nature astronomy

Supplementary information

https://doi.org/10.1038/s41550-025-02765-7

An interstellar energetic and non-aqueous pathway to peptide formation

In the format provided by the authors and unedited

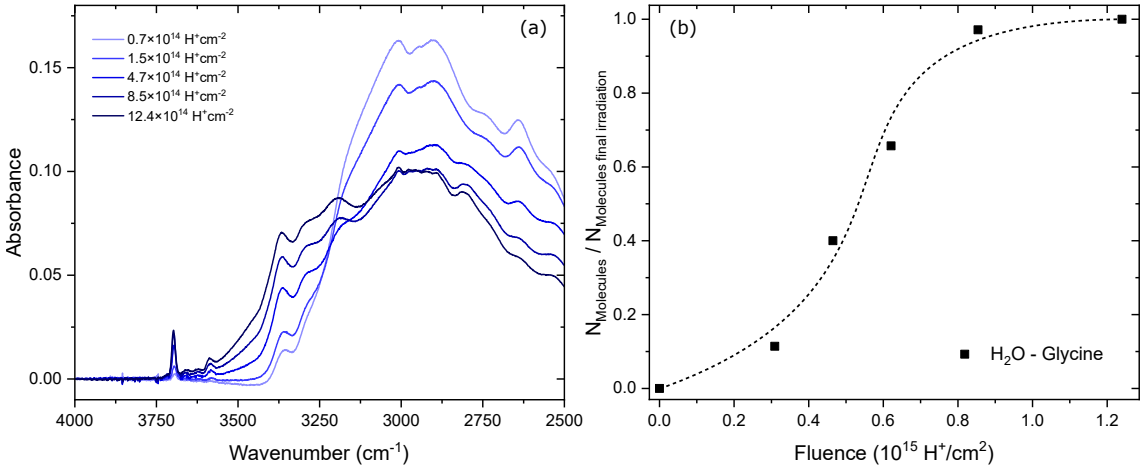
Supplementary Information

1 MeV proton irradiation and molecule formation

Glycine samples were irradiated with 1 MeV protons within the ICA chamber to investigate chemical transformations induced by high energy radiation. Such processing is expected to promote reactions analogous to those observed under lower-energy proton exposure, notably the formation of amide-bearing species and peptide bonds. Although initially masked by the strong infrared absorption features intrinsic to glycine in the relevant spectral region, the emergence of water was detected during the irradiation sequence. This water is plausibly formed via mechanisms similar to those proposed under 10 keV proton exposure, involving the stepwise degradation of glycine and the simultaneous formation of peptide linkages. In the later stages of irradiation, the infrared signature of water becomes increasingly prominent, as illustrated in the Supplementary Figure 1a, which shows a broad absorption band centred near 3350 cm^{-1}, corresponding to the symmetric and asymmetric OH stretching vibrations [1]. The intensity of this band increases with the cumulative proton fluence. While the appearance of water cannot be unambiguously ascribed to peptide-bond formation alone, the observed correlation supports the hypothesis that high-energy proton irradiation facilitates peptide synthesis, in line with the mechanistic insights discussed in the main text.

Supplementary Figure 1b shows the normalised number of water molecules detected in high energy irradiation. This was done by scaling a reference water spectrum to the irradiated spectra to give an estimate of the formation rate of the water at different proton fluences. As can be seen, the formation is low at first, but increases before slowing at the final irradiation steps. The profile of this formation is much more similar to that of water formation for D3 glycine shown in the main text and may indicate that higher energy protons destroy glycine more thoroughly and cause relatively more two-step water formation and less peptide formation than lower energy protons. Therefore, solar wind protons may engender peptide formation to a greater extent than galactic cosmic rays.

