# A rich hydrocarbon chemistry and high C to O ratio in the inner disk around a very low-mass star

Tabone, B.; Bettoni, G.; Dishoeck, E.F. van; Arabhavi, A.M.; Grant, S.; Gasman, D.; ... ; Wright, G.

**Citation**

Tabone, B., Bettoni, G., Dishoeck, E. F. van, Arabhavi, A. M., Grant, S., Gasman, D., … Wright, G. (2023). A rich hydrocarbon chemistry and high C to O ratio in the inner disk around a very low-mass star. *Nature Astronomy*, *7*, 805-814. doi:10.1038/s41550-023-01965-3 nature astronomy

Article

https://doi.org/10.1038/s41550-023-01965-3

# A rich hydrocarbon chemistry and high C to O ratio in the inner disk around a very low-mass star

Received: 20 December 2022

Accepted: 3 April 2023

**B. Tabone [1,2] ✉, G. Bettoni [2,3], E. F. van Dishoeck [2,3], A. M. Arabhavi [4], S. Grant [3], D. Gasman [5], Th. Henning [6], I. Kamp [4], M. Güdel [6,7,8], P. O. Lagage [9], T. Ray [10], B. Vandenbussche [5], A. Abergel [1], O. Absil [11], I. Argyriou [5], D. Barrado [12], A. Boccaletti [13], J. Bouwman [6], A. Caratti o Garatti [10,14], V. Geers [15], A. M. Glauser [7], K. Justannont [16], F. Lahuis [17], M. Mueller [4], C. Nehmé [9], G. Olofsson [18], E. Pantin [9], S. Scheithauer [6], C. Waelkens [5], L. B. F. M. Waters [19,20], J. H. Black [16], V. Christiaens [11], R. Guadarrama [8], M. Morales-Calderón [12], H. Jang [19], J. Kanwar [4,21], N. Pawellek [8,22], G. Perotti [6], A. Perrin [23], D. Rodgers-Lee [10], M. Samland [6], J. Schreiber [6], K. Schwarz [6], L. Colina [24], G. Östlin [18] & G. Wright [15]**

Carbon is an essential element for life but how much can be delivered to young planets is still an open question. The chemical characterization of planet-forming disks is a crucial step in our understanding of the diversity and habitability of exoplanets. Very low-mass stars (less than 0.2 $M_\odot$) are interesting targets because they host a rich population of terrestrial planets. Here we present the James Webb Space Telescope detection of abundant hydrocarbons in the disk of a very low-mass star obtained as part of the Mid-InfraRed Instrument mid-INfrared Disk Survey (MINDS). In addition to very strong and broad emission from $C_2H_2$ and its $^{13}C^{12}CH_2$ isotopologue, $C_4H_2$, benzene and possibly $CH_4$ are identified, but water, polycyclic aromatic hydrocarbons and silicate features are weak or absent. The lack of small silicate grains indicates that we can look deep down into this disk. These detections testify to an active warm hydrocarbon chemistry with a high C/O ratio larger than unity in the inner 0.1 astronomical units (AU) of this disk, perhaps due to destruction of carbonaceous grains. The exceptionally high $C_2H_2/CO_2$ and $C_2H_2/H_2O$ column density ratios indicate that oxygen is locked up in icy pebbles and planetesimals outside the water iceline. This, in turn, will have important consequences for the composition of forming exoplanets.

M dwarfs are the most common stars in the Galaxy and are known to host exoplanets in abundance[1,2]. However, the terrestrial planet-forming zones of the disks around M dwarfs have been largely inaccessible with previous observations due to limited spatial and spectral resolution and the dim nature of these objects. The source 2MASS-J16053215-1933159 (hereafter denoted J160532) is a member of the roughly 3–11-Myr-old Upper Scorpius star forming region at a distance of 152 ± 1 pc (ref. 3), with an age of 2.6 ± 1.6 Myr (ref. 4). Its spectral type of M4.75 points to a very low-mass young star ($M_* = 0.14\ M_\odot$, $L_* = 0.04\ L_\odot$) (refs. 5–7) that is still undergoing accretion at a rate of roughly $10^{-10}$–$10^{-9}\ M_\odot\ \mathrm{yr}^{-1}$. Its

A full list of affiliations appears at the end of the paper. ✉e-mail: benoit.tabone@universite-paris-saclay.fr

broadband infrared spectral energy distribution indicates the presence of circumstellar material in the form of a disk-like structure in which planets could originate. The non-detection of millimetre continuum emission[8] suggests that the current disk mass in millimetre-sized grains is less than 0.75 $M_{\mathrm{Earth}}$. For a standard gas to dust ratio of 100, this would indicate a gas mass that is less than 20% of that of Jupiter.

We observed J160532 with the James Webb Space Telescope (JWST)–Mid-InfraRed Instrument (MIRI)[9] Medium Resolution Spectrometer (MRS) with a spectral resolving power $R$ of roughly 1,500–4,000 covering 5–28 μm as part of the guaranteed time MIRI mid-INfrared Disk Survey (MINDS) (see Methods section for more details). The continuum-subtracted 5–18 μm part of the MIRI spectrum is presented in Fig. 1. Compared with mid-infrared spectra of other disks around low-mass stars obtained with the Spitzer Space Telescope[10,11], its shape is unusual[7]. The broad wavelength spectrum clearly shows two strong, broad bumps centred at 7.7 and 13.7 μm, not seen towards any other disk so far[12]. In contrast, no clear silicate emission features are found around 10 and 18 μm, nor any features due to polycyclic aromatic hydrocarbons (PAHs) at 6.2, 7.7, 8.6 or 11.3 μm (spectrum in Extended Data Fig. 1). These broadband characteristics suggest that the J160532 disk is settled and evolved, with silicate grains in the disk atmosphere that must have grown to at least 5 μm.

The much higher spectral resolution of MIRI–MRS compared with previous Spitzer data reveals numerous narrow hydrogen recombination lines as well as molecular features on top of the continuum (Fig. 1). $C_2H_2$ emission at 13.7 μm is particularly strong, consistent with earlier findings that this molecule is enhanced in disks around brown dwarfs and very low-mass stars[7,12,13]. We focus here on the analysis of the molecular lines and demonstrate that they can be ascribed to a mix of small aliphatic and aromatic hydrocarbon molecules plus $CO_2$, but that any water lines are weak. Column density ratios are found to be very different from those found in disks around the more massive T Tauri stars.

## Results

Molecular species were identified by matching the most prominent features in the continuum-subtracted JWST–MIRI spectrum with synthetic spectra (see Methods section for more details). Most of the bands involve vibration–rotation transitions. The synthetic spectrum of each molecule is calculated from a plane-parallel slab model, where the gas is assumed to have a uniform temperature $T$ and the excitation of the molecules to be in local thermodynamical equilibrium (LTE) at a single excitation temperature $T_{\mathrm{ex}}$ equal to $T$[14,15]. The other fitting parameters are the line of sight column density $N$ within a projected emitting area $\pi R^2$ given by its radius $R$. Note that $R$ does not need to correspond to a disk radius, but could also represent a ring with the same area. The best-fitting parameters are summarized in Table 1.

The shape and position of any Q branch, where lines with zero change in rotational quantum number $J$ pile up, are particularly sensitive to temperature. The full ro-vibrational bands of all considered species but $H_2O$ require a treatment of line overlap in the optically thick case. For most species with only a single feature, there is often a degeneracy between a high $T$, low $N$ optically thin solution and a lower $T$, high $N$ optically thick case. Uncertainties and degeneracies associated to the fits are evaluated using a $\chi^2$ approach following earlier studies[14,15] (Extended Data Fig. 2).

### $C_2H_2$ and $^{13}C^{12}CH_2$

The Q branch of $C_2H_2$ at 13.7 μm associated with the $\nu_5$ bending mode on top of the broad continuum is the most prominent feature in the entire MIRI–MRS spectrum (Fig. 1). At an $R$ of roughly 3,000, MIRI–MRS also reveals a series of P and R branches on top of both the 13.7 and 7.7 μm broad bumps. The fact that these bumps coincide in location with emission from gaseous $C_2H_2$ suggests that the carrier may be due to hot and very abundant $C_2H_2$ itself. No solid material can be identified that coincides with these broad bumps, and the spacing between the features is too broad to be due to silicate absorption in a near edge-on system suggested on the basis of *Spitzer* data[7]. We demonstrate here that both the broad and narrow components are well reproduced by a two-component model consisting of highly optically thick and more optically thin $C_2H_2$ emission.

Figure 1 shows that the overall shape of the 13.7 μm continuum bump can be well fit by a slab of gas at $T = 525$ K with a column density of $N(C_2H_2) = 2.4 \times 10^{20}$ cm$^{-2}$ within an emitting area of $\pi(0.033\ \mathrm{AU})^2$ (that is, $R = 0.033$ AU). Our fit includes the contribution of $^{13}C^{12}CH_2$, assuming a $C_2H_2/^{13}C^{12}CH_2$ ratio of 35 (ref. 16). In the following, this highly optically thick and compact component is called component I. Such an exceptionally high column density of $C_2H_2$ is required to fully saturate the blended molecular lines and produce a pseudo-continuum that masks any prominent features such as the Q branch (Extended Data Fig. 3). Some fraction of the 7.7 μm combination $\nu_4 + \nu_5$ band is also recovered, but with a too-high contrast between the amplitude of the narrow features and the level of the pseudo-continuum. At such high column densities, hot bands of $^{13}C^{12}CH_2$ that are not included in spectroscopic databases such as HITRAN should contribute as well to the 7.7 μm bump and result in a blending of the individual lines. Proper modelling must await more complete $^{13}C^{12}CH_2$ molecular spectroscopy including highly excited bands.

The presence of a prominent Q branch at 13.7 μm indicates a second physical component, called component II, producing less optically thick $C_2H_2$ emission. Our MIRI–MRS data allow to distinguish also the shortward peaks at 13.63 and 13.68 μm due to hot bands that were blended with the main peak in lower resolution Spitzer spectra[7,13] (Fig. 1, right). These features are not tracing the bulk reservoir of $C_2H_2$ but can either unveil the hotter layer at the surface of the thick component or a more radially extended emission. As an illustration, these features are indeed well reproduced by a more extended lower column density of $N(C_2H_2) = 2.5 \times 10^{17}$ cm$^{-2}$ with $R = 0.07$ AU at a temperature of 400 K. This component would then trace a physically distinct region, at the outer boundary of the $C_2H_2$-rich region of the disk. This is the assumption that we make for the analysis of most of the other molecules, which allows us to subtract the contribution of the two $C_2H_2$ bumps in the spectra (Methods and Extended Data Fig. 1) and fit their features without taking into account the masking of the features by optically thick $C_2H_2$ lines from component I.

The Q branch of $^{13}C^{12}CH_2$ is also detected at 13.73 μm above the optically thick component I. If this emission would originate from the optically thinner component II, a $C_2H_2/^{13}C^{12}CH_2$ abundance ratio of about three would be required to match the peak intensity, a value that is an order of magnitude lower than the interstellar $^{12}C/^{13}C$ ratio. More likely, this indicates a more complex layered structure than the slab model can simulate. Sophisticated models including radial and vertical gradients of temperature and a more complete spectroscopy of $^{13}C^{12}CH_2$ are needed to consistently interpret the prominent peaks of $C_2H_2$ and $^{13}C^{12}CH_2$.

