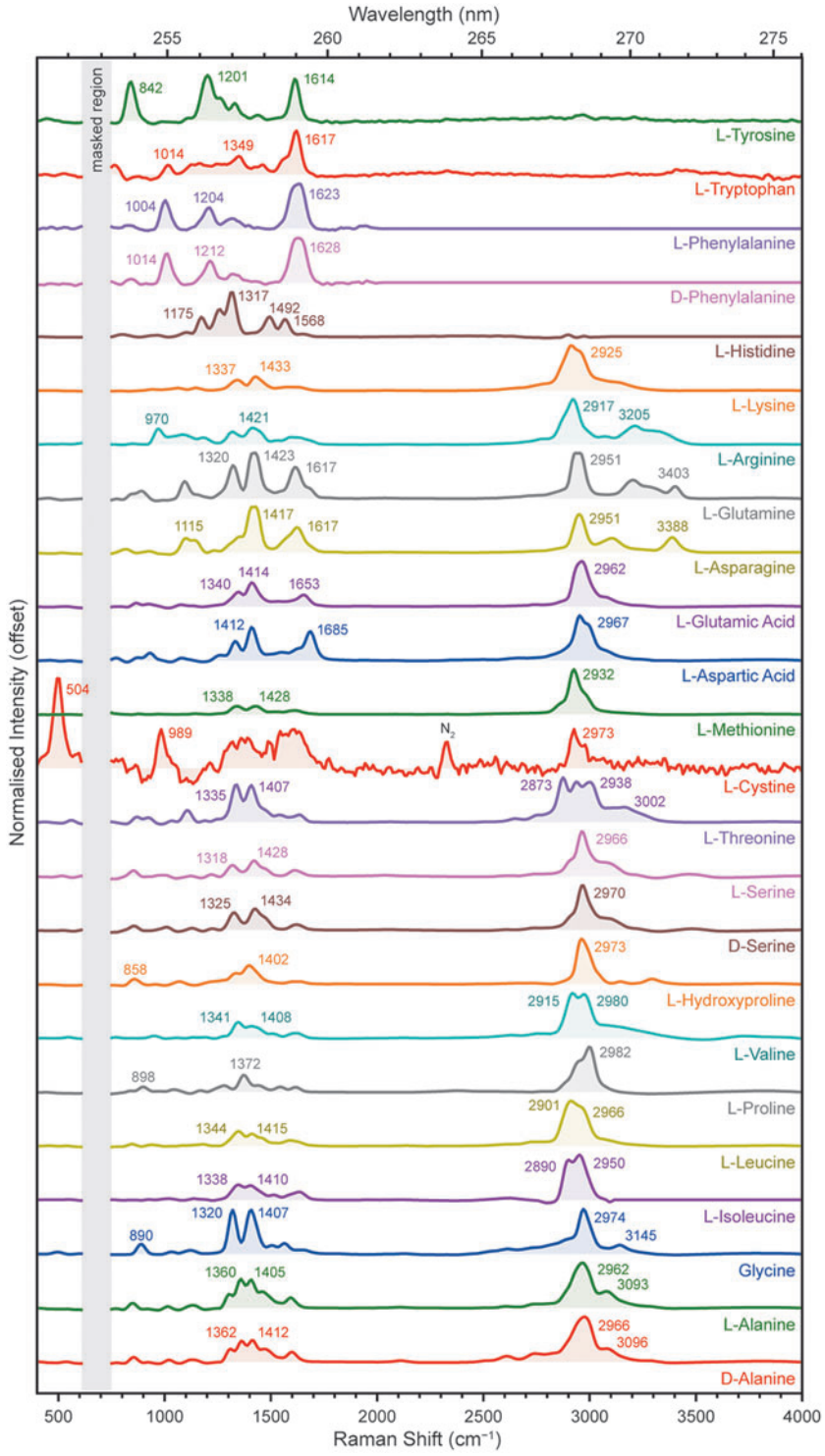

FIG. 9. Normalized DUV Raman spectra for 24 amino acids, measured under Earth ambient conditions. To avoid signal saturation, the L-histidine spectrum was acquired with 100 pulses at 20 Å (381 μJ).

group was also readily apparent in both L-asparagine and L-glutamine as a relatively narrow peak at 3388/3403 cm^{-1} (Moreno *et al.,* 1999).

The basic amino acids (L-arginine, L-lysine) both had two peaks at 1318/1337 and 1421/1433 cm^{-1}, which were the rocking mode of NH_3 and the bending mode of COOH, respectively. The primary amine of L-lysine did not appear to have a dramatic impact on its spectrum, which appeared similar to that of the aliphatic amino acids; however, the presence of guanidino groups in L-arginine was readily evident as two strong N–H stretching modes at 3205 and 3317 cm^{-1}, and a C–N stretching mode at 970 cm^{-1}.

Finally, the aromatic amino acids (L-histidine, D-phenylalanine, L-phenylalanine, L-tryptophan, and L-tyrosine) were dominated by multiple ring-stretching and C = C stretching modes around 800–1600 cm^{-1}, rather than C–H stretching modes around 3000 cm^{-1}. This is consistent with DUV resonance selectively enhancing Raman

Table 5. Spectral Position of the Fluorescence Maximum and Major Raman Peaks of Each Amino Acid, with Assignments to Generalized Vibrational Modes