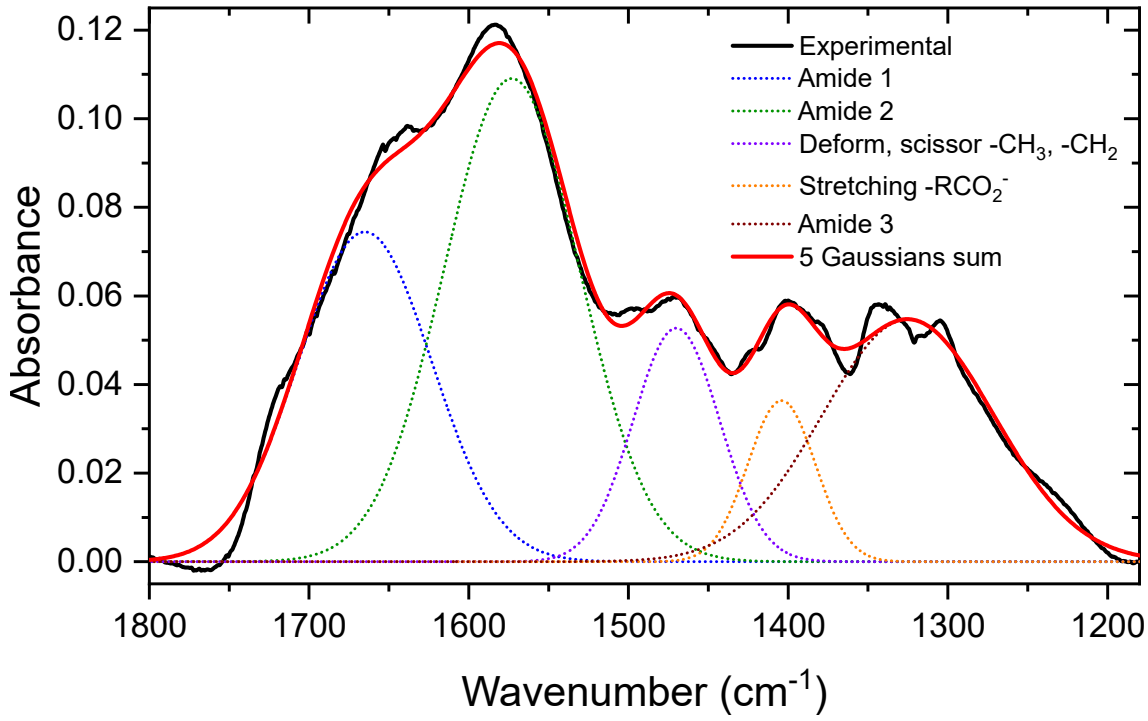
The fluence of the high energetic processing ended at a maximum of 1.24×10^{15} H^+cm^{-2} at which point the rate of water formation and processing has slowed. It is assumed that the flux of cosmic rays inside a dense molecular could can be approximated with an equivalent flux of 1 MeV protons of 1 proton $cm^{-2}s^{-1}$ [2, 3]. The lifetime for a dense cloud is on the order of 3×10^7 years which would give a fluence over its lifetime of 9.5×10^{14} H^+cm^{-2}. This is very close to the end point of the fluence as shown in Supplementary Figure 1 indicating the endpoint of the irradiation and its processing is similar to what could be expected in the ISM. If the time from collapse to star formation is included and estimated that the total lifetime is on the order of 10^8 years it gives a upper fluence of 3×10^{16} H^+cm^{-2}. The range then for glycine to experience processing from high energy energetic sources is within the range of the endpoint of the experiments where the water production had started to end and so is comparable to the ISM.



Supplementary Figure 1. IR spectra of glycine irradiated with 1 MeV protons and the rate of water formation from this. **Left panel:** IR spectra of glycine that has been irradiated with 1 MeV protons up to 1.24×10^{15} H^+cm^{-2}. It shows the spectra from 4000 to 2500 cm^{-1} focusing on the increasing absorbance of the O–H symmetric and anti-symmetric stretching modes around 3350 cm^{-1}. **Right panel:** The number of H_2O molecules formed during irradiation of glycine normalised to the final irradiation yield are plotted as a function of the fluence of 1 MeV proton bombardment. Line added to guide the eye.