### $C_4H_2$

The emission features at 15.92 μm highlighted in Fig. 2 correspond to the Q branch of the fundamental bending mode $\nu_8$ of di-acetylene, $C_4H_2$. As for other molecules, the shape of this feature depends strongly on temperature, becoming broader at higher $T$. An origin within a small emitting area of $R \lesssim 0.07$ AU is excluded as line overlap would mask the prominent features. This supports the scenario that the $C_4H_2$ features originate from the optically thinner component II. This also motivates the choice of our reference emitting area for the thin component to $R = 0.07$ AU. A model with $T = 330$ K is able to reproduce the feature at 15.92 μm as well as that at 15.88 μm. Moreover, the two emission features between 15.7 and 15.8 μm are well reproduced.

### $C_6H_6$

We identified three features around 14.85 μm to be the Q branches of the fundamental and hot bending mode $\nu_4$ of benzene, $C_6H_6$, presented in

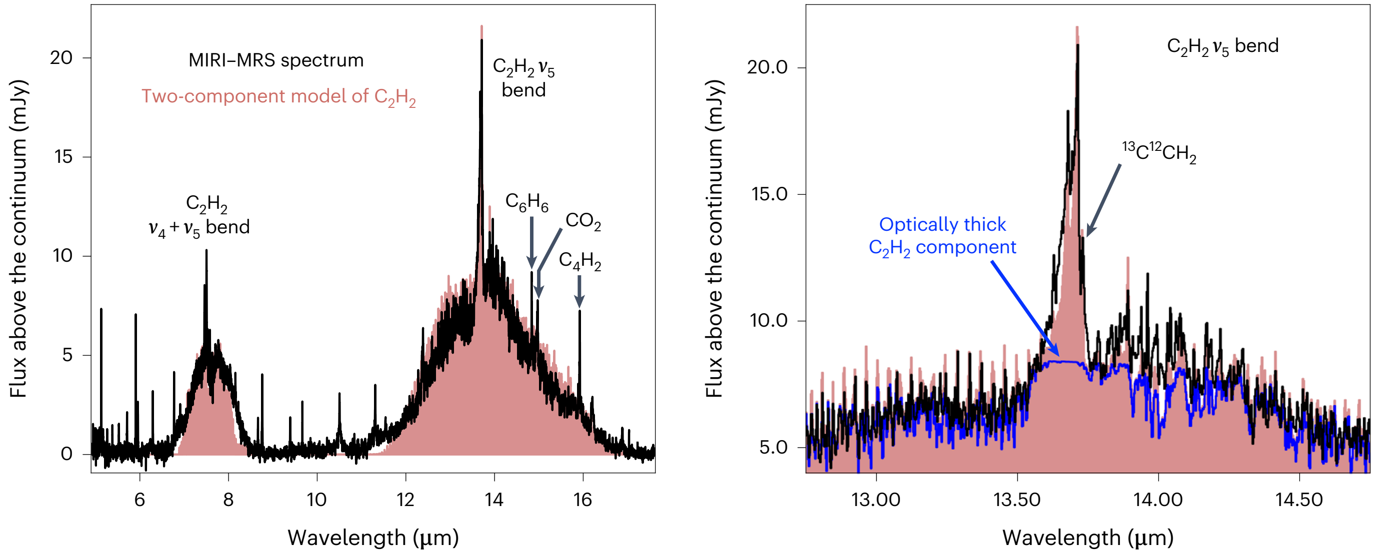


**Fig. 1 | JWST–MIRI–MRS spectrum of J160532 showing prominent $C_2H_2$ emission.** Left: continuum-subtracted MIRI–MRS spectrum of J160532 in the 5–17.5 μm range in black (Extended Data Fig. 1) compared with a simulated model spectrum of $C_2H_2$ in red. The two broad continuum bumps at 7.7 and 13.7 μm are reproduced by a high column density, highly optically thick $C_2H_2$ component I at 525 K with an emitting area of π(0.033 AU)$^2$ that masks the prominent Q branch at 13.7 μm (Extended Data Fig. 3). Right: zoom in on the 13–14.5 μm range, showing that the prominent $\nu_5$ Q branch of $C_2H_2$ is well matched by a second, more extended lower column density and less optically thick component II at 400 K with an emitting area of π(0.07 AU)$^2$. The blue line in this zoom in shows the contribution of the optically thick component I.

**Table 1 | Best-fit slab model results for molecules in the J160532 disk**

| | Component I[a] | | Component II | |
|---|---|---|---|---|
| **Molecule** | ***T*** | ***N*** | ***T*** | ***N*** |
| | **(K)** | **($10^{17}$ cm$^{-2}$)** | **(K)** | **($10^{17}$ cm$^{-2}$)** |
| $C_2H_2$ | 525 | $2,400^{+3,200}_{-1,400}$ | 400 | 2.5 |
| $C_4H_2$ | - | - | 330 | 0.7 |
| $C_6H_6$ | - | - | 400 | 0.7 |
| $CH_4$ | - | - | 400[b] | 1.5 |
| $CO_2$ | 430 | $20^{+55}_{-18}$ | 650 | 0.36 |
| HCN | - | - | 400[b] | ≤1.5 |
| $H_2O$ | 525[b] | ≤30 | 400[b] | ≤8 |

[a] For $H_2O$ and $CO_2$, the reported values correspond to an alternative fit to component II assuming that all the emission originates from the $C_2H_2$ thick component I ($R$ = 0.033 AU).
[b] Fit performed by fixing the temperature to that of the corresponding $C_2H_2$ component.
Uncertainties on the column densities correspond to the 1σ confidence interval obtained by fixing the emitting area and are only valid in the framework of our simple slab modelling. For clarity, uncertainties smaller than 0.5 dex are not reported.

Fig. 2. Their relative intensity is sensitive to temperature and indicates $T \simeq 400$ K. As for $C_4H_2$, compact emission ($R \lesssim 0.05$ AU) is excluded.

### $CH_4$

Extended Data Fig. 4 shows possible indications of $CH_4$ emission. $CH_4$ was previously seen in the GV Tau N disk[17], but only in absorption. We observe emission lines at 7.65–7.67 μm that are aligned with the Q branch of the $\upsilon_4$ mode of $CH_4$. $C_2H_2$ also has many emission lines in this region, but cannot reproduce by itself this broad feature.

### HCN

The ro-vibrational band from the fundamental $\upsilon_2$ bending mode of HCN is severely blended with the strong emission lines of $C_2H_2$. Extended Data Fig. 5 shows the maximum amount of HCN that could be present in the optically thinner $C_2H_2$ component II. If present in the $C_2H_2$ thick component I, HCN emission features would be severely masked and its column density in that region cannot be robustly constrained.

### $CO_2$

Figure 2 includes the fit to the $CO_2$ bending mode at 14.98 μm that is clearly detected. Assuming that $CO_2$ emission originates from the $C_2H_2$ thin component II, the shape of its Q branch indicates a high temperature around 650 K. The $\chi^2$ fit also points towards a smaller emitting area very close to that for the optically thick $C_2H_2$ component I with similar temperature and a column density as high as $2 \times 10^{18}$ cm$^{-2}$. In the latter case, $C_2H_2$ can partially mask $CO_2$ emission but we checked that the fitted column density is then underestimated by less than a factor of 2. The Q branch of $^{13}CO_2$ at 15.4 μm is not detected, in line with the column densities reported in Table 1.

### $H_2O$

Extended Data Fig. 6 presents enlargements of the regions of the MIRI spectrum where water can be observed: at 6 μm through the $\nu_2$ ro-vibrational lines, and longward of 10 μm through highly excited pure rotational transitions[15,18]. In the J160532 spectrum, neither set of lines are clearly seen but there are a few weak features around 17 μm that could potentially be consistent with water emission. In those spectral regions, $C_2H_2$ cannot mask $H_2O$ emission. The values listed in Table 1 could also be viewed as an upper limit on the amount of water hidden in this spectrum. No OH lines are found.

### $H_2$

Several $H_2$ pure rotational lines are clearly seen in the MIRI–MRS spectrum, which will be analysed in detail elsewhere. For a temperature of roughly 550 K, indicated by the $S(1)/S(3)$ line ratio, the total mass of warm $H_2$ is about $3 \times 10^{-5}$ $M_{\rm Jup}$.

### Other species

At the shortest MIRI wavelength range, between 4.9 and 5.1 μm, several CO $\upsilon = 1 - 0$ P branch lines are found, indicative of high temperature gas ($T > 1{,}000$ K) that will be analysed elsewhere. Several other hydrocarbon

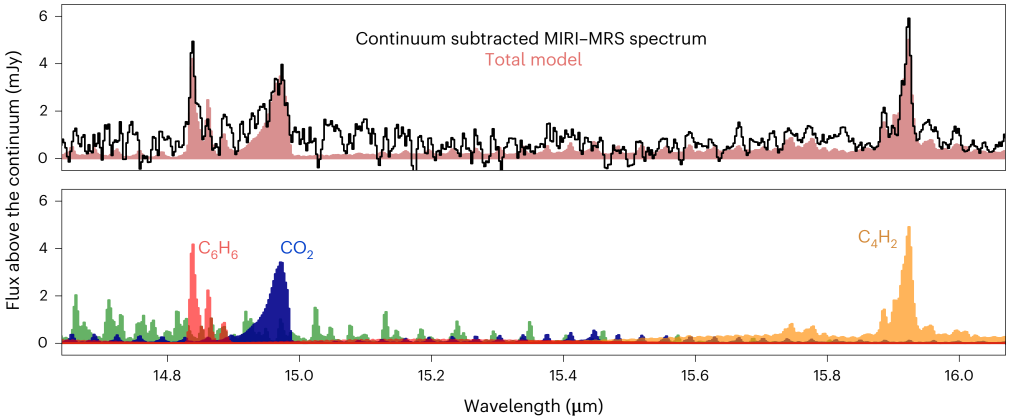


**Fig. 2 | Detection of $C_6H_6$ (benzene), $CO_2$ and $C_4H_2$.** Continuum-subtracted MIRI spectrum, showing a zoom in on a number of key molecular transitions in the 14.7–16 μm region (top panel) and their best-fit slab models assuming an emitting area of $\pi(0.07\ \mathrm{AU})^2$ (bottom panel). The contribution of the two $C_2H_2$ bumps (component I) has been subtracted (Extended Data Fig. 1). However, narrow $C_2H_2$ features are still present shortward of ≃15.5 μm as exemplified by a slab model of $C_2H_2$ (green spectrum). We note that $CO_2$ emission possibly originates from the $C_2H_2$ thick component I with a smaller emitting area of $\pi(0.033\ \mathrm{AU})^2$ (see alternative fit in Table 1).

species ($C_2H_4$, $HC_3N$) were searched for in the J160532 spectrum, but not identified. Also, $NH_3$, whose $\nu_6$ mode at 8.8 μm can be observed with MIRI, was not found in the current spectrum. The [Ne II] line at 12.8 μm is not detected.

## Discussion

The most striking feature of the J160532 MIRI spectrum is the dominance of hydrocarbon emission, most notably $C_2H_2$, but also $C_4H_2$, $C_6H_6$ (benzene) and possibly $CH_4$. In contrast, at best weak $H_2O$ emission is found, and $CO_2$ has a similar column density as most hydrocarbons in component II (Table 1). Hydrocarbon molecules such as $C_4H_2$ and benzene have been found previously in some astrophysical environments, including asymptotic giant branch stars, comets and moons in our own Solar System[19–21], but not yet in the planet-forming zones of disks. These detections therefore highlight that the inner disks around very low-mass objects are indeed very rich in carbon-bearing molecules as suggested on the basis of Spitzer data[7,12,13]. The high column densities point towards being able to probe deep down layers, probably due to a lack of grains in the inner disk.