Sample	*Fluor, nm*	*Raman, cm^{-1}*	*Assignment*	*References*
Glycine	n.d.	1320	NH_3^+ (*R.*)	Zhu et al (2011)
		1407	COO^- (*S.*)	Zhu et al (2011)
		2974	C–H (*S.*)	
D-Alanine	n.d.	1362, 1412	NH_2 (*R.*), CH_3 (*B.*)	Kumar et al (2006); Zhu et al (2011)
		2966	C–H (*S.*)	Kumar et al (2006)
		3096	NH_2 (*S.*)	Kumar et al (2006)
L-Alanine	n.d.	1360, 1405	NH_2 (*R.*), CH_3 (*B.*)	Kumar et al (2006); Zhu et al (2011)
		2962	C–H (*S.*)	Kumar et al (2006)
		3093	NH_2 (*S.*)	Kumar et al (2006)
L-Arginine	n.d.	1421	COOH (*S.*)	Zhu et al (2011)
		2917	C–H (*S.*)	Lima et al (2002)
		3205	NH_2 (*S.*)	Lima et al (2002)
L-Asparagine	n.d.	1417	COOH (*S.*)	Zhu et al (2011)
		1617	C–N (*S.*)	Guicheteau et al (2006)
		2951	C–H (*S.*)	Moreno et al (1998)
L-Aspartic acid	n.d.	1412	COOH (*S.*)	Guicheteau et al (2006); Zhu et al (2011)
		1685	C=O (*S.*)	Guicheteau et al (2006)
		2967	C–H (*S.*)	Rajkumar et al (1988)
L-Cystine	n.d.	504	S–S (*S.*)	Johnson and Asher (1987); Zhu et al (2011)
		989	Overtone	
		2937	C–H (*S.*)	Megevand et al (2021)
L-Glutamine	n.d.	1320	NH_2 (*R.*)	Guicheteau et al (2006)
		1423	C–N (*S.*)	Guicheteau et al (2006)
		2945	C–H (*S.*)	
L-Glutamic acid	n.d.	1340	NH_2 (*R.*)	Zhu et al (2011)
		1414	COOH (*S.*)	Zhu et al (2011)
		1653	C=O (*S.*)	
		2962	C–H (*S.*)	
L-Histidine	278	1317	NH_2 (*R.*)	Zhu et al (2011)
		1492	Imidazole (*B.*)	Zhu et al (2011)
		1568	Imidazole (*B.*)	Zhu et al (2011)
L-Hydroxyproline	n.d.	858	Ring (*S.*)	Deveney et al (1971)
		1402	COO^- (*S.*)	Deveney et al (1971)
		2973	C–H (*S.*)	Deveney et al (1971)
L-Isoleucine	293 (*c.*)	1338, 1410	C–H (*B.*)	Zhu et al (2011)
		2890, 2950	C–H (*S.*)	
L-Leucine	n.d.	1344, 1415	C–H (*B.*)	Zhu et al (2011)
		2901, 2966	C–H (*S.*)	
L-Lysine	351 (*c.*)	1337	NH_2 (*R.*)	Zhu et al (2011)
		1433	COOH (*S.*)	Zhu et al (2011)
		2925	C–H (*S.*)	
L-Methionine	n.d.	1338	NH_2 (*R.*)	Zhu et al (2011)
		1428	C–H (*B.*)	Zhu et al (2011)
		2932	C–H (*S.*)	
D-Phenylalanine	293	1014	Ring (*Br.*)	Asher et al (1986); Jenkins et al (2005)
		1212	Ph–C (*S.*)	Asher et al (1986)
		1628	Ring (*S.*)	Asher et al (1986)
L-Phenylalanine	294	1004	Ring (*Br.*)	Asher et al (1986); Jenkins et al (2005)
		1204	Ph–C (*S.*)	Asher et al (1986)
		1623	Ring (*S.*)	Asher et al (1986)
L-Proline	n.d.	898	Ring (*S.*)	Deveney et al (1971)
		1372	COO^- (*S.*)	Deveney et al (1971)
		2982	C–H (*S.*)	Deveney et al (1971)
D-Serine	n.d.	1325	NH_2 (*R.*)	Zhu et al (2011)
		1434	COOH (*S.*)	Zhu et al (2011)
		2970	C–H (*S.*)	
L-Serine	n.d.	1318	NH_2 (*R.*)	Zhu et al (2011)
		1428	COOH (*S.*)	Zhu et al (2011)
		2966	C–H (*S.*)	
L-Threonine	n.d.	1335	C–H (*B.*)	Zhu et al (2011)
		1407	COO^- (*S.*)	Zhu et al (2011)
		2873, 2938, 3002	C–H (*S.*)	Pawlukoj et al (2001)

(*continued*)

TABLE 5. (CONTINUED)

Sample	*Fluor, nm*	*Raman, cm^{-1}*	*Assignment*	*References*
L-Tryptophan	333	1014	Both rings (*Br.*)	Asher et al (1986); Jenkins et al (2005)
		1349	Pyrrole (*S.*)	Asher et al (1986)
		1617	Phenyl (*S.*)	Asher et al (1986)
L-Tyrosine	303	842	Ring (*Br.*)	Asher et al (1986); Jenkins et al (2005)
		1201	Ring (*S.*)	Asher et al (1986)
		1614	Ring (*S.*)	Asher et al (1986)
L-Valine	292 (*c.*)	1341	NH_2 (*R.*)	Zhu et al (2011)
		1408	COOH (*S.*)	Zhu et al (2011)
		2915, 2980	C–H (*S.*)	

All Raman peak positions have an estimated uncertainty of ±5 cm^{-1}.
B. = bending mode; *Br.* = ring-breathing mode; *c.* = contaminant; *R.* = rocking mode.

scattering for only those vibrations associated with the aromatic rings, leading to aromatic spectra where the ring-stretching vibrations are orders of magnitude stronger than nonresonant vibrations from other parts of the molecule (Asher *et al.,* 1986; Rava and Spiro, 1985).

In addition to the primary C=C stretching mode at ~1600 cm^{-1}, D- and L-phenylalanine exhibited a ring-breathing mode at 1004/1014 cm^{-1}, which was shifted to 842 cm^{-1} in tyrosine due to the addition of the phenolic OH group. The L-tryptophan spectrum exhibited several vibrational modes between 1000 and 1600 cm^{-1} due to the complex symmetry of the bicyclic indole unit: the peak at 1014 cm^{-1} was assigned to the collective breathing mode of both rings, while the peaks at 1349 and 1617 cm^{-1} were assigned to stretching modes of the individual pyrrole and phenyl rings, respectively (Asher *et al.,* 1986).

3.4. Nucleobases. We measured six nucleobases, which could be put in two groups depending on whether they were based on the bicyclic purine structure (purine, adenine, guanine) or the monocyclic pyrimidine structure (cytosine, thymine, uracil). Of these, adenine, guanine, cytosine, and thymine are the four nucleobases of terrestrial DNA, with uracil replacing thymine in RNA. These nucleobases are known to have exceptionally strong UV absorption, with molar absorption coefficients on the order of 5000–20,000 M^{-1} cm^{-1} at our excitation wavelength, 248.6 nm (Dixon *et al.,* 2005). In each case, the nucleobase was characterized as a solid powder.

Figure 10 shows the DUV fluorescence spectra for each nucleobase, acquired with 100 laser pulses at 20 Å (250–350 μJ). All of the nucleobases exhibited detectable fluorescence within SHERLOC's spectral range; adenine and thymine had λ_{max} around 300–310 nm while purine, guanine, cytosine, and uracil had λ_{max} around 340–350 nm. There was no clear distinction between purine and pyrimidine-derived bases in terms of λ_{max}, as both groups have some nucleobases with λ_{max} around 305 nm and some around 345 nm.

This is consistent with what we observed for the aromatic carboxylic acids and PAHs, in which the addition of functional groups to the aromatic unit can result in fairly large (~30 nm) changes in fluorescence wavelength, comparable with the difference in λ_{max} between 1- and 2-ring aromatic

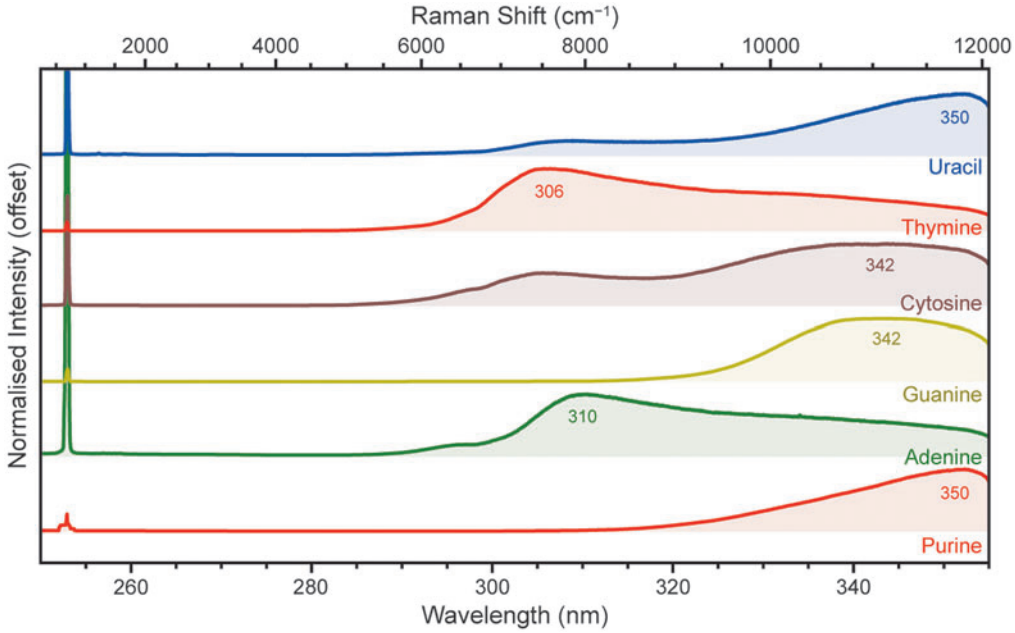

FIG. 10. Normalized DUV fluorescence spectra for six nucleobases, measured under Earth ambient conditions. To avoid signal saturation, the thymine spectrum was acquired with 25 laser pulses (109 μJ).