Irradiation with 1 MeV protons also resulted in the formation of a residue, which was subsequently analysed by infrared (IR) spectroscopy, as shown in Figure 5 in the main text of the article. This IR spectrum was deconvoluted using six Gaussian components. In contrast, the spectrum presented in the Supplementary Figure 2 was fit using five Gaussian bands. The

primary difference between these two fits lies in the inclusion or omission of the NH_2 scissoring mode. This mode has been inconsistently included in previous studies [4–6], and its inclusion here allows a comparative analysis of its influence on spectral fitting. Although omitting the NH_2 scissoring band results in a slightly poorer visual agreement with the experimental spectrum, it does not significantly alter the positions of the remaining five band centres. These band positions are listed in Supplementary Table 1 and remain consistent with those reported in related studies [4–6], supporting the interpretation that they are characteristic of amide bond formation. However, while these spectral features are indicative of peptide-like structures, their presence alone remains insufficient for a definitive identification of peptide-bond synthesis.



Supplementary Figure 2. IR spectra of the residue obtained after processing glycine with 1 MeV protons with a fluence of 1.24×10^{15} H^+cm^{-2}. The residue spectrum is fitted with five Gaussians of different IR bands based off of fits found in previous work [4].

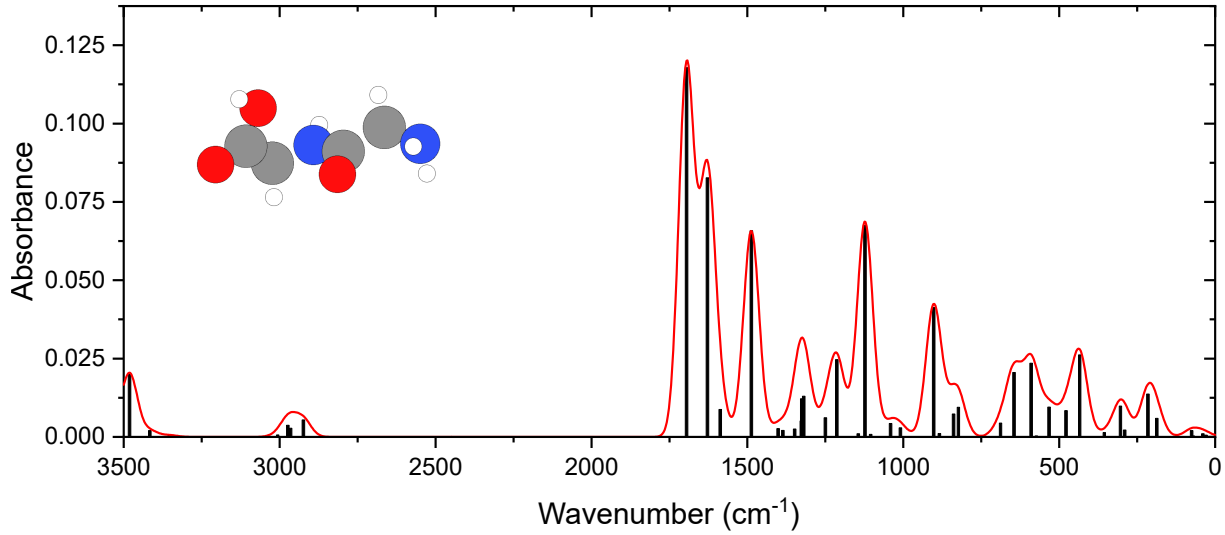
Supplementary Table 1. Tentative assignment of IR bands of the residue obtained upon exposing glycine to 1.24×10^{15} 1 MeV H^+/cm^2.