Figure 3 summarizes the observed column density ratios of key molecules in J160532 with those found in disks around more massive T Tauri stars[15]. For the latter, the sources listed in table 8 of Salyk et al.[15] with detected $C_2H_2$ and $H_2O$ are taken. The $C_2H_2/H_2O$ ratio for the optically thick component I is up to five orders of magnitude higher in the J160532 disk than for T Tauri disks, and even more if $H_2O$ is treated as an upper limit. Similarly, the $C_2H_2/CO_2$ ratio is two orders of magnitude higher. These ratios are much higher than just the flux ratios shown in Extended Data Fig. 7 since our analysis, including the hot bands and a detailed treatment of line overlap, clearly demonstrates that the $C_2H_2$ emission is highly optically thick boosting its column density by orders of magnitude. Note that these column density ratios should not be viewed as local abundance ratios since each molecular band may originate from a different part of the disk, with abundances known to vary radially and vertically[22].

Nevertheless, the J160532 disk is clearly rich in hydrocarbon molecules and the observed chemical differences indicate that hydrocarbon molecules either form more efficiently in disks around very low-mass stars, or that the conditions for their survival are more favourable there. One difference is the ultraviolet (UV) spectrum of the central star, which has far fewer high energy photons that can photodissociate molecules for an M-type star than for an early K- or G-type star. However, J160532 still does have some accretion consistent with its relatively young age of 2.5 Myr; its estimated far UV (FUV) luminosity of $\log(L_{\mathrm{FUV}}) = -3.59\ L_\odot$ (ref. 7) is comparable to that of T Tauri stars. As an M star, J160532 may also have chromospheric activity and flares producing enhanced UV and X-rays, the latter estimated[7] at $\log(L_{\mathrm{X}}) = 28.8\ \mathrm{erg\ s^{-1}}$.

The alternative possibility is enhanced $C_2H_2$ production. Carbonaceous grains and PAHs can be destroyed in the inner disk due to UV radiation, chemical processes or sublimation producing abundant $C_2H_2$ (refs. 23,24). The so-called hydrocarbon 'soot' line as defined by the sublimation front of refractory carbon is estimated to lie at 500 K (ref. 25), the temperature found for the abundant optically thick $C_2H_2$ component I. The exact location of the soot line is however uncertain and depends on the type of carbonaceous material, with some laboratory data putting sublimation of amorphous carbon grains at higher temperatures, up to 1,200 K (ref. 26). One possibility is therefore that we are witnessing carbon grain destruction in the inner disk. Is carbon grain destruction observable in other disks? This remains an open question to be tackled with the upcoming JWST data. If carbon grains are destroyed by UV photolysis, it could be dramatically enhanced in J160532 due to dust growth and settling, which increase the penetration depth of the UV. Alternatively, carbon grain destruction could be happening primarily in the high temperature midplane that is uniquely visible in the J160532 disk but hidden from our view in most disks due to the presence of dust with high infrared optical depth.

Further insight into the hydrocarbon chemistry can be obtained from comparison of our observed column density ratios with those found in thermochemical models[27] of disks around low-mass stars (see Methods section for more details). Both benzene and $C_4H_2$ were predicted to be abundant in inner disk regions[16,28]. For the optically thin $C_2H_2$ component II in which these molecules are not masked by optically thick $C_2H_2$, we find relatively good agreement with the models even though benzene may be underestimated by the models (Extended Data Table 1). Under these conditions, warm dense gas with high $C_2H_2$ abundance, one would expect also efficient PAH formation up to temperatures where erosion starts to take over[29,30]. To what extent the absence of PAH features in the J160532 spectrum also suggests absence of PAHs, or whether there is simply not enough UV radiation to excite and make them visible[31], needs to be quantified. A detailed comparison using a physical–chemical model appropriate for the J160532 disk is postponed to a future study.

An additional source of $C_2H_2$ production due to carbon grain destruction in the J160532 disk would also be consistent with the fact that the observed $C_2H_2/H_2O$ ratio is four orders of magnitude higher

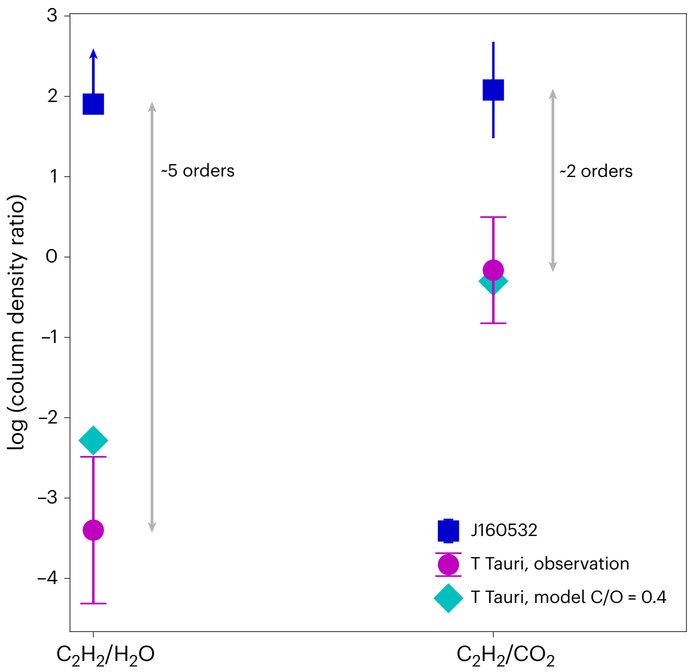


**Fig. 3 | Column density ratios in the inner disk of J160532.** Comparison of the observed column density ratios in component I with an emitting area of $\pi(0.033\ \mathrm{AU})^2$ in the disk of J160532 with those found for T Tauri disks with detected $C_2H_2$ and $H_2O$ (ref. 15). The ratios are also compared with those found in thermochemical disk models that assume solar C/O elemental ratios[16]. Note that column density ratios should not be equated with abundance ratios since abundances of individual molecules vary strongly radially and vertically. The error bars on the measured J160532 ratios are $1\sigma$ confidence interval estimated from our $\chi^2$ fit assuming similar emitting areas for $CO_2$ and $C_2H_2$. The error bars on the ratios measured in the T Tauri sample correspond to the $\pm1\sigma$ of the distribution.

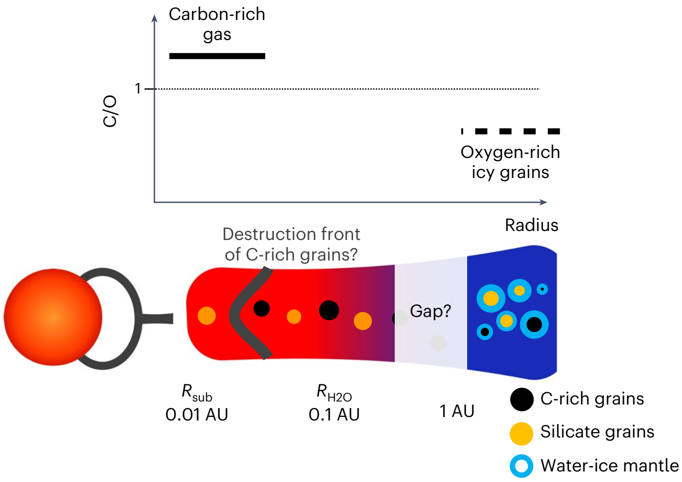


**Fig. 4 | The possible structure of the inner part of the J160532 disk.** This illustration shows the silicate dust sublimation radius around 0.01 AU, the destruction front of carbonaceous grains around 0.033 AU (by sublimation, UV photolysis or chemical processes) and the water snowline around 0.1 AU. In this schematic view, the outer radius of the optically thick $C_2H_2$ component I is put at 0.033 AU derived from the fitted emitting area of $\pi(0.033\ \mathrm{AU})^2$. The location of the silicate and water sublimation fronts is estimated from the luminosity of the star. The top shows the high C/O ratio in the gas in the very inner disk and the low C/O ratio in icy grains at larger radius. The inner disk contains only large (more than 5 μm) silicate grains. The location of any dust trap locking up water ice is unconstrained, except that it must be outside the water snowline at roughly 0.1 AU. Note the log scale for the distance to the star.

than what is found in those models. $CO_2$ is underabundant as well by two orders of magnitude. This conclusion holds irrespective of the stellar mass and adopted UV field in the models[27].

Najita et al.[32,33] suggested that the range in observed HCN/$H_2O$ ratios for T Tauri stars indicates different C/O ratios in the inner disk. High HCN/$H_2O$ ratios would indicate that $H_2O$ is locked up in non-migrating pebbles and planetesimals in the outer disk beyond the water iceline. In the J160532 disk, the HCN/$H_2O$ ratio cannot be robustly constrained but the $C_2H_2/H_2O$ ratio is much higher than for T Tauri disks[34]. In fact, thermochemical disk models of cool M-type stars predict a low $H_2O$ abundance due to the lower temperature of their disks driving much of the oxygen into $O_2$ (ref. 27). Still, for solar C/O ratios, $CO_2$ is predicted to be one of the main oxygen carriers after CO and $O_2$, in stark contrast with our estimates of the $C_2H_2/CO_2$ ratio. Such high abundance ratios as found here can therefore only be reproduced if the C/O ratio of the gas in the inner disk is significantly increased compared with standard values of C/O = 0.4 (ref. 7). In fact, the models of Najita et al.[32], Woitke et al.[22] and Anderson et al.[35] show that values of C/O ≥ 1 are needed to ensure that the bulk of the volatile oxygen is contained in CO leaving little room for $H_2O$ and $CO_2$ production, and permitting the formation of abundant $C_2H_2$.

What could cause the inner disk to be depleted in oxygen compared to carbon? Destruction of carbon-rich grains helps to boost carbon and thus the gas-phase C/O ratio by about a factor of two compared to the volatile carbon in the interstellar medium (about 50% of the total elemental carbon), enough to put the C/O ratio close to unity. However, the very high C/O ratio inferred for the inner disk of J160532 may also point towards a depletion of oxygen. To deplete the inner disk of oxygen the most plausible mechanism would be to lock most of it up in water ice in pebbles and planetesimals in the outer disk, beyond the water snowline, which for such a low-mass star is around 0.1 AU (ref. 36). Gas with high C/O, as often found in the outer disk[37], could still be able to cross the gap. Little is known about the small scale structure of disks around very low-mass stars such as J160532, but our findings indicate that such dust disks are not smooth but must have substructures with dust and ice traps on (sub)AU scales[38,39], as illustrated in Fig. 4. To avoid much oxygen crossing these traps, they must develop early in the disk's evolution, perhaps due to a companion that has formed there. Atacama large millimetre/submillimetre array (ALMA) observations of the outer disk, and analysis of $H_2$ and ro-vibrational CO emission from the inner disk, combined with detailed modelling, are required to constrain the gas-phase C/O and O/H elemental ratios in the inner versus outer disk and further quantify the destruction of carbon grains and the efficiency of oxygen trapping in the outer disk.

What are the implications for any planets forming around J160532? The current amount of solid material contained in millimetre-sized grains in its disk after 2.5 Myr of evolution is less than an Earth mass[8], but the system could have started forming planetesimals and building planets earlier in its lifetime. In fact, there is ample evidence for efficient terrestrial planet formation around very low-mass objects. The occurrence rate per M dwarf is 2.5 ± 0.2 planets with radii of 1–4 $R_{\mathrm{Earth}}$ and periods shorter than 200 days (ref. 1). Surveys also find that short period (fewer than 10 days) planets around stars with $M_* < 0.34\ M_\odot$ are significantly overabundant relative to more massive stars[2]. Our JWST data reveal that the chemistry in disks around such very low-mass stars may have an even higher gaseous C/O ratio in the planet-forming zones than thought before. This, in turn, could significantly affect the composition of planets that may form around them.