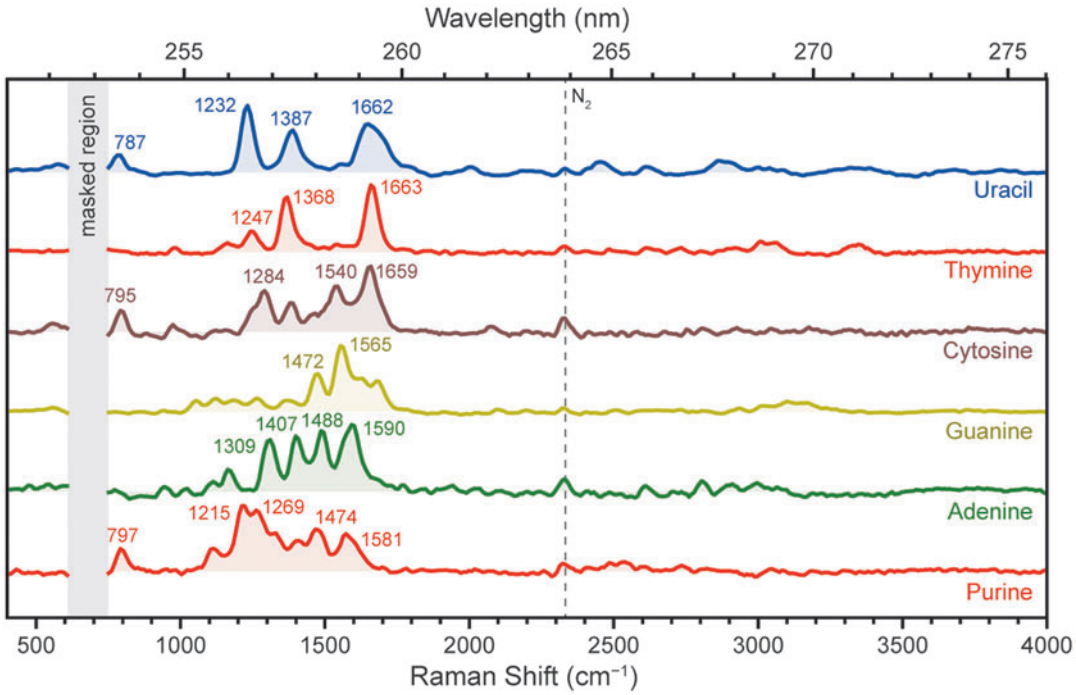

FIG. 11. Normalized DUV Raman spectra for six nucleobases, measured under Earth ambient conditions. The vertical dashed lines indicate the Raman peak of atmospheric N_2.

systems. Multiple nucleobases also exhibited a relatively minor, higher energy shoulder, around 295 nm for adenine, 305 nm for cytosine, and 307 nm for uracil.

Figure 11 shows the Raman spectra for each nucleobase, which were generally dominated by aromatic ring-stretching modes between 1200 and 1600 cm^{-1} due to resonance with the aromatic rings present in all nucleobases, specifically the fused pyridine and imidazole rings in the purine bases and the pyrimidine unit in the pyrimidine bases (Table 6). The purine bases typically exhibited several overlapping peaks of similar intensity, while the pyrimidine bases were typically dominated by two or three major modes, which may be attributed to the larger number of distinct vibrational modes associated with the bicyclic purine unit versus the simpler

Table 6. Spectral Position of the Fluorescence Maximum and Major Raman Peaks of Each Nucleobase, with Assignments to Generalized Vibrational Modes

Sample	*Fluor, nm*	*Raman, cm^{-1}*	*Assignment*	*References*
Purine	350	797	Ring (*Br.*)	Majoube et al (1995)
		1215, 1269	C=N (*S.*)	Majoube et al (1995)
		1474	C=N (*S.*), C–H (*B.*)	Majoube et al (1995)
		1581	C=C (*S.*)	Majoube et al (1995)
Adenine	310	1309	Both rings (*S.*)	Mathlouthi et al (1984); Toyama et al (1994)
		1407	Im. ring (*S.*)	Toyama et al (1994)
		1488	Both rings (*S.*)	Mathlouthi et al (1984); Toyama et al (1994)
		1590	Pyr. ring (*S.*)	Mathlouthi et al (1984); Toyama et al (1994)
Cytosine	342	795	Ring (*Br.*)	
		1284	C–H (*B.*)	D'Amico et al (2015)
		1540	Ring (*S.*)	D'Amico et al (2015)
		1659	C=O (*S.*), NH_2 (*B.*)	D'Amico et al (2015)
Guanine	342	1472	Ring (*S.*)	D'Amico et al (2015); Wen et al (1997)
		1565	Ring (*B.*), C–H (*B.*)	D'Amico et al (2015); Wen et al (1997)
Thymine	306	1247	Ring (*S.*)	Guicheteau et al (2006)
		1368	Ring (*S.*)	Guicheteau et al (2006), D'Amico et al (2015)
		1663	C=O (*S.*), Ring (*S.*)	D'Amico et al (2015)
Uracil	350	787	Ring (*Br.*)	
		1232	Ring (*S.*)	Guicheteau et al (2006); Peticolas and Rush (1995)
		1387	Ring (*S.*)	Guicheteau et al (2006); Peticolas and Rush (1995)
		1662	C=C (*S.*)	Guicheteau et al (2006); Peticolas and Rush (1995)

All Raman peak positions have an estimated uncertainty of ±5 cm^{-1}.
Im. = imidazole ring; Pyr. = pyrrole ring.

pyrimidine unit. However, each individual nucleobase produced a unique spectrum that was very distinct from chemically similar bases (*e.g.*, adenine vs. guanine) due to the impact of different functional groups on the vibrational patterns of the aromatic unit.

The purine spectrum exhibited several overlapping peaks, dominated by a pair of peaks at 1215 and 1269 cm^{-1}, which were assigned to stretching modes of different C=N bonds on the aromatic rings (Majoube *et al.,* 1995). Three additional modes could be distinguished and assigned with confidence: the peak at 797 cm^{-1} corresponded to an out-of-phase bicyclic ring-breathing mode, the peak at 1474 cm^{-1} was assigned to a combined C=N stretching and C–H bending mode, while the asymmetric peak at 1581 cm^{-1} was assigned to a C=C stretching mode (Majoube *et al.,* 1995).

The adenine spectrum exhibited a set of four strong, well-defined peaks at 1309, 1407, 1488, and 1590 cm^{-1}, which were all assigned to collective stretching modes of either one or both of the imidazole (5-member) and pyrimidine (6-member) rings of the purine unit (Mathlouthi *et al.,* 1984; Toyama *et al.,* 1994).