Band assignment	Band centre (cm^{-1}) Fig. 5	Band centre (cm^{-1}) Supplementary Fig. 2	Literature band center [4]	Literature band center [5]
Amide 1	1665	1665	1673	1666-1670
Scissor NH_2	1599	-	-	1592-1604
Amide 2	1574	1573	1573	1548-1569
Deform, scissor $-CH_3$, $-CH_2$	1475	1470	1480	-
Stretching $-RCO_2^-$	1402	1404	1405	-
Amide 3	1328	1325	1319	-

Glycylglycine calculated infrared spectrum

The vibrational bands associated with peptides, particularly amide modes, are well characterised in the literature [4–6]. To complement these experimental observations, we calculated the theoretical IR spectrum of glycylglycine using the methodology described in the Methods section. This simulated spectrum is only directly applicable to gas phase peptides, however, only small deviations are induced in such spectra through the introduction of physisorption onto a surface and thus these gas phase spectra provide relevant insight into the surface state. The computed spectrum is shown in Supplementary Figure 3, with the

predicted band positions indicated by black lines and overlaid red Gaussian profiles to aid in visual interpretation. In particular, several calculated features closely correspond to the experimentally observed bands. Of particular interest is the spectral region from 1800 to 1000 cm^{-1}, where prominent absorption features characteristic of peptide (amide) bonds appear. These bands, also observed in the experimental residue spectra (see the main text and Supplementary Information), are consistent with prior assignments in the literature. There are some differences between the bands shown and those detected, but this is likely due to contributions from other molecules on the surface that make precise band positions difficult to determine. The agreement between the computed and measured peak positions supports the assignment of these features to molecules containing amide bonds, such as glycylglycine, and further confirms peptide formation under the investigated conditions.

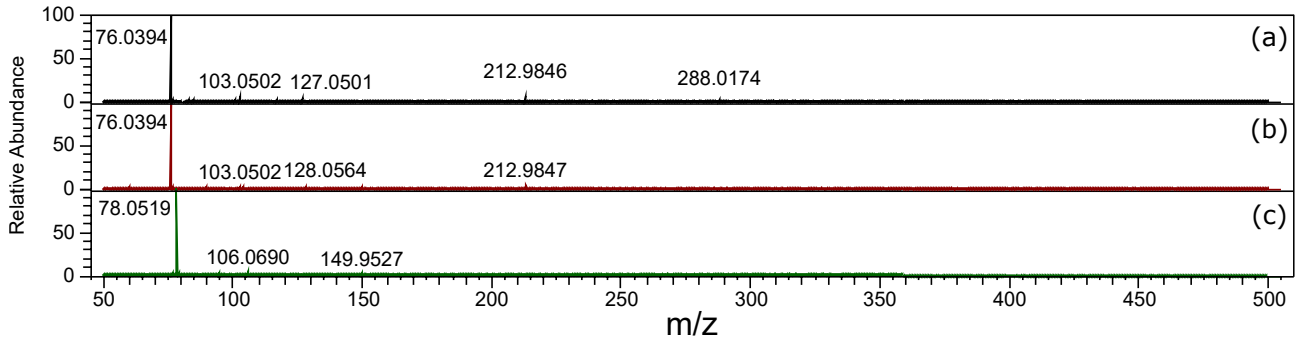


Supplementary Figure 3. Theoretically calculated IR spectrum of glycylglycine, the dipeptide that was identified through its amide bands and ESI-MS detection. Its spectrum contains bands with positions similar to those of the gaussians assigned in the experimental residue spectra further giving evidence of peptide formation.

ESI-MS measurements

When irradiated glycine samples were investigated using electrospray ionisation mass spectrometry (ESI-MS), the major detection was glycine. This can be seen in Supplementary Figure 4 where the main detection with 100 relative abundance is m/z = 76.039 for the glycine sample (a) and the partially deuterated D3-glycine sample (b) and m/z = 78.052 for the fully deuterated D5-glycine sample in (c).

The hypothesis for why glycine is the most abundant molecule is based on several key factors. First, the initial glycine may have been deposited onto a substrate larger than the proton beam itself, which results in some glycine remaining unprocessed by the protons. Because the ESI-MS measurements are conducted on the substrate, they capture both this unaltered glycine and the newly formed molecules. However, the beam size and substrate size are roughly equal which makes this less likely. The second explanation is that, when broken down into its constituent parts by energetic processing, glycine may re-form into glycine once again. Furthermore, it is possible that some glycine in the proton beam survives irradiation and can still be detected. These factors contribute to the abundance of glycine, while also explaining why larger and more complex species that are generated have relatively lower abundances.