To what extent the planets inside 0.1 AU are also carbon rich will depend on whether they are formed mostly from 'dry' planetesimals in this inner disk, or whether their bulk composition and atmospheres result primarily from impacting icy planetesimals from the outer disk, as found in scenarios for making terrestrial planets[40,41]. In fact, it is not clear whether planetesimals formed in the inner disk are carbon rich: our data indicate the presence of some warm silicate refractory dust producing continuum mid-infrared emission, whereas an important fraction of the carbon may be in the gas-phase and even refractory carbon grains may have been destroyed. This carbon-rich gas could be lost from the system over time, with only a small fraction of carbon eventually included in planets[42]. Such a scenario probably holds for our own Earth, which is known to be very carbon poor[25]. These two competing scenarios—making carbon-rich versus carbon-poor terrestrial planets—can be tested by comparing high sensitivity JWST observations of the chemical composition of significant samples of disks around very low-mass young stars with that of the atmospheres of terrestrial-sized planets around mature brown dwarfs and M stars such as the Trappist I system[43].

## Methods

### Observations and data reduction

J160532 was observed with the JWST–MIRI[9,44,45] in Medium Resolution Spectroscopy mode (MRS)[46] on 1 August 2022, from 05:50:01 UT, for a total observation time of 2.22 h. The observation is number 47 of the Cycle 1 Guaranteed Time Observation programme 1282 (PI: Thomas Henning). Following target acquisition, the three grating settings short (A), medium (B) and long (C) were used in each of the four channels observations carried out in parallel, providing full coverage of the MIRI spectral window 4.9–28.1 µm. For each of the three sub-bands, the FASTR1 readout pattern was adopted with a point source four-point dither pattern, an exposure time of 308 s and an integration time of 74.9 s.

For the reduction of the uncalibrated raw data, we used v.1.8.4 of the JWST Science Calibration Pipeline[47] and the CRDS context jwst_1017.pmap. The uncal files were first processed with the default class of the pipeline Detector1. Next, we performed a background correction by subtracting each dither pattern from the other associated pattern (one from four, two from three and vice versa). We then applied the default class Spec2, skipping the residual fringe correction. The data were then processed by the Spec3 class, which combines the calibrated data from the different dither observations into a final level 3 spectral cube. We skipped the outlier_detection and master_background methods of the class, since we already performed the background subtraction after Detector1. We set Spec3 to produce one spectral cube for each sub-band, from which we extracted the spectrum using the pipeline method extract1d. This was done to provide the best input for a residual fringe correction that we applied at the spectrum level.

### Local continuum fit

Extended Data Fig. 1 shows the baseline fit used to produce the spectrum in Figs. 1 and 2. Starting from the original spectrum, first a low-order continuum due to warm dust emission has been removed over the entire 5–20 µm range, producing Fig. 1 and Extended Data Fig. 5. Subsequently, the two broad bumps at 7.7 and 13.7 µm have been subtracted in Fig. 2 and Extended Data Figs. 4 and 6 to further analyse all the molecular features except that of the very optically thick bumps of $C_2H_2$ (component I).

### Slab model fits

The molecular lines are analysed using a slab approach that takes into account optical depth effects. The level populations are assumed to be in LTE and the line profile function to be Gaussian with an intrinsic broadening of $\sigma = 2$ km s$^{-1}$ (full-width at half-maximum of $\Delta V = 4.7$ km s$^{-1}$) to include the effect of turbulence. We note that for optically thin lines, the inferred column densities are independent of the value of $\Delta V$ whereas for optically thick emission, the inferred column densities scale approximatively as $1/\Delta V$. The line emission is assumed to originate from a layer of gas with a temperature $T$ and a line of sight column density of $N$. Because most of the species analysed here produce lines that are close to each other in frequency, we adopt a detailed treatment of line overlap by first computing the wavelength dependent opacity over a fine grid of wavelength ($\lambda/\Delta\lambda \simeq 10^6$)

$$\tau(\lambda) = \sum_i \tau_{0,i} \mathrm{e}^{-(\lambda-\lambda_{0,i})^2/2\sigma_\lambda^2}, \tag{1}$$

where $i$ is the line index, $\lambda_{0,i}$ is the rest wavelength of line $i$, $\sigma_\lambda$ is the intrinsic broadening of the line in µm, and $\tau_{0,i}$ is the optical depth at the centre of line $i$ given by

$$\tau_0 = \sqrt{\frac{\ln 2}{\pi}} \frac{A_{ul} N \lambda_0^3}{4\pi\Delta V}\left(x_l \frac{g_u}{g_l} - x_u\right). \tag{2}$$

In this equation, $x_u$ and $x_l$ denote the population level of the upper and lower states, $g_u$ and $g_l$ their respective statistical weights, and $A_{ul}$ the spontaneous downwards rate of the transition. The flux density $F(\lambda)$ is then computed assuming an emitting area of $\pi R^2$ and a distance to the source $d$ as:

$$F(\lambda) = \pi\left(\frac{R}{d}\right)^2 B_\nu(T)(1-\mathrm{e}^{-\tau(\lambda)}), \tag{3}$$

and convolved at MIRI–MRS spectral resolution. This special treatment, although computationally expensive, is particularly crucial at high column densities for which overlapping lines can form an effectively optically thick continuum across a relatively broad spectral range (Extended Data Fig. 3). Our code has been benchmarked against the publicly available code slabspec[48].

The molecular data, that is, line positions, Einstein $A$ coefficients, statistical weights and partition functions, were taken from the HITRAN 2020 database[49], except for $C_6H_6$ for which the molecular parameters were provided based on the GEISA database[50]. We provide further details about the $C_6H_6$ line list used in the next section.

Protoplanetary disks are obviously not isothermal (vertically or radially), but previous studies have shown that the LTE assumption is a good first step approximation to determine the relative column densities of molecules and physical parameters of the line-emitting regions[14,15]. Non-LTE effects can play a role for the higher excited energy levels[51] if the local density is less than around $10^{15}$ cm$^{-3}$. Differences up to factors of three in inferred column densities have been found in LTE versus non-LTE comparisons for the case of HCN ro-vibrational lines[52]. Since non-LTE effects are expected to be comparable and in the same direction for different molecules, the effect on column density ratios is expected to be smaller than such a factor of three.

Given the overwhelming presence of the $C_2H_2$ band in the MIRI spectrum, we first fit the broad continuum bump (component I) between 12 and 17 µm using a $\chi^2$ approach (Extended Data Fig. 2, left) and including the contribution of $^{13}C^{12}CH_2$ with a $C_2H_2/^{13}C^{12}CH_2$ ratio of 35, half of the interstellar medium value to account for two carbon atoms[16]. To avoid the contribution of the other molecular features, the $\chi^2$ fit is computed using four spectral windows (12.1–12.2, 12.65–12.9, 14.6–14.85 and 15.5–15.7 µm). We then find a more optically thin model that reproduces well the main Q branches of $C_2H_2$ and $^{13}C^{12}CH_2$, assuming that this second component II originates from a different region of the disk. This model, with an emitting area corresponding to $R = 0.07$ AU, is only illustrative since the accuracy of this fit is limited by the lack of spectroscopic data of $^{13}C^{12}CH_2$. Excited states of the $\nu_5$ band of $^{13}C^{12}CH_2$ are indeed missing in the molecular databases whereas the contribution of those states to the Q branch of $C_2H_2$ is expected to be substantial.

The resulting $C_2H_2$ models for both component I and II are then subsequently used in the analysis of the other species to identify spectral windows that are free of contamination by narrow $C_2H_2$ features (for example, Extended Data Fig. 4). However, to subtract the contribution of optically thick $C_2H_2$ (component I), we do not use the best-fit model but subtract a spline fit through the two broad $C_2H_2$ bumps. This strategy avoids artefacts in the subtracted spectra that would be due to the imperfect $C_2H_2$ model.

When subtracting the prominent $C_2H_2$ bumps before the fit of the other features, we implicitly neglect mutual line overlap between $C_2H_2$ and the other species that can be relevant if the emission of the species originates from the optically thick $C_2H_2$ component I. In fact, as discussed in the main text, only the analyses of $CO_2$ and HCN are affected by $C_2H_2$ opacity. This is because $CO_2$ and HCN features could originate from component I and are located close to prominent $C_2H_2$ lines. For HCN, masking by $C_2H_2$ is too substantial to put constraints on the amount of HCN in component I. For $CO_2$, we conducted additional tests and find that the derived column densities of $CO_2$ for component I change by only a factor of less than two when including mutual shielding of the lines.

The best-fit slab model parameters ($N$, $T$, emitting area characterized by $R$) for the species other than $C_2H_2$ are then estimated by a $\chi^2$ approach (Extended Data Fig. 2). Extensive grids of models varying the column density from $10^{15}$ up to $10^{22}$ cm$^{-2}$, in steps of 0.17 in $\log_{10}$ space and temperature from 100 up to 1,500 K in steps of 25 K were computed. Given that $C_2H_2$ emission is highly optically thick, the fit of the 13.7 μm bump allows us to determine the emitting radius of $R = 0.033$ AU. The fit of the $CO_2$ feature points towards a similarly small emitting area. In contrast, for the $C_4H_2$ and benzene prominent features, a compact emission with $R \lesssim 0.07$ AU is excluded. When the $\chi^2$ cannot constrain the emitting size (optically thin lines), the same emitting radius of 0.07 AU is used to evaluate or place upper limits on column densities ($CH_4$, $C_6H_6$, $C_4H_2$, HCN). For $H_2O$, the emitting area is unconstrained and we provide (upper limit) column densities for either component I or II in Table 1.

In the calculation of the $\chi^2$, specific spectral windows are chosen to avoid contamination by other molecular features. The 1, 2 and 3 $\sigma$ confidence intervals are estimated by drawing the contours of $\Delta\chi^2_{\rm red} = \chi^2_{\rm red}(N, T) - \chi^2_{\rm red,min}$ corresponding to values of 2.3, 6.2 and 11.8, respectively, and using a representative noise level of $\sigma = 0.14$ mJy, where $\chi^2_{\rm red}(N, T)$ is the reduced $\chi^2$ obtained by fitting the emitting area for a given value of ($N$, $T$)[53].

### Benzene spectroscopy

Benzene is included in the GEISA database[50], but the existing line list does not provide Einstein A coefficients nor statistical weights and involves only the cold $\nu_4$ band centred at 14.837 μm (673.975 cm$^{-1}$). Therefore, the missing spectroscopic parameters that are necessary for the present study have been generated.

For the cold $\nu_4$ band, we completed the line list, in terms of nuclear spin statistical weights and Einstein coefficients, using the method described in ref. 54, and the available spectroscopic constants[55]. For the partition function, that involves a vibrational and a rotational contribution, we used the empirical equations of ref. 55. Before this, we made extensive calculations to check that these equations are usable for the $50 < T < 500$ K temperature range with an error that is less than 0.5%. The contribution of hot bands at 14.9 μm is missing in the GEISA line list. For this heavy molecule, these hot bands contribute about 45% to the infrared activity at 14.9 μm at room temperature. To account for these contributions, we generated empirical line lists, using the cold $\nu_4$ band as a 'guide list' and the cross-section measurements of benzene performed at high resolution and for different temperatures[56].