The guanine spectrum exhibited several peaks between 1050 and 1680 cm^{-1}, but only two peaks were strong enough to be distinct, at 1472 and 1565 cm^{-1}. These were assigned, respectively, to a stretching mode of both rings, and a combined pyrimidine ring-stretching and C–H bending mode (D'Amico *et al.,* 2015; Wen *et al.,* 1997).

For the pyrimidine-derived bases, the cytosine spectrum exhibited multiple peaks of medium-to-strong intensity at 1284, 1540, and 1649 cm^{-1}, which were assigned, respectively, to a C–H bending mode, a ring-stretching mode, and a combined C=O stretching and NH_2 deformation mode (D'Amico *et al.,* 2015). There was also a peak at 795 cm^{-1} that was tentatively assigned to the breathing mode of the aromatic ring based on its position.

The thymine spectrum exhibited just two major peaks, at 1368 and 1663 cm^{-1}, which were assigned, respectively, to a ring-stretching mode and a combined in-phase C=O stretching mode of the two carbonyl groups (D'Amico *et al.,* 2015). A relatively weak peak at 1247 cm^{-1} was assigned to a ring-stretching mode (Guicheteau *et al.,* 2006). There were no apparent peaks attributable to the methyl group of thymine.

Finally, the uracil spectrum exhibited three major peaks at 1232, 1387, and 1662 cm^{-1}, with the latter appearing unusually broad and asymmetric—likely indicating convolution with an overlapping peak. The three observed peaks were assigned as follows: ring-stretching mode, ring-stretching mode, and a combined in-phase C=O stretching mode of the two carbonyl groups (Guicheteau *et al.,* 2006; Peticolas and Rush, 1995). A minor peak at 787 cm^{-1} was tentatively assigned to the breathing mode of the aromatic ring.

We note that there are a number of differences between the spectra shown in Fig. 11 and Raman spectra reported in the literature for the same nucleobases (D'Amico *et al.,* 2015; Majoube *et al.,* 1995; Toyama *et al.,* 1994), including those we have previously reported under 248.6 nm excitation (Razzell Hollis *et al.,* 2021; Sapers *et al.,* 2019). The largest differences involve the relative intensities of the various peaks in a particular spectrum, for example, the 1409 cm^{-1} mode of adenine, which was a strong peak in Fig. 11 but reported to be very weak in the literature. These differences may be attributed to the fact that we examined nucleobases in the solid state, while the majority of literature studies referenced above focused on nucleobases in aqueous solution.

3.5. Macromolecular carbons. We measured six macromolecular carbons (also known as kerogens, carbonaceous material, insoluble organic matter, and refractory carbon) that make up the continuum of material produced by geothermal degradation of organic matter, which gradually increase in carbon content, decrease in heteroatomic content, and increase in graphitization as they mature (Galvez *et al.,* 2012; Quirico *et al.,* 2009).

Humic acid (which consists of highly crosslinked aromatic moieties rich in phenolic/carboxylic groups) was used to represent the extreme of thermally immature organic matter, while HOPG (pure graphitic carbon) was used to represent the other extreme of thermally mature material. Four grades of coal obtained from the Pennsylvania State University Coal Bank were also selected to represent intermediate states in the continuum between humic acid and graphite. In order of increasing thermal maturity, the coal standards were: lignite (PSOC 1533), sub-bituminous B (DECS 25), bituminous hvC (DECS 5), and anthracite (PSOC 1468).

The majority of the macromolecular carbons shown in Fig. 12a exhibited little fluorescence within SHERLOC's wavelength range; only bituminous and sub-bituminous coals exhibited a very weak fluorescence band around ~305 nm that was only slightly greater than the dark noise background. In all cases, the strongest feature was a single Raman peak at ~1600 cm^{-1} (258.9 nm), assigned to the G (graphitic) band of macromolecular carbon. Figure 12b shows that in less thermally mature, less graphitic samples the G band was accompanied by a broad, relatively weak peak at ~1400 cm^{-1}, consistent with the D (defective/disordered) band of macromolecular carbon.

The G and D bands are usually interpreted as representing the polyaromatic (more graphite-like) and heteroatomic/aliphatic (more disordered) components of the material, respectively (Schopf *et al.,* 2005). The use of DUV excitation, which is resonant with the UV-absorbing polyaromatic component of the material, enhances Raman scattering from those units, resulting in a far stronger G band compared with reported Raman spectra under visible excitation (Quirico *et al.,* 2020; Quirico *et al.,* 2009). A pair of very weak, very broad peaks around 2700–3300 cm^{-1} were second-order overtones of the primary D and G peaks. A peak observed at 500 cm^{-1} for HOPG was attributed to a very minor laser emission line, due to the highly reflective planar surface of HOPG.

The G and D bands were fitted using Gaussian functions to derive their spectral parameters for each sample. Only one function was needed to adequately fit the D band under DUV excitation, consistent with previous analyses of DUV Raman spectra for carbonaceous materials (Quirico *et al.,* 2020). Fitting showed that with increasing thermal maturity, the D/(D + G) intensity ratio and the full-width half-maximum (FWHM) of the G band decreased (Table 7).

The first trend, decreasing D intensity, is attributed to removal of defects and increasing aromatization, while the second trend, decreasing G band FWHM, indicates the conversion of polyaromatic units into a more organized graphitic lattice; similar trends have been reported for the Raman spectra of carbonaceous materials under both visible