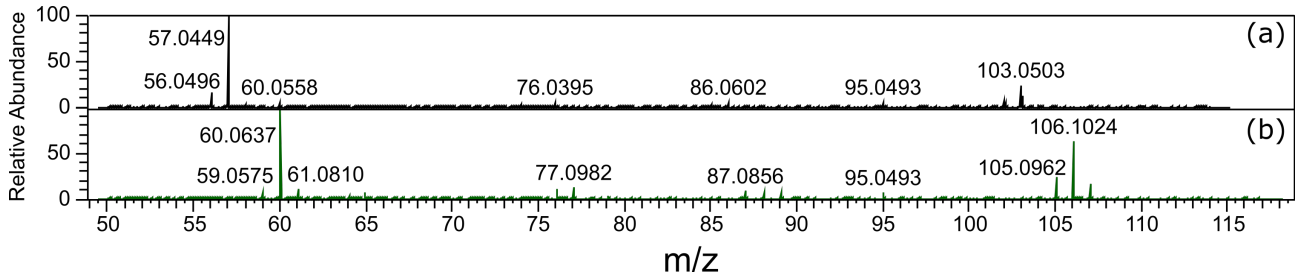


Supplementary Figure 4. Total abundance of the ESI-MS data of the irradiated samples where (a) is the sample that had non-deuterated glycine deposited on it before irradiation, (b) the D3-glycine and (c) the D5-glycine. The major detection for each was glycine at m/z = 76.0394 for (a) and (b) and m/z = 78.0519 for (c).

ESI-MS molecule Fragmentation

The proposed molecular species with an m/z of 103.05, N-formylglycinamide, is not the only possible structure corresponding to this nominal mass. However, its identification is supported by two key factors. First, the observed mass-to-charge ratio matches the theoretical value for N-formylglycinamide with high precision. Second, the ion at m/z = 103.0502 was isolated and subjected to tandem mass spectrometry (MS/MS), which produced fragmentation patterns consistent with the proposed molecular structure. This combination of accurate mass measurement and corroborative fragmentation data provides strong but not conclusive evidence for the formation of N-formylglycinamide under experimental conditions. This is shown in Supplementary Figure 5 where (a) is the irradiated non-deuterated glycine sample and (b) is the irradiated D5-glycine sample and both were carried out with collisional energies of 30. (a) shows that m/z = 103.05 and that the main fragment of this molecule has m/z = 57.04. (b) shows the D5-glycine sample with an m/z = 106.07 and its main fragment with an m/z of 60.06. The deuterated sample has a mass of 3 more than the nondeuterated sample, meaning it has maintained three deuterium atoms that have not been exchanged in solution. These are likely to be on a central carbon atom because it does not undergo exchange reactions with hydrogen in solution, and therefore retains its deuterium. The deuterium bonded to the peptide bonded nitrogen next to an aldehyde may be difficult to exchange with hydrogen, possibly due to steric or electronic effects, therefore it is most likely retained as a deuterium atom.

The fragment with a m/z = 57.04 is proposed to be (CH_2CONH), and therefore the corresponding fragment for the D5-glycine sample would be (CD_2COND) and should be 3 amu heavier with a m/z of 60 as measured experimentally with an m/z = 60.06 as seen in the Supplementary Figure 5 (b). This means that the molecule likely has one peptide bond and one central carbon atom containing these deuterium atoms therefore *N*-Formylglycinamide ($NH_2CH_2CONHCHO$) being suggested as a candidate for the m/z of 103.05 and deuterated *N*-Formylglycinamide ($NH_2CD_2CONDCHO$) for the m/z of 106.07. The different proposed structures of these molecules are shown in Supplementary Table 2.