### Details of disk models

The observational results are compared with a number of state-of-the-art thermochemical disk models in Extended Data Table 1. In this comparison, we show two observational values of the ratios; assuming either an emitting region corresponding to the highly optically thick component I of $C_2H_2$, or from the less optically thick component II. The thermochemical models assume a gas surface density structure, often taken to be the self-similar solution of a viscously evolving disk. Since we are only interested in the inner few AU of the disk, the precise shape and size of the outer disk are not relevant. The gas distribution in the vertical direction is characterized by a scale height and flaring index. The disk is irradiated by the star whose spectrum is given by its effective temperature $T_{\rm eff}$ and strength by its luminosity. Extra UV due to accretion is modelled by adding a $10^4$ K black body with a strength proportional to the observed accretion rate or $L_{\rm FUV}$. X-rays and cosmic rays are also included, the latter usually at a generic rate of roughly $10^{-17}$ s$^{-1}$.

The models first solve for the dust temperature given the source's luminosity, and then either assume that the gas temperature is equal to the dust temperature or solve explicitly for the gas temperature by iterating over the heating and cooling balance with a small chemical network. Typical gas temperatures in the inner (less than 1 AU) disk are a few thousand K at the top of the atmosphere, dropping to several hundred K deeper in the disk where most of the molecular emission originates[22]. Such temperatures are consistent with our inferred values of 300–600 K for $C_2H_2$ and other molecules. Finally, the 2D abundance distribution of each molecule can be determined by solving the chemistry at each grid point using a more extensive chemical network. Details can be found in refs. 22,32,57.

Most relevant for comparison with our observations are those models that include a large hydrocarbon network. In particular, Woods and Willacy[16] have developed a detailed chemical model appropriate for a disk around a T Tauri star including not just $^{12}$C but also $^{13}$C isotopologues. We take the column densities at 1 AU from their table 3 for comparison in Fig. 3. Walsh et al.[27] have run chemical models for disks around a M dwarf, a T Tauri star and a Herbig star, to investigate how the chemistry differs across the stellar mass range. We focus here on their results for the M-dwarf disk. Walsh et al. show not only total column densities as function of disk radius, but also the column densities above the dust $\tau = 1$ surface at 14 μm, since mid-infrared observations do not probe down to the midplane. We take the latter for our comparison. None of these models vary the input volatile C and O abundances (that is, the amount of carbon and oxygen that can cycle between gas and ice), which are usually taken such that oxygen is more abundant than carbon at C/O = 0.4.

Najita et al.[32] have also presented sophisticated inner disk models focusing on smaller molecules up to $C_2H_2$, HCN and $H_2O$ to investigate trends with disk parameters and C/O. Their models include accretional heating and stellar X-rays, but not FUV radiation in the heating and chemistry. Their model abundances of hydrocarbon molecules are clearly increased when C/O is increased. The same is found in more recent models by Woike et al.[22] and Anderson et al.[35].

## Data availability

The original data analysed in this work are part of the Guaranteed Time Observation-MIRI programme 'MIRI EC Protoplanetary and Debris Disks Survey' (ID 1282) with number 47 and will become public on 1 August 2023 on the MAST database https://archive.stsci.edu/. The continuum-subtracted spectra presented in Fig. 1 (right) and in Fig. 2 are available on Zenodo at https://zenodo.org/record/7850667. The spectroscopic data for all the species but benzene are available on the HITRAN database (https://hitran.org/). For benzene, the data will be shared on request to the corresponding author.

## Code availability

The slab model used in this work is a private code developed by B.T. and collaborators. It is available from the corresponding author upon request. The synthetic spectra presented in this work can be reproduced using the slabspec code, which is publicly available at https://doi.org/10.5281/zenodo.4037306.

## References

1. Dressing, C. D. & Charbonneau, D. The occurrence of potentially habitable planets orbiting M dwarfs estimated from the full Kepler dataset and an empirical measurement of the detection sensitivity. *Astrophys. J.* **807**, 45 (2015).
2. Sabotta, S. et al. The CARMENES search for exoplanets around M dwarfs. Planet occurrence rates from a subsample of 71 stars. *Astron. Astrophys.* **653**, A114 (2021).
3. Gaia Collaboration et al. Gaia Data Release 3: summary of the content and survey properties. Preprint at *arXiv* https://doi.org/10.48550/arXiv.2208.00211 (2022).
4. Miret-Roig, N. et al. The star formation history of Upper Scorpius and Ophiuchus. A 7D picture: positions, kinematics, and dynamical traceback ages. *Astron. Astrophys.* **667**, A163 (2022).
5. Carpenter, J. M., Ricci, L. & Isella, A. An ALMA continuum survey of circumstellar disks in the upper Scorpius OB association. *Astrophys. J.* **787**, 42 (2014).
6. Luhman, K. L., Herrmann, K. A., Mamajek, E. E., Esplin, T. L. & Pecaut, M. J. New young stars and brown dwarfs in the upper Scorpius association. *Astron. J.* **156**, 76 (2018).
7. Pascucci, I., Herczeg, G., Carr, J. S. & Bruderer, S. The atomic and molecular content of disks around very low-mass stars and brown dwarfs. *Astrophys. J.* **779**, 178 (2013).
8. Barenfeld, S. A., Carpenter, J. M., Ricci, L. & Isella, A. ALMA observations of circumstellar disks in the upper Scorpius OB association. *Astrophys. J.* **827**, 142 (2016).
9. Wright, G. S. et al. The Mid-Infrared Instrument for the James Webb Space Telescope, II: design and build. *Publ. Astron. Soc. Pac.* **127**, 595 (2015).
10. Kessler-Silacci, J. et al. c2d Spitzer IRS spectra of disks around T tauri stars. I. Silicate emission and grain growth. *Astrophys. J.* **639**, 275–291 (2006).
11. Furlan, E. et al. A survey and analysis of spitzer infrared spectrograph spectra of T tauri stars in taurus. *Astrophys. J. Suppl. Ser.* **165**, 568–605 (2006).
12. Dahm, S. E. & Carpenter, J. M. Spitzer spectroscopy of circumstellar disks in the 5 Myr old upper Scorpius OB association. *Astron. J.* **137**, 4024–4045 (2009).
13. Pascucci, I. et al. The different evolution of gas and dust in disks around sun-like and cool stars. *Astrophys. J.* **696**, 143–159 (2009).
14. Carr, J. S. & Najita, J. R. Organic molecules and water in the inner disks of T tauri stars. *Astrophys. J.* **733**, 102 (2011).
15. Salyk, C., Pontoppidan, K. M., Blake, G. A., Najita, J. R. & Carr, J. S. A Spitzer survey of mid-infrared molecular emission from protoplanetary disks. II. Correlations and local thermal equilibrium models. *Astrophys. J.* **731**, 130 (2011).
16. Woods, P. M. & Willacy, K. Carbon isotope fractionation in protoplanetary disks. *Astrophys. J.* **693**, 1360–1378 (2009).
17. Gibb, E. L. & Horne, D. Detection of $CH_4$ in the GV Tau N protoplanetary disk. *Astrophys. J. Lett.* **776**, L28 (2013).
18. Carr, J. S. & Najita, J. R. Organic molecules and water in the planet formation region of young circumstellar disks. *Science* **319**, 1504 (2008).
19. Cernicharo, J. et al. Infrared Space Observatory's discovery of $C_4H_2$, $C_6H_2$, and benzene in CRL 618. *Astrophys. J. Lett.* **546**, L123–L126 (2001).
20. Coustenis, A. et al. The composition of Titan's stratosphere from Cassini/CIRS mid-infrared spectra. *Icarus* **189**, 35–62 (2007).
21. Schuhmann, M. et al. Aliphatic and aromatic hydrocarbons in comet 67P/Churyumov-Gerasimenko seen by ROSINA. *Astron. Astrophys.* **630**, A31 (2019).
22. Woitke, P. et al. Modelling mid-infrared molecular emission lines from T Tauri stars. *Astron. Astrophys.* **618**, A57 (2018).
23. Kress, M. E., Tielens, A. G. G. M. & Frenklach, M. The 'soot line': destruction of presolar polycyclic aromatic hydrocarbons in the terrestrial planet-forming region of disks. *Adv. Space Res.* **46**, 44–49 (2010).
24. Anderson, D. E. et al. Destruction of refractory carbon in protoplanetary disks. *Astrophys. J.* **845**, 13 (2017).
25. Li, J., Bergin, E. A., Blake, G. A., Ciesla, F. J. & Hirschmann, M. M. Earth's carbon deficit caused by early loss through irreversible sublimation. *Sci. Adv.* **7**, eabd3632 (2021).
26. Gail, H.-P. & Trieloff, M. Spatial distribution of carbon dust in the early solar nebula and the carbon content of planetesimals. *Astron. Astrophys.* **606**, A16 (2017).
27. Walsh, C., Nomura, H. & van Dishoeck, E. The molecular composition of the planet-forming regions of protoplanetary disks across the luminosity regime. *Astron. Astrophys.* **582**, A88 (2015).
28. Woods, P. M. & Willacy, K. Benzene formation in the inner regions of protostellar disks. *Astrophys. J. Lett.* **655**, L49–L52 (2007).
29. Frenklach, M. & Feigelson, E. D. Formation of polycyclic aromatic hydrocarbons in circumstellar envelopes. *Astrophys. J.* **341**, 372 (1989).
30. Morgan, J. W. A., Feigelson, E. D., Wang, H. & Frenklach, M. A new mechanism for the formation of meteoritic kerogen-like material. *Meteoritics* **26**, 374 (1991).
31. Geers, V. C. et al. C2D Spitzer-IRS spectra of disks around T Tauri stars. II. PAH emission features. *Astron. Astrophys.* **459**, 545–556 (2006).
32. Najita, J. R., Ádámkovics, M. & Glassgold, A. E. Formation of organic molecules and water in warm disk atmospheres. *Astrophys. J.* **743**, 147 (2011).
33. Najita, J. R. et al. The HCN-water ratio in the planet formation region of disks. *Astrophys. J.* **766**, 134 (2013).
34. van Dishoeck, E. F. et al. Water in star-forming regions: physics and chemistry from clouds to disks as probed by Herschel spectroscopy. *Astron. Astrophys.* **648**, A24 (2021).
35. Anderson, D. E. et al. Observing carbon and oxygen carriers in protoplanetary disks at mid-infrared wavelengths. *Astrophys. J.* **909**, 55 (2021).
36. Mulders, G. D., Ciesla, F. J., Min, M. & Pascucci, I. The snow line in viscous disks around low-mass stars: implications for water delivery to terrestrial planets in the habitable zone. *Astrophys. J.* **807**, 9 (2015).
37. Bosman, A. D. et al. Molecules with ALMA at planet-forming scales (MAPS). VII. Substellar O/H and C/H and superstellar C/O in planet-feeding gas. *Astrophys. J. Suppl. Ser.* **257**, 7 (2021).
38. Pinilla, P. et al. Explaining millimeter-sized particles in brown dwarf disks. *Astron. Astrophys.* **554**, A95 (2013).
39. Kurtovic, N. T. et al. Size and structures of disks around very low mass stars in the taurus star-forming region. *Astron. Astrophys.* **645**, A139 (2021).
40. Morbidelli, A., Lunine, J. I., O'Brien, D. P., Raymond, S. N. & Walsh, K. J. Building terrestrial planets. *Ann. Rev. Earth Planetary Sci.* **40**, 251–275 (2012).
41. Ormel, C. W., Liu, B. & Schoonenberg, D. Formation of TRAPPIST-1 and other compact systems. *Astron. Astrophys.* **604**, A1 (2017).
42. Lee, J.-E., Bergin, E. A. & Nomura, H. The solar nebula on fire: a solution to the carbon deficit in the inner Solar System. *Astrophys. J. Lett.* **710**, L21–L25 (2010).
43. Greene, T. P. et al. Thermal emission from the Earth-sized exoplanet TRAPPIST-1 b using JWST. *Nature* https://doi.org/10.1038/s41586-023-05951-7 (2023).
44. Rieke, G. H. et al. The Mid-Infrared Instrument for the James Webb Space Telescope, I: introduction. *Publ. Astron. Soc. Pac.* **127**, 584 (2015).