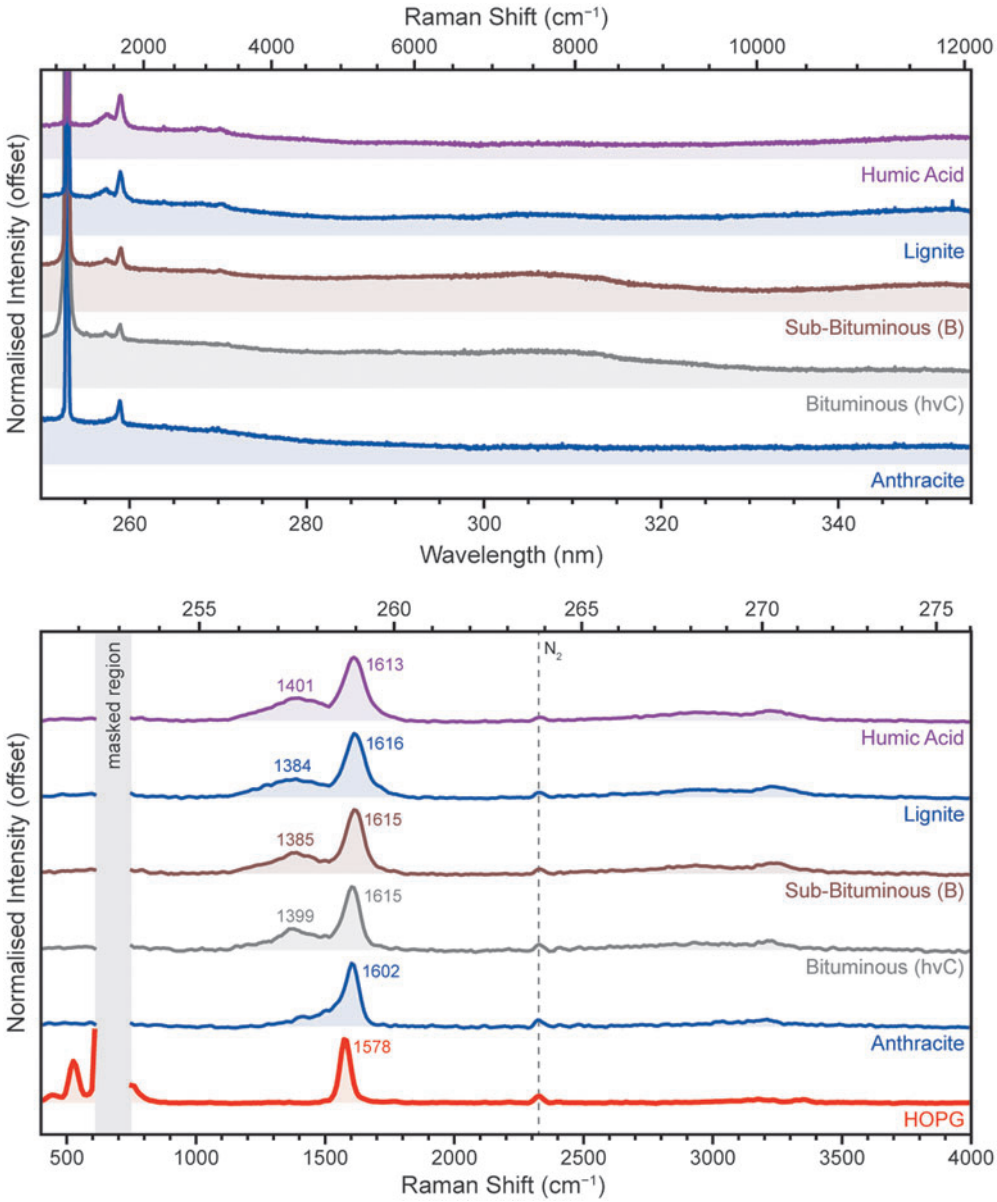

FIG. 12. Normalized DUV spectra (top) and the baselined Raman region (bottom) for four macromolecular carbons, measured under Earth ambient conditions.

TABLE 7. SPECTRAL PROPERTIES OF THE GRAPHITIC (G) AND DEFECTIVE (D) CARBON-STRETCHING MODES FOR EACH MACROMOLECULAR CARBON SAMPLE

Sample	*D band, cm⁻¹*	*G band, cm⁻¹*	*G FWHM, cm⁻¹*	*D/(D + G) intensity ratio*
HOPG	n.d.	1578	48.6	—
Anthracite	n.d.	1602	74.7	—
Bituminous (hvC)	1399	1615	81.5	0.29
Sub-bituminous (B)	1385	1615	85.7	0.28
Lignite	1384	1616	89.0	0.27
Humic acid	1401	1613	95.5	0.34

All Raman peak positions have an estimated uncertainty of ±5 cm^{-1}.

FWHM = full-width half-maximum.

and DUV excitation (Czaja *et al.,* 2009; Delarue *et al.,* 2016; Quirico *et al.,* 2020; Schopf *et al.,* 2005). Those studies show that spectral parameters such as FWHM and D/G intensity ratios can be used to interrogate the geothermal history of macromolecular carbons based on Raman spectra (Quirico *et al.,* 2020), providing additional insight beyond simple detection/identification of carbonaceous material.

4. Discussion

4.1. Implications for organic detection

SHERLOC is capable of detecting organic compounds by their DUV fluorescence and/or their Raman scattering, with fluorescence typically 10^1–10^4 times stronger than Raman scattering when both are observable (see signal yield values in Supplementary Table S1). However, only aromatic organics possess a chromophore capable of absorbing UV

light and producing fluorescence, and we have found that only 1- and 2-ring aromatics fluoresce in the wavelength range that SHERLOC can detect (250–355 nm), but that many of these aromatics have enough fluorescence that they can be easily detected at short exposures similar to a SHERLOC survey scan ($\sim$45 μJ per point).

The detection of larger aromatics, as well as aliphatic organics, will instead depend on Raman scattering signal. We were able to detect Raman peaks from all samples in this study, but due to the generally lower yield of Raman scattering compared with fluorescence it generally required long exposures (more laser pulses) to yield substantial SNR values, but should be detectable during longer SHERLOC acquisitions, such as detail scans ($\sim$4500 μJ per point).

The use of DUV excitation, which is resonant with aromatic organic compounds, significantly increases Raman scattering yield per molecule compared with compounds that are nonresonant, for example, aliphatic organics (Asher, 1993; Asher *et al.,* 1986). Despite the anticipated difference in yield, we still obtained good Raman spectra from both aliphatic and aromatic organics using 800 laser pulses per point (1800–2600 μJ) and many samples from both groups were still detectable at shorter exposures – absolute signal yields are shown in Suplementary Table S1. This phenomenon may be explained by considering the impact of UV absorption by the pure, solid-state organic material, which will influence the volume probed by the DUV laser and thus the number of molecules that can contribute to the overall signal detected.

A previous study showed that for aromatic molecules like adenine, which is an exceptionally strong UV absorber, the measured DUV signal becomes self-absorption-limited at 1 m*M* ($\sim$18 ppm in H_2O) and will no longer increase with concentration (Razzell Hollis *et al.,* 2020). By comparison, nonresonant compounds like aliphatic organics are relatively transparent to UV and their Raman signal remains proportional to concentration. Signal being limited by self-absorption in pure aromatic samples but maximized by laser penetration in pure aliphatic samples go some way to explaining how aliphatic samples (like palmitic acid) provided such strong spectra compared with samples that the DUV instrument is more sensitive to (like adenine).

Because of the dependence on absorption, not just scattering cross-section/quantum yield, absolute signals from pure samples (such as those in Table S1) should not be considered representative of relative detectability for trace organics in the samples that SHERLOC might encounter on Mars. Past analyses of martian meteorites and measurements by Curiosity's SAM instrument in Gale Crater have estimated the bulk organic concentration of martian rocks ranging between 0.07 and 11 ppm (Eigenbrode et al, 2018; Freissinet *et al.,* 2015; Steele *et al.,* 2016). Using the adenine absorption limit as an upper bound, this suggests that any organic signal measured by SHERLOC will not be limited by self-absorption, meaning that aromatic and aliphatic organics will exhibit DUV signals in proportion to their concentrations and relative Raman scattering cross-sections.

SHERLOC's ability to detect organic compounds on Mars will also be affected by UV absorption from the surrounding mineral matrix; certain minerals have been shown to be capable of efficiently absorbing UV light, which will reduce laser penetration and the overall Raman/fluorescence signal of any organics present (Carrier *et al.,* 2019). Accurately estimating detection limits for any given organic compound will require a systematic analysis of absorption coefficients for both the organic in question and the local mineral matrix, at both the excitation wavelength (248.6 nm) and the wavelengths of Raman scattering and/or fluorescence.

4.2. Implications for organic identification

Once SHERLOC has detected organics by their DUV Raman or fluorescence, there is the question of identifying the particular compound(s) producing that signal. While generally speaking, 2-ring aromatics tend to fluoresce at longer wavelengths than 1-ring aromatics, we have observed that λ_{max} and peak shape vary dramatically depending on what functional groups are attached to the aromatic unit. As a result, the potential range of wavelengths attributable to 1-ring aromatics significantly overlaps with that for 2-ring aromatics, and fluorescence wavelength is perhaps best considered a first-order approximation rather than diagnostic of chemical structure or identity.

By comparison, DUV Raman spectra tend to be highly specific to each compound, as the number of Raman peaks and their positions (in wavenumbers) are dictated by the vibrational modes unique to that compound's chemical structure. The sensitivity of DUV Raman to small changes of structure on or near the aromatic unit is evident in the significant variation in spectra between the nucleobases purine, adenine, and guanine, which share a common aromatic unit (Fig. 11). The majority of compounds we have studied exhibit two or more dominant vibrational modes, plus multiple minor modes, providing a distinctive spectral signature that will contribute to the identification of a specific organic, even over chemically similar compounds.