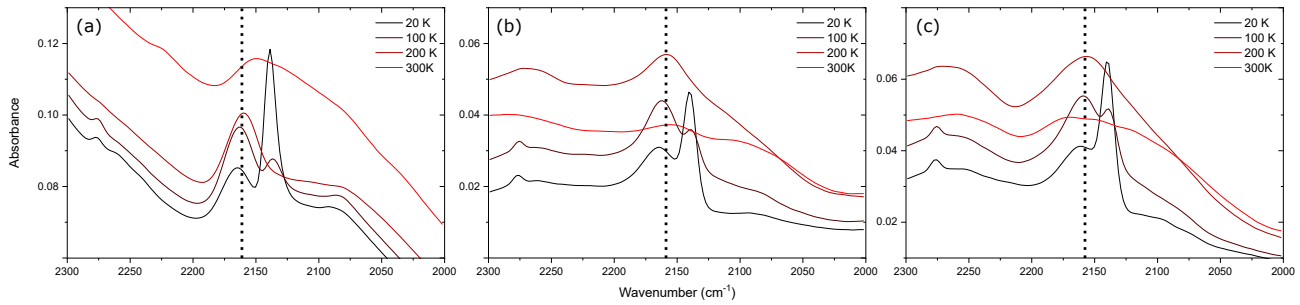


Supplementary Figure 5. (a) shows the ESI-MS of the isolated m/z = 103.05, of the irradiated non-deuterated glycine sample. It shows its main residual mass fragment at 57.04. (b) is the isolated m/z = 106.07 of the irradiated D5-glycine sample which has a main fragment at 60.06.

OCN-type IR band contribution

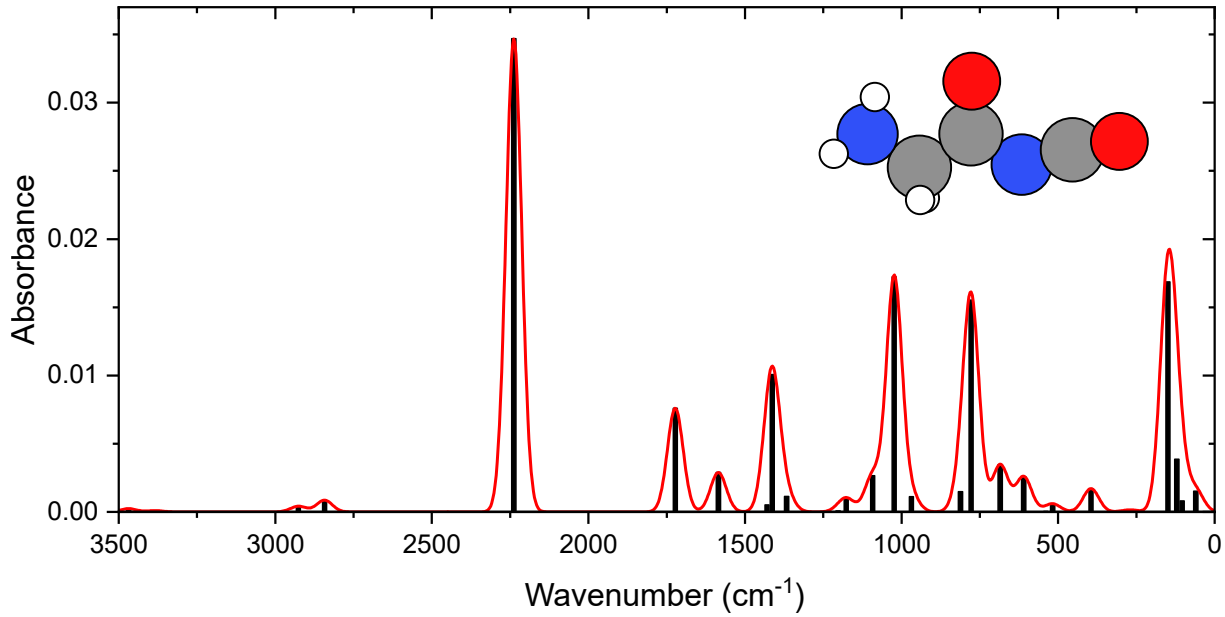
During irradiation of the three glycine samples, a band near 2160 cm^{-1} was observed in the IR spectra. This band has been reported in previous studies and has been attributed to the OCN^- ion [4], as it corresponds to the asymmetric stretching mode of the NCO group [7–9]. However, as shown in Supplementary Figure 6, this band is consistently present across a range of temperatures, with a distinct signal for each sample at 200 K. It also appears to persist—albeit shifted in wavenumbers—even at