45. Labiano, A. et al. Wavelength calibration and resolving power of the JWST MIRI Medium Resolution Spectrometer. *Astron. Astrophys.* **656**, A57 (2021).
46. Wells, M. et al. The Mid-Infrared Instrument for the James Webb Space Telescope, VI: the Medium Resolution Spectrometer. *Publ. Astron. Soc. Pac.* **127**, 646 (2015).
47. Bushouse, H. et al. JWST Calibration Pipeline. Zenodo. https://doi.org/10.5281/zenodo.7325378 (2022).
48. Salyk, C. slabspec: Python code for producing LTE slab model molecular spectra. Zenodo. https://doi.org/10.5281/zenodo.4037306 (2020).
49. Gordon, I. et al. The HITRAN2020 molecular spectroscopic database. *J. Quantit. Spectrosc. Radiative Trans.* **277**, 107949 (2022).
50. Delahaye, T. et al. The 2020 edition of the GEISA spectroscopic database. *J. Mol. Spectrosc.* **380**, 111510 (2021).
51. Meijerink, R., Pontoppidan, K. M., Blake, G. A., Poelman, D. R. & Dullemond, C. P. Radiative transfer models of mid-infrared $H_2O$ lines in the planet-forming region of circumstellar disks. *Astrophys. J.* **704**, 1471–1481 (2009).
52. Bruderer, S., Harsono, D. & van Dishoeck, E. F. Ro-vibrational excitation of an organic molecule (HCN) in protoplanetary disks. *Astron. Astrophys.* **575**, A94 (2015).
53. Avni, Y. Energy spectra of X-ray clusters of galaxies. *Astrophys. J.* **210**, 642–646 (1976).
54. Šimečková, M., Jacquemart, D., Rothman, L. S., Gamache, R. R. & Goldman, A. Einstein A-coefficients and statistical weights for molecular absorption transitions in the HITRAN database. *J. Quant. Spectrosc. Radiat. Transf.* **98**, 130–155 (2006).
55. Dang-Nhu, M. & Plíva, J. Intensities in the $\nu_4$, $\nu_{12}$, $\nu_{13}$, and $\nu_{14}$ bands of benzene. *J. Mol. Spectrosc.* **138**, 423–429 (1989).
56. Sung, K., Toon, G. C. & Crawford, T. J. $N_2$- and ($H_2$+He)-broadened cross sections of benzene ($C_6H_6$) in the 7-15 μm region for the Titan and Jovian atmospheres. *Icarus* **271**, 438–452 (2016).
57. Bruderer, S. Survival of molecular gas in cavities of transition disks. I. CO. *Astron. Astrophys.* **559**, A46 (2013).
58. Lebouteiller, V. et al. CASSIS: the Cornell Atlas of Spitzer/infrared spectrograph sources. *Astrophys. J. Suppl. Ser.* **196**, 8 (2011).
59. Banzatti, A. et al. The kinematics and excitation of infrared water vapor emission from planet-forming disks: results from spectrally resolved surveys and guidelines for JWST spectra. *Astron. J.* **165**, 72 (2023).
60. Banzatti, A. et al. Hints for icy pebble migration feeding an oxygen-rich chemistry in the inner planet-forming region of disks. *Astrophys. J.* **903**, 124 (2020).

## Acknowledgements

The MIRI mid-INfrared Disk Survey team thank the entire MIRI European and United States instrument team. Support from StScI is also appreciated. The following National and International Funding Agencies funded and supported the MIRI development: NASA; ESA; Belgian Science Policy Office (BELSPO); Centre Nationale d'Etudes Spatiales (CNES); Danish National Space Centre; Deutsches Zentrum fur Luftund Raumfahrt; Enterprise Ireland; Ministerio De Economiá y Competividad; Netherlands Research School for Astronomy (NOVA); Netherlands Organisation for Scientific Research (NWO); Science and Technology Facilities Council; Swiss Space Office; Swedish National Space Agency and UK Space Agency. B.T. is a Laureate of the Paris Region fellowship programme, which is supported by the Ile-de-France Region and has received funding under Marie Sklodowska-Curie grant agreement no. 945298. B.T. acknowledges support from the Programme National 'Physique et Chimie du Milieu Interstellaire' (PCMI) of CNRS/INSU with INC/INP and cofunded by CNES. G.B. thanks the Deutsche Forschungsgemeinschaft (DFG), grant no. 325594231, FOR 2634/2. E.F.v.D. acknowledges support from the EU ERC grant no. 101019751 MOLDISK and the Danish National Research Foundation through the Center of Excellence 'InterCat' (DNRF150). D.G. thanks the Research Foundation Flanders for cofinancing the present research (grant no. V435622N). T.H. and K.S. acknowledge support from the ERC Advanced grant no. Origins 83 24 28. I.K., A.M.A. and E.F.v.D. acknowledge support from grant no. TOP-1614.001.751 from the Dutch Research Council (NWO). I.K. and J.K. acknowledge funding from H2020-MSCA-ITN- 2019, grant no. 860470 (CHAMELEON). O.A. and V.C. acknowledge funding from the Belgian F.R.S.-FNRS. I.A. and D.G. thank the European Space Agency (ESA) and the Belgian Federal Science Policy Office (BELSPO) for their support in the framework of the PRODEX Programme. D.B. has been funded by Spanish grant no. MCIN/AEI/10.13039/501100011033 grant nos. PID2019- 107061GB-C61 and no. MDM-2017-0737. A.C.G. has been supported by grant nos. PRIN-INAF MAIN-STREAM 2017 and PRIN-INAF 2019 (STRADE). T.P.R. acknowledges support from ERC grant no. 743029 EASY. D.R.L. acknowledges support from Science Foundation Ireland, grant no. 21/PATH-S/9339. L.C. acknowledges support by grant no. PIB2021-127718NB-I00, from the Spanish Ministry of Science and Innovation/State Agency of Research grant no. MCIN/AEI/10.13039/501100011033.

## Author contributions

B.T. and G.B. did the analysis using molecular data files created by A.M.A. and A.P. and a model developed by B.T. and J.H.B. G.B., S.G. and D.G. performed the data reduction, supported by I.A., J.S., M.S., G.P., V.C. and J.B. E.F.v.D., B.T. and G.B. wrote the manuscript. T.H. and I.K. planned and co-led the MIRI guaranteed time project on disks. All authors participated in either the development and testing of the MIRI instrument and its data reduction, in the discussion of the results and/or commented on the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

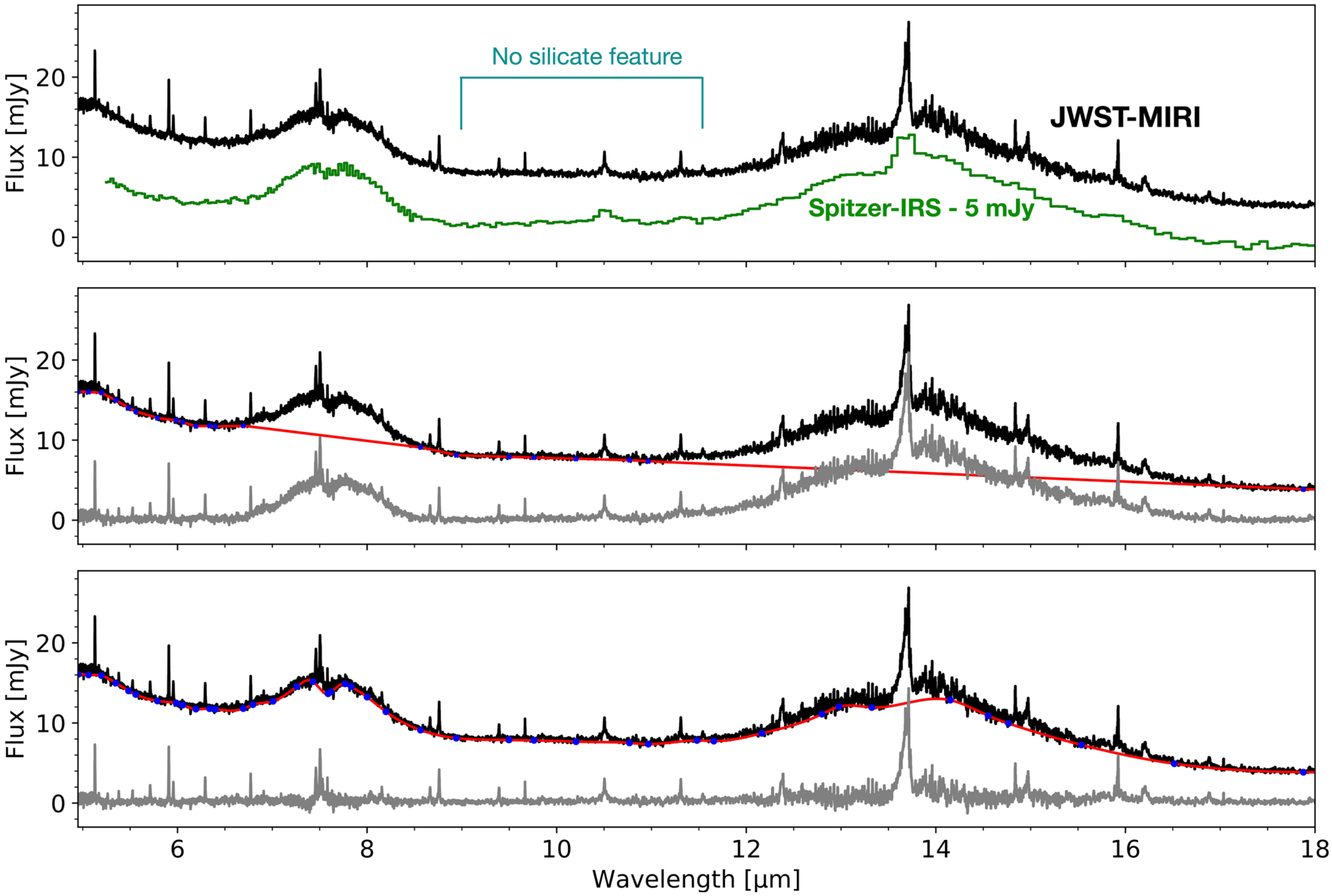
**Extended data** is available for this paper at https://doi.org/10.1038/s41550-023-01965-3.

**Correspondence and requests for materials** should be addressed to B. Tabone.

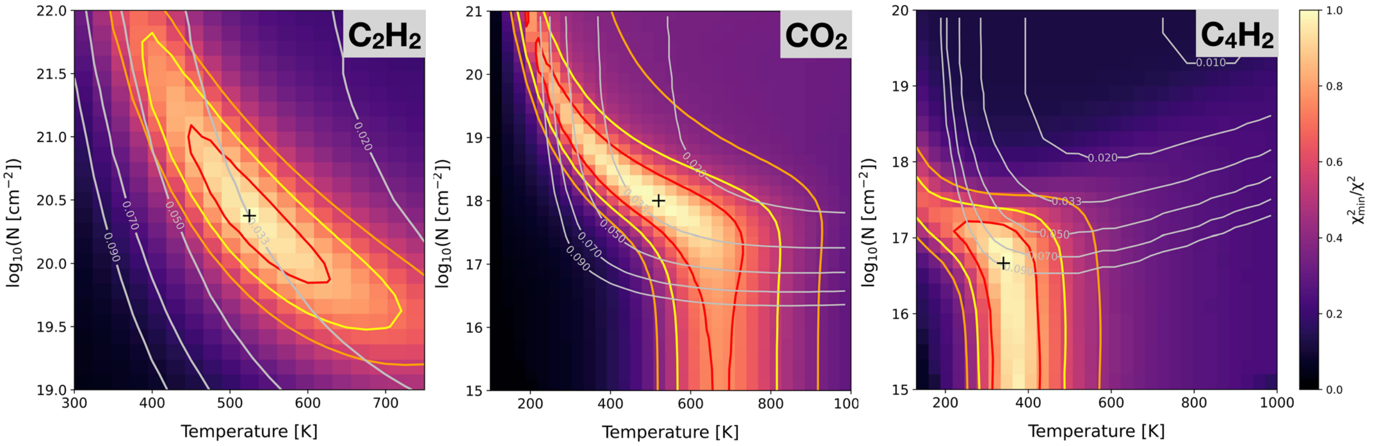
**Peer review information** *Nature Astronomy* thanks the anonymous reviewers for their contribution to the peer review of this work.