In more general terms, it may be possible to classify unknown organics by their spectra, even if the relevant standards are not currently in the library. The number of major peaks and their general positions can still provide some indication as to what class it belongs to, as shown in Fig. 1. For example, the appearance of two bands around $\sim$1330 and $\sim$1420 cm^{-1}, in combination with at least one peak around $\sim$2950 cm^{-1}, would suggest detection of an aliphatic amino acid based on comparison with available library spectra.

The relative intensity of hydrogen vibrations around 2800–3300 cm^{-1} (*e.g.*, C–H/N–H/O–H stretching modes) versus skeletal vibrations around 800–1800 cm^{-1} (*e.g.*, C–C, C = C, ring-stretching modes) may provide further insight into whether the organic material is more aromatic or aliphatic. We found that aromatic compounds tend to exhibit spectra that are dominated by skeletal modes and have a minimal contribution from hydrogen modes, due to selective resonant enhancement of the skeletal modes of the aromatic unit, while nonresonant aliphatic compounds exhibit spectra that have hydrogen modes of similar or greater intensity to their skeletal modes.

Samples containing a mixture of organic compounds may present a greater challenge to identifying organics with SHERLOC; both abiotic and biological sources of organic material tend to leave behind a wide variety of organic compounds rather than a single, isolated species, as shown by the reported organic inventories of meteorites, martian rocks in Gale Crater, and in microbial fossils (Burton *et al.,* 2012; Eigenbrode *et al.,* 2018; Freissinet *et al.,* 2015; Sapers *et al.,* 2019; Sephton and Botta, 2005; Summons *et al.,*

2021). The Raman spectra of such mixtures tend to be linear combinations of individual spectra for each compound present, weighted by relative concentration and Raman scattering cross-section (Razzell Hollis *et al.,* 2020).

As a result, mixed Raman spectra may be deconvolved into their different contributions, provided that the constituent compounds can be guessed and their Raman scattering cross-sections are not too dissimilar in magnitude; such deconvolutions can provide additional insight into structural complexity and relative concentrations (Razzell Hollis, 2020; Sapers *et al.,* 2019). The challenge of identifying individual molecules in complex mixtures is somewhat mitigated by the selective enhancement of DUV Raman scattering from aromatic organics, which will generally dominate the spectrum over any aliphatic components present: although a microbial cell contains many thousands of different molecules, the DUV Raman spectrum of *Escherichia coli* has been accounted for using only eight aromatics, the five nucleobases of RNA and DNA, and the three aromatic amino acids in protein (Sapers *et al.,* 2019; Wu *et al.,* 2001).

While a rigorous analytical process for identifying/deconvolving unknown samples based on their spectra is possible, it will require extensive further work beyond the scope of this study, but the reference spectra in the Library will provide the groundwork for any future deconvolution of mixed spectra observed by SHERLOC.

5. Conclusions

The Perseverance rover's SHERLOC instrument is a DUV Raman and fluorescence spectrometer capable of detecting and identifying a wide range of organic compounds. Fluorescence offers high-sensitivity detection of aromatic organics at low concentrations, with 1- and 2-ring aromatic units producing detectable fluorescence within SHERLOC's wavelength range, 250–355 nm. Raman scattering generally provides lower signal, requiring longer exposures for detection, but offers detection of both aromatic and aliphatic organics and provides spectra that are highly specific to molecular structure, enabling identification of particular compounds.

This Spectral Library provides the initial reference spectra for identifying organics from SHERLOC measurements taken on Mars, including standards for a number of astrobiologically relevant compounds such as amino acids, nucleobases, and PAHs. Even when considering the spectrum of an unknown compound that is not in the library, it may still be possible to classify it by type (aliphatic amino acid, nucleobase, PAH, etc.) according to its similarity to existing spectra in the library, based on the number of Raman peaks and the pattern of their positions.

Acknowledgments

The work described in this article was partially carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. We thank Universities Space Research Association for the NASA Postdoctoral Program fellowship awarded to Joseph Razzell Hollis, administered by the Universities Space Research Association on behalf of NASA. We also thank the members of the *Mars 2020* science team for useful discussions on the development of this Spectral Library.

Author Disclosure Statement

No competing financial interests exist.

Funding Information

National Aeronautics and Space Administration (80NM 0018D0004), and supported by 107415 Mars 2020 Phase-E grant.

Supplementary Material

Supplementary Figure S1
Supplementary Figure S2
Supplementary Figure S3
Supplementary Figure S4
Supplementary Figure S5
Supplementary Figure S6
Supplementary Figure S7
Supplementary Figure S8
Supplementary Figure S9
Supplementary Figure S10
Supplementary Figure S11
Supplementary Table S1

References

Abbey WJ, Bhartia R, Beegle LW, *et al.* Deep UV Raman spectroscopy for planetary exploration: The search for in situ organics. Icarus 2017;290:201–214; doi: 10.1016/j.icarus.2017.01.039

Alajtal AI, Edwards HGM, Elbagerma MA, *et al.* The effect of laser wavelength on the Raman spectra of phenanthrene, chrysene, and tetracene: Implications for extra-terrestrial detection of polyaromatic hydrocarbons. Spectrochim Acta A Mol Biomol Spectrosc 2010;76:1–5; doi: 10.1016/j.saa.2010.01.009

Asher SA. Ultraviolet resonance Raman spectrometry for detection and speciation of trace polycyclic aromatic hydrocarbons. Anal Chem 1984;56(4):720–724; doi: 10.1021/ac00268a029

Asher SA. UV resonance Raman spectroscopy for analytical, physical, and biophysical chemistry. Anal Chem 1993;65(4):201A–210A; doi: 10.1021/ac00052a715

Asher SA, Ludwig M, Johnson CR. UV resonance Raman excitation profiles of the aromatic amino acids. J Am Chem Soc 1986;108(12):3186–3197; doi: 10.1021/ja00272a005

Azumi T, McGlynn SP. Polarization of the luminescence of phenanthrene. J Chem Phys 1962;37(10):2413–2420; doi: 10.1063/1.1733019

Beaty DW, Grady MM, McSween HY, *et al.* The potential science and engineering value of samples delivered to Earth by Mars sample return: International MSR Objectives and Samples Team (iMOST). Meteorit Planet Sci 2019;54(3):S3–S152; doi: 10.1111/maps.13242

Benner SA, Devine KG, Matveeva LN, *et al.* The missing organic molecules on Mars. Proc Natl Acad Sci U S A 2000;97(6):2425–2430; doi: 10.1073/pnas.040539497

Bhartia R, Hug WF, Salas EC, *et al.* Classification of organic and biological materials with deep ultraviolet excitation. Appl Spectrosc 2008;62(10):1070–1077; doi: 10.1366/000370208786049123

Bhartia R, Beegle LW, DeFlores L, *et al.* Perseverance's Scanning Habitable Environments with Raman and Luminescence for Organics and Chemicals (SHERLOC) investigation. Space Sci Rev 2021;217:58; doi: 10.1007/s11214-021-00812-z

Burton AS, Stern JC, Elsila JE, *et al.* Understanding prebiotic chemistry through the analysis of extraterrestrial amino acids and nucleobases in meteorites. Chem Soc Rev 2012;41(16): 5459–5472; doi: 10.1039/c2cs35109a