300 K. Given that HNCO desorbs between 100 and 120 K [8, 9], the presence of this band at significantly higher temperatures suggests additional contributors beyond OCN^-. We hypothesize that other larger COMs containing the OCN functional group also contribute to this spectral feature. These COMs would remain in the solid phase at higher temperatures, which explains the continued presence of the band.



Supplementary Figure 6. The spectra of glycine (a), D3-gycine (b) and D5-glycine (c) in the range of 2000 to 2300 cm^{-1} at temperatures of 20, 100, 200 and 300 K. All three figures show a band at around 2160 cm^{-1}, indicated with dotted lines, that remains present at all temperatures. This band is commonly associated with OCN^- and has been done so in previous work, however here is also attributed to larger COMs that have formed as a result of the irradiation that contain an OCN group and at high temperatures will not desorb.

One potential complex organic molecule that could contribute to this observed band is N-formylglycinamide, which may form prior to dissolution. This molecule includes an OCN moiety and, upon formation, is hypothesized to lack a hydrogen bond to the carbon atom within the functional group, resulting in the structure shown in Supplementary Figure 7. The IR spectrum of N-formylglycinamide was simulated as described in the Methods section. Supplementary Figure 7 presents the calculated spectrum, where the black lines indicate the computed band centres, and the red Gaussians aid in visualizing the corresponding IR features. A prominent band at 2262 cm^{-1} is assigned to the asymmetric stretching mode of the NCO group, consistent with the experimentally observed feature shown in Supplementary Figure 6. We attribute the offset between the calculated and measured peak positions to a combination of factors. Firstly, the molecules are simulated in the gas phase with the CPCM assumption, which neglects van der Waals interactions with the surface. These weak interactions can contribute a downward shift in frequency. Secondly, a multiplicative correction is applied according to empirical benchmarking performed for peptides, which is subject to further deviation. Due to its relatively large molecular size, N-formylglycinamide would not desorb from the surface at low temperatures and could therefore contribute to the persistent IR absorption band near 2160 cm^{-1}. Upon dissolution, this molecule may also be responsible for the ex situ detection of $m/z = 103$.

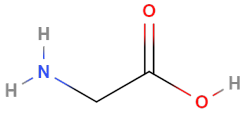
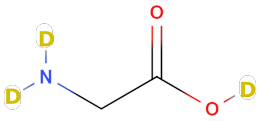
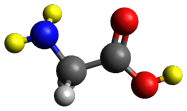
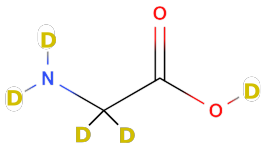
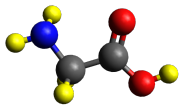
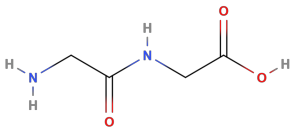
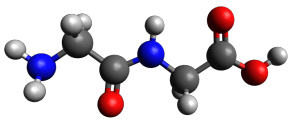
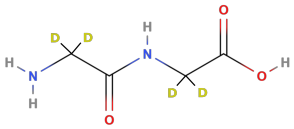
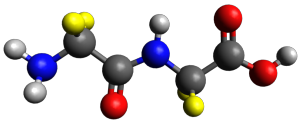
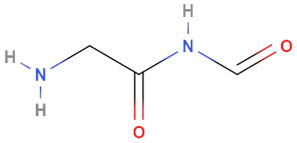
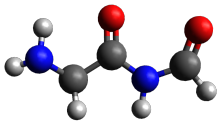
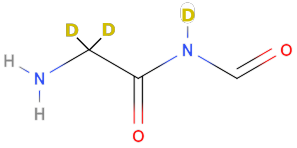
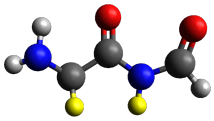