**Reprints and permissions information** is available at www.nature.com/reprints.

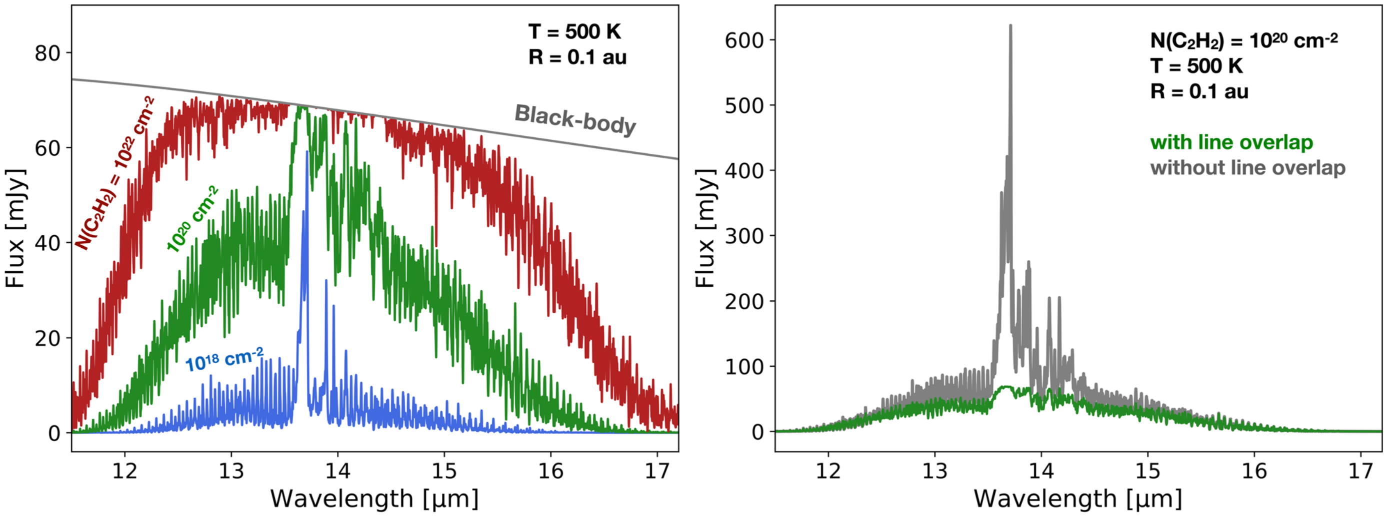
[1]Université Paris-Saclay, CNRS, Institut d'Astrophysique Spatiale, Orsay, France. [2]Leiden Observatory, Leiden University, Leiden, the Netherlands. [3]Max-Planck Institute for Extraterrestrial Physics (MPE), Garching, Germany. [4]Kapteyn Astronomical Institute, University of Groningen, Groningen, the Netherlands. [5]Institute of Astronomy, KU Leuven, Leuven, Belgium. [6]Max-Planck Institute for Astronomy (MPIA), Heidelberg, Germany. [7]Institute for Particle Physics and Astrophysics, ETH Zürich, Zürich, Switzerland. [8]Department of Astrophysics, University of Vienna, Vienna, Austria. [9]Université Paris-Saclay, Université Paris Cité, CEA, CNRS, AIM, Gif-sur-Yvette, France. [10]Dublin Institute for Advanced Studies, Dublin, Ireland. [11]STAR Institute, University of Liège, Liège, Belgium. [12]Centre for Astrobiology (CAB), CSIC-INTA, Madrid, Spain. [13]LESIA, Observatoire de Paris, University PSL, CNRS, Sorbonne University, University of Paris, Meudon, France. [14]INAF—Astronomical Observatory of Capodimonte, Napoli, Italy. [15]UK Astronomy Technology Centre, Royal Observatory Edinburgh, Edinburgh, UK. [16]Department of Space, Earth and Environment, Onsala Space Observatory, Chalmers University of Technology, Onsala, Sweden. [17]SRON Netherlands Institute for Space Research, Groningen, the Netherlands. [18]Department of Astronomy, Stockholm University, AlbaNova University Center, Stockholm, Sweden. [19]Department of Astrophysics/IMAPP, Radboud University, Nijmegen, the Netherlands. [20]SRON Netherlands Institute for Space Research, Leiden, the Netherlands. [21]Space Research Institute, Austrian Academy of Sciences, Graz, Austria. [22]Konkoly Observatory, Research Centre for Astronomy and Earth Sciences, Budapest, Hungary. [23]Laboratoire de Météorologie Dynamique/IPSL, CNRS, Ecole Polytechnique, Institut Polytechnique de Paris, Sorbonne Université, PSL Research University, Palaiseau, France. [24]Centre for Astrobiology (CAB, CSIC-INTA), Carretera de Ajalvir, Madrid, Spain. ✉e-mail: benoit.tabone@universite-paris-saclay.fr



**Extended Data Fig. 1 | Comparison between the *Spitzer*-IRS spectrum and the MIRI-MRS spectrum and baseline fits of the MIRI-MRS spectrum.** The *Spitzer*-IRS low-resolution spectrum[58] has been shifted by 5 mJy to ease the comparison with the MIRI-MRS spectrum (top panel). Baseline fits used in the continuum subtracted spectrum presented in Fig. 1 and 2 in the Results section are shown in the middle and bottom panels, respectively. The blue dots represent the location where the continuum is evaluated. The red curve is the interpolated continuum used to produce continuum-subtracted spectra (in grey). The presence of warm dust is evidenced by the infrared continuum emission on either side of the two $C_2H_2$ bumps but no silicate feature is detected. HI and $H_2$ lines are present in the spectrum and will be analysed in a next paper.

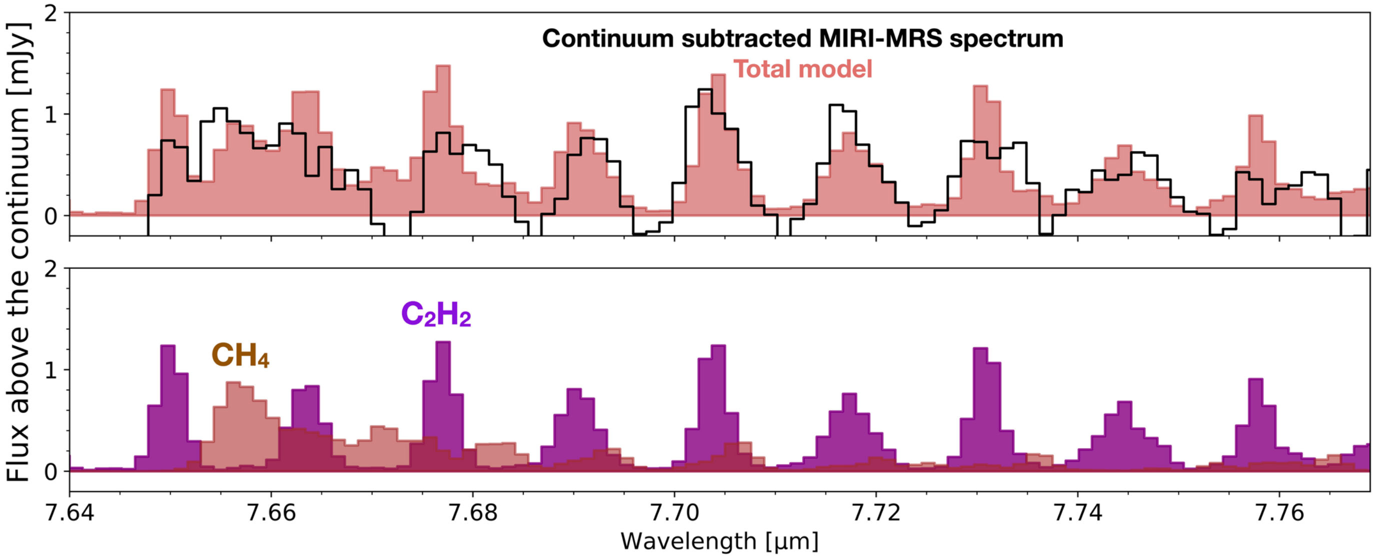


**Extended Data Fig. 2 | Constraints obtained from $\chi^2$ fits.** The $\chi^2$ maps for the fit of the 13.7 $\mu$m broad bump associated with $C_2H_2$ (left), and the $CO_2$ (middle) and $C_4H_2$ (right) features are shown. The 1$\sigma$, 2$\sigma$, and 3$\sigma$ confidence intervals are pictured in red, yellow, and orange, respectively. The best-fitting emitting radius $R$ for each value of $N$ and $T$ is indicated as grey lines. In general, we find a degeneracy between a high $T$ and low $N$ solutions, and a low $T$ and high $N$ solutions. For $CO_2$ the best fit corresponding to an emitting area of 0.033 au is chosen to alleviate the degeneracy and compare with the optically thick component of $C_2H_2$ (component I). We note that for $R = 0.07$ au, corresponding to component II, the $CO_2$ feature can be fitted by either a hot and thin model or a cold and thick model. However, the thick solution over-predicts $^{13}CO_2$ emission which is not detected. We therefore report in Table 1 the column density of the optically thin solution for component II.