Carrier BL, Abbey WJ, Beegle LW, *et al.* Attenuation of ultraviolet radiation in rocks and minerals: Implications for Mars science. J Geophys Res Planets 2019;124(10):2599–2612; doi: 10.1029/2018JE005758

Clairemidi J, Moreels G, Mousis O, *et al.* Identification of anthracene in Comet 1P/Halley. Astron Astrophys 2008;492(1): 245–250; doi: 10.1051/0004-6361:200809497

Cloutis E, Szymanski P, Applin D, *et al.* Identification and discrimination of polycyclic aromatic hydrocarbons using Raman spectroscopy. Icarus 2016;274:211–230; doi: 10.1016/j.icarus.2016.03.023

Czaja AD, Kudryavtsev AB, Cody GD, *et al.* Characterization of permineralized kerogen from an Eocene fossil fern. Org Geochem 2009;40(3):353–364; doi: 10.1016/j.orggeochem.2008.12.002

D'Amico F, Cammisuli F, Addobbati R, *et al.* Oxidative damage in DNA bases revealed by UV resonant Raman spectroscopy. Analyst 2015;140:1477–1485; doi: 10.1039/C4AN02364A

Dartnell LR, Patel MR, Storrie-Lombardi MC, *et al.* Experimental determination of photostability and fluorescence-based detection of PAHs on the martian surface. Meteorit Planet Sci 2012;47(5):806–819; doi: 10.1111/j.1945-5100.2012.01351.x

Delarue F, Rouzaud JN, Derenne S, *et al.* The Raman-derived carbonization continuum: A tool to select the best preserved molecular structures in Archean kerogens. Astrobiology 2016;16(6):407–417; doi: 10.1089/ast.2015.1392

Deveney MJ, Walton AG, Koenig JL. Raman spectra of amino acids and poly-L-hydroxyproline. Biopolymers 1971;10(4): 615–630; doi: 10.1002/bip.360100403

Dixon JM, Taniguchi M, Lindsey JS. PhotochemCAD 2: A refined program with accompanying spectral databases for photochemical calculations. Photochem Photobiol 2005; 81(1):212–213; doi: 10.1562/2004-11-06-TSN-361

Eigenbrode JL, Summons RE, Steele A, *et al.* Organic matter preserved in 3-billion-year-old mudstones at Gale crater, Mars. Science 2018;360(6393):1096–1101; doi: 10.1126/science.aas9185

Eshelman E, Daly MG, Slater G, *et al.* Detecting aromatic compounds on planetary surfaces using ultraviolet time-resolved fluorescence spectroscopy. Planet Space Sci 2018; 151:1–10; doi: 10.1016/j.pss.2017.09.003

Farley KA, Williford KH, Stack KM, *et al.* Mars 2020 mission overview. Space Sci Rev 2020;216(8);142: doi: 10.1007/s11214-020-00762-y

Freissinet C, Glavin DP, Mahaffy PR, *et al.* Organic molecules in the Sheepbed Mudstone, Gale crater, Mars. J Geophys Res Planets 2015;120(3):495–514; doi: 10.1002/2014JE004737

Galvez ME, Beyssac O, Benzerara K, *et al.* Morphological preservation of carbonaceous plant fossils in blueschist metamorphic rocks from New Zealand. Geobiology 2012; 10(2):118–129; doi: 10.1111/j.1472-4669.2011.00316.x

Guicheteau J, Argue L, Hyre A, *et al.* Raman and surface-enhanced Raman spectroscopy of amino acids and nucleotide bases for target bacterial vibrational mode identification. Proc SPIE 2006;6218:62180O; doi: 10.1117/12.670294

Jenkins AL, Larsen RA, Williams TB. Characterization of amino acids using Raman spectroscopy. Spectrochim Acta A Mol Biomol Spectrosc 2005;61(7):1585–1594; doi: 10.1016/j.saa.2004.11.055

Johnson CR, Asher SA. UV resonance Raman excitation profiles of L-cystine. J Raman Spectrosc 1987;18(5):345–349; doi: 10.1002/jrs.1250180509

Jones E, Oliphant E, Peterson P, *et al.* SciPy: Open source scientific tools for Python. SciPy; 2001. Available from: www.scipy.org [Last accessed: July 20, 2018].

Kumar S, Kumar Rai A, Rai SB, *et al.* Infrared, Raman and electronic spectra of alanine: A comparison with ab intio calculation. J Mol Struct 2006;791(1–3):23–29; doi: 10.1016/j.molstruc.2006.01.004

Lakowicz JR. Principles of Fluorescence Spectroscopy, 3rd ed. Springer: New York; 2006.

Lima RJC, Freire PTC, Sasaki JM, *et al.* Temperature-dependent Raman study of L-arginine hydrochloride monohydrate single crystal. J Raman Spectrosc 2002;33(8): 625–630; doi: 10.1002/jrs.890

Lin-Vien D, Colthup NB, Fateley WG, *et al.* The Handbook of Infrared and Raman Characteristic Frequencies of Organic Molecules. Academic Press: Boston; 1991.

Long DA. Raman Spectroscopy. McGraw-Hill: New York; 1977.

Loppnow GR, Shoute L, Schmidt KJ, *et al.* UV Raman spectroscopy of hydrocarbons. Philos Trans R Soc A Math Phys Eng Sci 2004;362(1864):2461–2476; doi: 10.1098/rsta.2004.1449

Majoube M, Millié P, Chinsky L, *et al.* Resonance Raman spectra for purine. J Mol Struct 1995;355(2):147–158; doi: 10.1016/0022-2860(95)08896-4

Mathlouthi M, Seuvre AM, Koenig JL. FTIR and laser-Raman spectra of adenine and adenosine. Carbohydr Res 1984; 131(1):1–15; doi: 10.1016/0008-6215(84)85398-7

McDonald GD, de Vanssay E, Buckley JR. Oxidation of organic macromolecules by hydrogen peroxide: Implications for stability of biomarkers on Mars. Icarus 1998;132(1):170–175; doi: 10.1006/icar.1998.5896

Megevand V, Viennet JC, Balan E, *et al.* Impact of UV radiation on the Raman signal of cystine: Implications for the detection of S-rich organics on Mars. Astrobiology 2021; 21(5):566–574; doi: 10.1089/ast.2020.2340

MEPAG E2E-iSAG. Planning for Mars returned sample science: Final report of the MSR End-to-End International Science Analysis Group (E2E-iSAG). Astrobiology 2012; 12(3):175–230; doi: 10.1089/ast.2011.0805

Moores JE, Schuerger AC. UV degradation of accreted organics on Mars: IDP longevity, surface reservoir of organics, and relevance to the detection of methane in the atmosphere. J Geophys Res Planets 2012;117(8):1–14; doi: 10.1029/2012JE004060

Moreno AJD, Freire PTC, Guedes I, *et al.* Raman scattering studies of monohydrated L-asparagine. Brazilian J Phys 1999; 29(2):380–387; doi: 10.1590/S0103-97331999000200019

Newville M, Stensitzki T, Allen DB, *et al.* LMFIT: Non-linear least-square minimization and curve-fitting for Python. 2014. Available from: https://lmfit.github.io/lmfit-py/index.html [Last accessed: March 10, 2018].