Supplementary Figure 7. Theoretically calculated IR spectrum of *N*-Formylglycinamide, one of the proposed molecules that was formed via energetic processing before being placed in solution for the ESI-MS. This molecule is theorized to be present after irradiation and will not desorb off the surface at high temperatures. Its spectrum does contain a band at 2262 cm^{-1} attributed to the NCO asymmetric stretch.

Deposited and detected molecules

Supplementary Table 2 shows the names, molecular formula, chemical skeleton and ball and stick figures of the COMs deposited and those formed and detected. The three different starting isotopologues of glycine are shown followed by the glycylglycine dipeptide and *N*-Formylglycinamide.

Supplementary Table 2. Names, molecular formula, skeleton structure and ball-and-stick structure of COM species mentioned in the manuscript.

Molecule name	Molecular formula	Skeleton figure	Ball and stick figure
Glycine	NH_2CH_2COOH		
D3-Glycine	ND_2CH_2COOD		
D5-Glycine	ND_2CD_2COOD		
Glycylglycine	$NH_2CH_2CONHCH_2COOH$		
Deuterated Glycylglycine	$NH_2CD_2CONHCD_2COOH$		
N-Formylglycinamide	$NH_2CH_2CONHCHO$		
Deuterated *N*-Formylglycinamide	$NH_2CD_2CONDCHO$		

Notes. Red is an oxygen atom, blue is nitrogen, black is carbon, white is hydrogen and yellow is deuterium. The ball-and-stick figures were made using the Avogadro molecular builder software [10].

References

1. Palumbo, M., Baratta, G., Leto, G. & Strazzulla, G. H bonds in astrophysical ices. *J. Mol. Struct.* **972,** 64–67. ISSN: 0022-2860 (2010).
2. Peláez, R. J. *et al.* Plasma generation and processing of interstellar carbonaceous dust analogs. *Plasma Sources Sci. Technol.* **27,** 035007 (2018).
3. Mennella, V., Baratta, G. A., Esposito, A., Ferini, G. & Pendleton, Y. J. The Effects of Ion Irradiation on the Evolution of the Carrier of the 3.4 Micron Interstellar Absorption Band. *Astrophys. J.* **587,** 727–738 (2003).
4. Maté, B., Tanarro, I., Escribano, R., Moreno Alba, M. & Herrero, V. Stability of extraterrestrial glycine under energetic particle radiation estimated from 2 keV electron bombardment experiments. *Astrophys. J.* **806,** 151 (2015).
5. Krasnokutski, S. A., Chuang, K.-J., Jäger, C., Ueberschaar, N. & Henning, T. A pathway to peptides in space through the condensation of atomic carbon. *Nat. Astron.* **6,** 381–386. arXiv: 2202.12170 [physics.bio-ph] (2022).
6. Kaiser, R. I., Stockton, A. M., Kim, Y. S., Jensen, E. C. & Mathies, R. A. On the formation of dipeptides in interstellar model ices. *Astrophys. J.* **765,** 111 (2013).
7. Theule, P. *et al.* Kinetics of the OCN^- and HOCN formation from the HNCO + H_2O thermal reaction in interstellar ice analogs. *Astron. Astrophys.* **530,** A96 (2011).
8. Farkas, A., Berkó, A. & Solymosi, F. Interaction of HNCO with Au(111) surfaces. *Surf. Sci.* **606,** 1345–1349. ISSN: 0039-6028 (2012).
9. Noble, J. A. *et al.* Hydrogenation at low temperatures does not always lead to saturation: the case of HNCO. *Astron. Astrophys.* **576,** A91 (2015).
10. *Avogadro: an open-source molecular builder and visualization tool. Version 1.2.* https://avogadro.cc/. 2023.