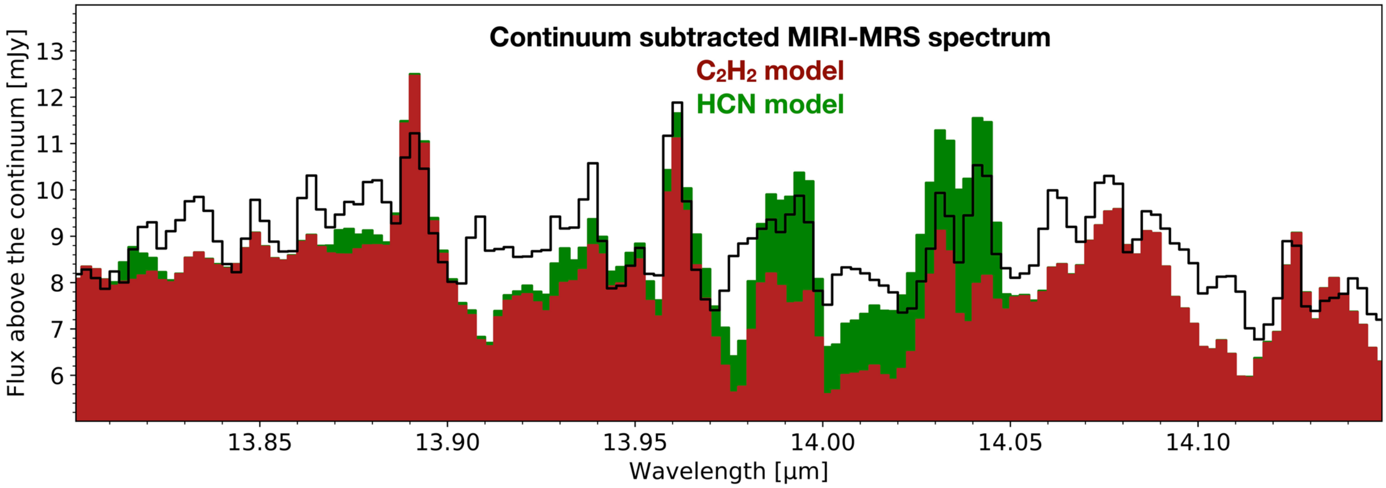


**Extended Data Fig. 3 | Effect of line overlap on the main $C_2H_2$ feature at 13.7 μm.** *Left:* $C_2H_2$ emission as function of column density for $T = 500$ K and $R = 0.1$au. Note that the $Q$-branch becomes highly optically thick above $N(C_2H_2) = 10^{20}$ cm$^{-2}$ and flattens. The contrast between the amplitude of the narrow features on either side of the $Q-$branch and the continuum level decreases by increasing $N(C_2H_2)$. A column density of at least $N(C_2H_2) \simeq 10^{20}$ cm$^{-2}$ is required to fit the observations. *Right:* Importance of line overlap in slab models. For highly optically thick lines that are close to each other such as in the $Q-$branch of $C_2H_2$, slab models neglecting line overlap overestimate the fluxes. For $C_2H_2$, this effect dominates for $N \gtrsim 10^{19}$ cm$^{-2}$.

https://doi.org/10.1038/s41550-023-01965-3

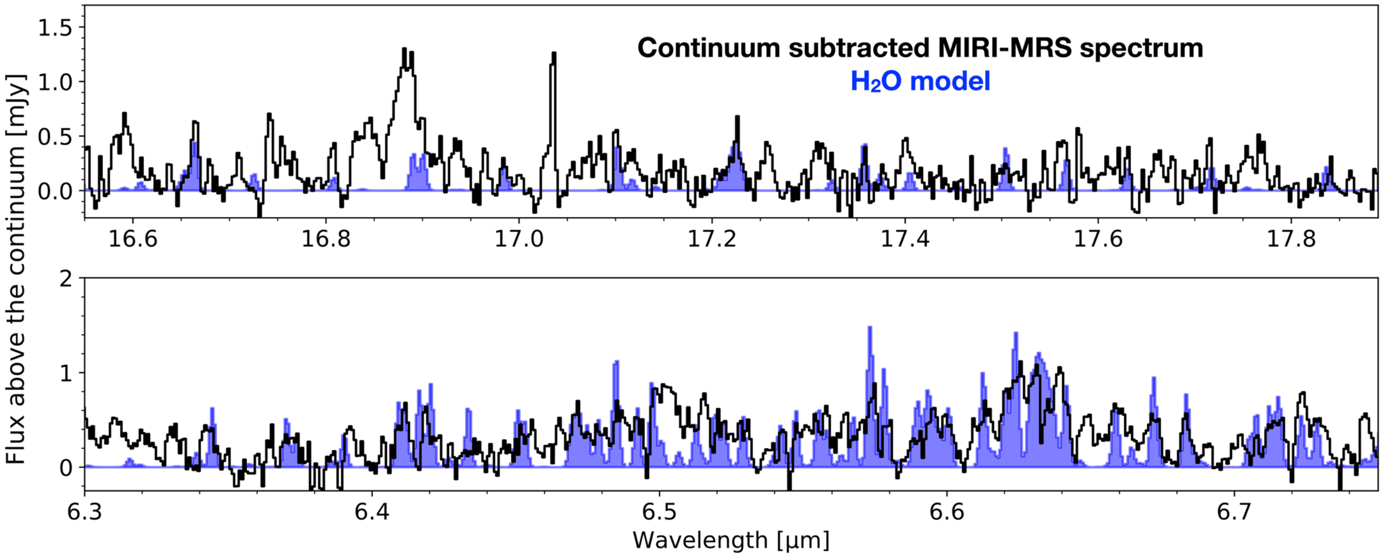


**Extended Data Fig. 4 | Possible indication for $CH_4$ emission in the 7.64-7.77 μm range.** $CH_4$ emission could be present at 7.655 μm in addition to the many $C_2H_2$ lines in this region. The column density of $CH_4$ is estimated assuming that the emission originates from component II (see main text, Table 1). The $C_2H_2$ model in purple corresponds to the component II for which the best-fit column density has been increased by a factor of 4 to better match the series of $C_2H_2$ lines in that specific spectral region.

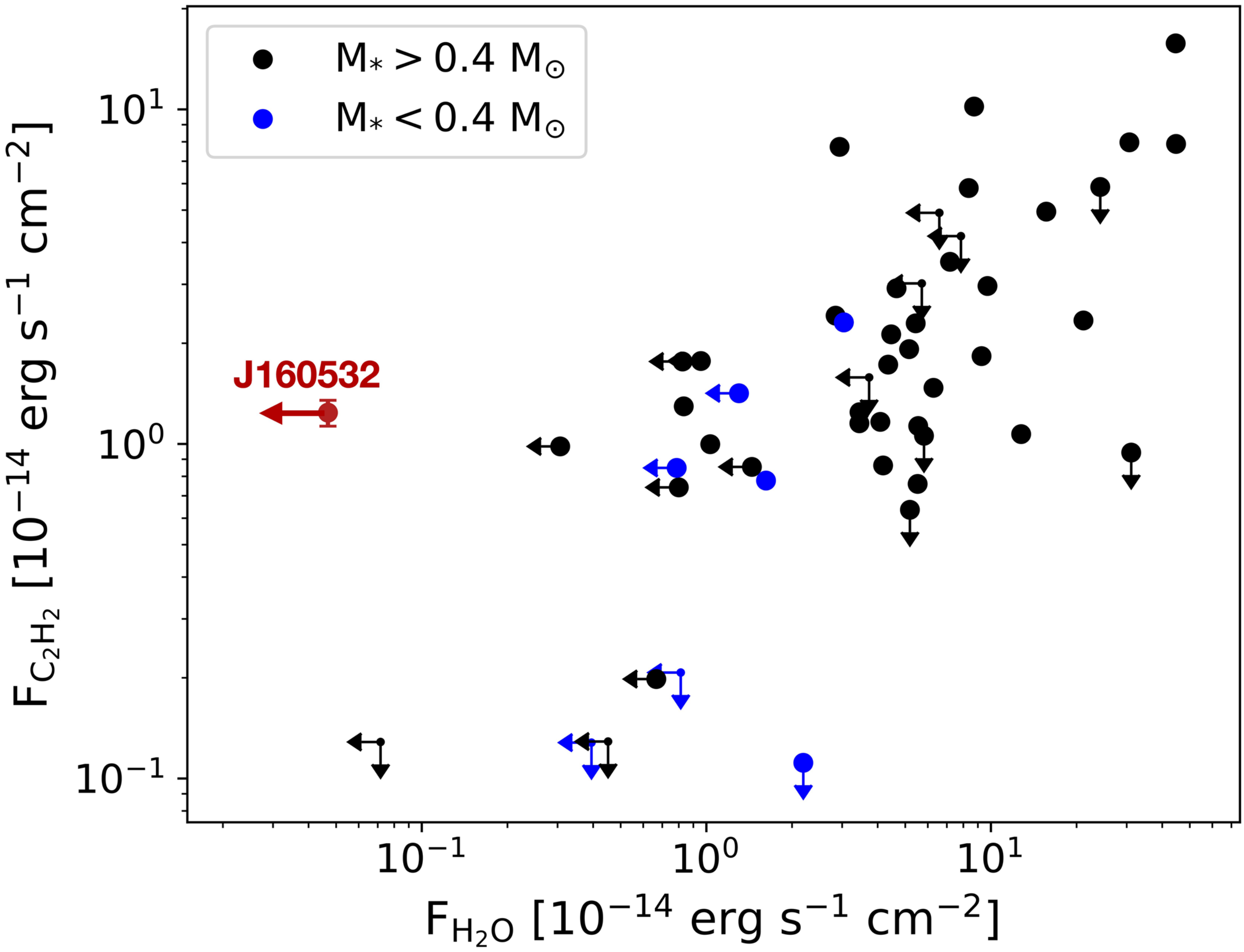


**Extended Data Fig. 5 | Constraints of the amount of HCN in the 14 *µ*m region.** The $C_2H_2$ model, including both component I and II is shown in red on top of the MIRI spectrum where the contribution of the $C_2H_2$ thick component is not subtracted. This figure shows that a maximum column density of HCN of $N = 1.5 \times 10^{17}$ cm$^{-2}$ can be hidden in the $C_2H_2$ line forest in this region assuming an origin in the optically thin component II ($R = 0.07$ au and $T = 400$K). HCN emission from the $C_2H_2$ thick component I would be highly masked by $C_2H_2$ and therefore its column density remains unconstrained.

https://doi.org/10.1038/s41550-023-01965-3



**Extended Data Fig. 6 | Possible detection of weak $H_2O$ lines in the 17.2 $\mu$m and 6.5 $\mu$m regions.** The pure rotational lines at 16.5–18 $\mu$m can hide as much as $N(H_2O)=3\times10^{18}$ cm$^{-2}$ assuming a fixed temperature of 525 K and a characteristic emitting radius of $R$ = 0.033au, corresponding to the optically thick $C_2H_2$ component I. These lines are not affected by masking of $C_2H_2$ since only very weak lines of $C_2H_2$ are present in these spectral ranges. Some lines in the 6.3-6.8 $\mu$m range are somewhat overestimated by our LTE model but non-LTE effects will tend to quench these lines compared to the pure rotational lines longward of ~ 12$\mu$m[59].



**Extended Data Fig. 7 | J160532 line fluxes compared to other disks.** This figure presents a comparison of $C_2H_2$ versus $H_2O$ line flux scaled to 140 pc for a number of T Tauri disks observed with *Spitzer* and complied by ref. 60 and J160532 observed with JWST MIRI. The line fluxes for J160532 are consistently calculated by integrating the flux of the three water features at 17.12 $\mu$m, 17.22 $\mu$m, and 17.36 $\mu$m, and the $C_2H_2$ feature over a window between 13.65-13.72 $\mu$m as explained in ref. 60. Leftward (resp. downward) arrows represent upper limits on $H_2O$ (resp. $C_2H_2$) line flux.

**Extended Data Table 1 | Observed column density ratios of various species compared with disk models with solar C/O elemental ratio in the gas-phase**

| Ratio | Observations[a] | | T Tauri model | BD model |
|---|---|---|---|---|
| | Component I | Component II | WW 09[b] | WNvD15[c] |
| $C_4H_2/C_2H_2$ | - | 0.3 | 0.18 | - |
| $CH_4/C_2H_2$ | - | 0.6 | 0.18 | - |
| $C_6H_6/C_2H_2$ | - | 0.3 | $\geq 0.02$[d] | - |
| $CO_2/C_2H_2$ | $8\ 10^{-3}$ | 0.14 | 2.0 | - |
| $H_2O/C_2H_2$ | $\leq 1\ 10^{-2}$ | $\leq 3$ | 192 | 100 |

a. Assuming that the emission is either confined to the $C_2H_2$ highly optically thick component I or the optically thinner component II. The prominent features of the small hydrocarbons are associated with the extended component II only whereas the $CO_2$ emission originates more likely from the thick component I (see main text). b. Woods & Willacy (2009). c. Walsh, Nomura & van Dishoeck (2015). d. Woods & Willacy (2007).