Ong KK, Jensen JO, Hameka HF. Theoretical studies of the infrared and Raman spectra of perylene. J Mol Struct Theochem 1999;459(1–3):131–144; doi: 10.1016/S0166-1280(98)00266-8

Osterrothová K, Jehlička J. Raman spectroscopic identification of phthalic and mellitic acids in mineral matrices. Spectrochim Acta A Mol Biomol Spectrosc 2010;77(5):1092–1098; doi: 10.1016/j.saa.2010.08.081

Pawlukoj A, Leciejewicz J, Tomkinson J, *et al.* Neutron scattering, infrared, Raman spectroscopy and ab initio study of L-threonine. Spectrochim Acta A Mol Biomol Spectrosc 2001; 57(12):2513–2523; doi: 10.1016/S1386-1425(01)00508-X

Peticolas WL, Rush T. Ab initio calculations of the ultraviolet resonance Raman spectra of uracil. J Comput Chem 1995; 16(10):1261–1270; doi: 10.1002/jcc.540161008

Pizzarello S, Cooper GF, Flynn G. The nature and distribution of the organic material in carbonaceous chondrites and interplanetary dust particles. In: Meteorites and the Early Solar System II. (Lauretta DS, McSween HY. eds.) University of Arizona Press: Tucson, AZ; 2006; pp. 625–651.

Quirico E, Bonal L, Montagnac G, *et al.* New insights into the structure and formation of coals, terrestrial and extraterrestrial kerogens from resonant UV Raman spectroscopy. Geochim Cosmochim Acta 2020;282:156–176; doi: 10.1016/j.gca.2020.05.028

Quirico E, Montagnac G, Rouzaud JN, *et al.* Precursor and metamorphic condition effects on Raman spectra of poorly ordered carbonaceous matter in chondrites and coals. Earth Planet Sci Lett 2009;287(1–2):185–193; doi: 10.1016/j.epsl.2009.07.041

Rajkumar BJM, Ramakrishnan V, Rajaram RK. Infrared and Raman spectra of DL-aspartic acid nitrate monohydrate. Spectrochim Acta A Mol Biomol Spectrosc 1998;54(10): 1527–1532; doi: 10.1016/S1386-1425(98)00198-X

Rava RP, Spiro TG. Resonance enhancement in the ultraviolet Raman spectra of aromatic amino acids. J Phys Chem 1985; 89(10):1856–1861; doi: 10.1021/j100256a007

Razzell Hollis J, Abbey W, Beegle LW, *et al.* A deep-ultraviolet Raman and fluorescence spectral library of 62 minerals for the SHERLOC instrument onboard Mars 2020. Planet Space Sci 2021;209:105356; doi: 10.1016/j.pss.2021.105356

Razzell Hollis J, Rheingold D, Bhartia R, *et al.* An optical model for quantitative Raman microspectroscopy. Appl Spectrosc 2020;74(6):684–700; doi: 10.1177/0003702819895299

Sapers HM, Razzell Hollis J, Bhartia R, *et al.* The cell and the sum of its parts: Patterns of complexity in biosignatures as revealed by deep UV Raman spectroscopy. Front Microbiol 2019;10:679; doi: 10.3389/fmicb.2019.00679

Schmitt-Kopplin P, Gabelica Z, Gougeon RD, *et al.* High molecular diversity of extraterrestrial organic matter in Murchison meteorite revealed 40 years after its fall. Proc Natl Acad Sci U S A 2010;107(7):2763–2768; doi: 10.1073/pnas.0912157107

Schopf JW, Kudryavtsev AB, Agresti DG, *et al.* Raman imagery: A new approach to assess the geochemical maturity and biogenicity of permineralized Precambrian fossils. Astrobiology 2005;5(3):333–371; doi: 10.1089/ast.2005.5.333

Sephton MA, Botta O. Recognizing life in the Solar System: Guidance from meteoritic organic matter. Int J Astrobiol 2005;4(3–4):269–276; doi: 10.1017/S1473550405002806

Socrates G. Infrared and Raman Characteristic Group Frequencies: Tables and Charts. John Wiley & Sons: Chichester, United Kingdom; 2001.

Steele A, McCubbin FM, Fries MD. The provenance, formation, and implications of reduced carbon phases in martian meteorites. Meteorit Planet Sci 2016;51(11):2203–2225; doi: 10.1111/maps.12670

Summons R, Welander PV, Gold DA. Lipid biomarkers: Molecular tools for illuminating the history of microbial life. Nat Rev Microbiol 2021;20(3):174–185; doi: 10.1038/s41579-021-00636-2

Toyama A, Hanada N, Abe Y, *et al.* Assignment of adenine ring in-plane vibrations in adenosine on the basis of ^{15}N and ^{13}C isotopic frequency shifts and UV resonance Raman enhancement. J Raman Spectrosc 1994;25(7–8):623–630; doi: 10.1002/jrs.1250250724

Uckert K, Bhartia R, Beegle LW, *et al.* Calibration of the SHERLOC deep UV fluorescence/Raman spectrometer on the Perseverance rover. Appl Spectrosc 2021;75(7):763–773; doi: 10.1177/00037028211013368

Uckert K, Bhartia R, Michel J. A semi-autonomous method to detect cosmic rays in Raman hyperspectral data sets. Appl Spectros 2019;73(9):1019–1027; doi: 10.1177/0003702819850584

van der Walt S, Colbert SC, Varoquaux G. The NumPy array: A structure for efficient numerical computation. Comput Sci Eng 2011;13(2):22–30; doi: 10.1109/MCSE.2011.37

Wen ZQ, Overman SA, Thomas GJ. Structure and interactions of the single-stranded DNA genome of filamentous virus fd: Investigation by ultraviolet resonance Raman spectroscopy. Biochemistry 1997;36(25):7810–7820; doi: 10.1021/bi970342q

Williford KH, Farley KA, Stack KM, *et al.* The NASA Mars 2020 rover mission and the search for extraterrestrial life. In: From Habitability to Life on Mars. (Cabrol NA, Grin EA. eds.) Elsevier: Amsterdam; 2018; pp. 275–308.

Wu Q, Hamilton T, Nelson WH, *et al.* UV Raman spectral intensities of *E. coli* and other bacteria excited at 228.9, 244.0, and 248.2 nm. Anal Chem 2001;73(14):3432–3440; doi: 10.1021/ac001268b

Zhu G, Zhu X, Fan Q, *et al.* Raman spectra of amino acids and their aqueous solutions. Spectrochim Acta A Mol Biomol Spectrosc 2011;78(3):1187–1195; doi: 10.1016/j.saa.2010.12.079

Address correspondence to:
Joseph Razzell Hollis
Department of Life Sciences
The Natural History Museum
Exhibition Road
South Kensington
London SW7 5HD
United Kingdom

E-mail: joseph.razzell-hollis1@nhm.ac.uk

Submitted 15 February 2022
Accepted 1 August 2022
Associate Editor: Michael C. Storrie-Lombardi

Abbreviations Used

AA = amino acid
B. = bending mode
Br. = ring-breathing mode
c. = contaminant
CCD = charge-coupled device
DECS = Department of Energy Coal Sample
DNA = deoxyribonucleic acid
DUV = deep ultraviolet
FWHM = full-width half-maximum
HOPG = highly oriented pyrolytic graphite
Im. = imidazole ring
n.d. = not detected
PAH = polycyclic aromatic hydrocarbon
PSOC = Penn. State/Office of Coal Research
Pyr. = pyrrole ring
R. = rocking mode
RNA = ribonucleic acid
S. = stretching mode
SHERLOC = Scanning Habitable Environments with Raman and Luminescence for Organics and Chemicals
SNR = signal:noise ratio
UV = ultraviolet