The Astrophysical Journal Letters, 929:L11 (9pp), 2022 April 10 https://doi.org/10.3847/2041-8213/ac6186

Precursors of the RNA World in Space: Detection of (*Z*)-1,2-ethenediol in the Interstellar Medium, a Key Intermediate in Sugar Formation

Víctor M. Rivilla[1,2], Laura Colzi[1,2], Izaskun Jiménez-Serra[1], Jesús Martín-Pintado[1], Andrés Megías[1], Mattia Melosso[3,4], Luca Bizzocchi[3], Álvaro López-Gallifa[1], Antonio Martínez-Henares[1], Sarah Massalkhi[1], Belén Tercero[5], Pablo de Vicente[6], Jean-Claude Guillemin[7], Juan García de la Concepción[1], Fernando Rico-Villas[1], Shaoshan Zeng[8], Sergio Martín[9,10], Miguel A. Requena-Torres[11,12], Francesca Tonolo[3,13], Silvia Alessandrini[3,13], Luca Dore[3], Vincenzo Barone[13], and Cristina Puzzarini[3]

[1] Centro de Astrobiología (CSIC-INTA), Ctra. de Ajalvir Km. 4, Torrejón de Ardoz, E-28850 Madrid, Spain
[2] INAF-Osservatorio Astrofisico di Arcetri, Largo Enrico Fermi 5, 50125, Florence, Italy
[3] Department of Chemistry "Giacomo Ciamician", University of Bologna, Via F. Selmi 2, Bologna, I-40126, Italy
[4] Scuola Superiore Meridionale, Università di Napoli Federico II, Largo San Marcellino 10, Naples, I-80138, Italy
[5] Observatorio Astronómico Nacional (OAN-IGN), Calle Alfonso XII, 3, E-28014 Madrid, Spain
[6] Observatorio de Yebes (OY-IGN), Cerro de la Palera SN, Yebes, Guadalajara, Spain
[7] Univ Rennes, École Nationale Supérieure de Chimie de Rennes, CNRS, ISCR—UMR6226, F-35000 Rennes, France
[8] Star and Planet Formation Laboratory, Cluster for Pioneering Research, RIKEN, 2-1 Hirosawa, Wako, Saitama, 351-0198, Japan
[9] European Southern Observatory, Alonso de Córdova 3107, Vitacura 763 0355, Santiago, Chile
[10] Joint ALMA Observatory, Alonso de Córdova 3107, Vitacura 763 0355, Santiago, Chile
[11] University of Maryland, College Park, ND 20742-2421 USA
[12] Department of Physics, Astronomy and Geosciences, Towson University, Towson, MD 21252, USA
[13] Scuola Normale Superiore, Piazza del Cavalieri 7, I-56126, Pisa, Italy
Received 2022 February 10; revised 2022 March 14; accepted 2022 March 27; published 2022 April 12

Abstract

We present the first detection of (*Z*)-1,2-ethenediol, $(CHOH)_2$, the enol form of glycolaldehyde, in the interstellar medium toward the G+0.693−0.027 molecular cloud located in the Galactic Center. We have derived a column density of $(1.8 \pm 0.1) \times 10^{13}$ cm^{-2}, which translates into a molecular abundance with respect to molecular hydrogen of 1.3×10^{-10}. The abundance ratio between glycolaldehyde and (*Z*)-1,2-ethenediol is ∼5.2. We discuss several viable formation routes through chemical reactions from precursors such as HCO, H_2CO, CHOH, or CH_2CHOH. We also propose that this species might be an important precursor in the formation of glyceraldehyde ($HOCH_2CHOHCHO$) in the interstellar medium through combination with the hydroxymethylene (CHOH) radical.

Unified Astronomy Thesaurus concepts: Astrochemistry (75); Pre-biotic astrochemistry (2079); Molecular clouds (1072); Galactic centre (565); Interstellar medium (847); Interstellar molecules (849)

1. Introduction

One of the most prominent hypotheses to explain the origin of life is the so-called RNA world (Gilbert 1986), in which RNA molecules could have performed at the same time the genetic and metabolic functions that nowadays are carried out by DNA and proteins, respectively. In the last decade, several prebiotic chemistry experiments in the laboratory have shown that the building blocks of RNA, ribonucleotides, can be formed from simple organic molecules (Powner et al. 2009; Patel et al. 2015; Becker et al. 2016, 2019). Were these simple RNA-world precursors available on early Earth? While some authors have discussed the possibility of an endogenous formation on the planet (e.g., Ruiz-Mirazo et al. 2014; Benner et al. 2019), a possible (and complementary) origin is extraterrestrial delivery (Oró 1961; Chyba & Sagan 1992; Cooper et al. 2001; Pasek 2008; Pearce et al. 2017; Rubin et al. 2019; Rivilla et al. 2020a) during the heavy cometary and meteoritic bombardment that occurred around 3.9 Gyr ago (Marchi et al. 2014). In this way, a broad repertoire of molecules could have landed on Earth's surface from outer space, which implies that they may have already been formed in the interstellar medium (ISM) in the natal molecular cloud prior to the formation of the solar system.

Among the key molecular precursors of the RNA world, sugars have a prominent role. Powner et al. (2009) proposed a viable formation route of RNA nucleotides in prebiotic conditions starting from simpler sugars, like glycolaldehyde ($HOCH_2CHO$) and glyceraldehyde ($HOCH_2CHOHCHO$). New laboratory experiments by Becker et al. (2019) have also shown that glycolaldehyde is a key intermediate in the formation of RNA nucleotides. From all this, a prerequisite for the RNA world is the availability of simple sugars. Glycolaldehyde has been detected in the ISM in different sources (Hollis et al. 2000; Requena-Torres et al. 2008; Beltrán et al. 2009; Jorgensen et al. 2012), and several works have addressed its formation routes through interstellar chemistry (e.g., Bennett & Kaiser 2007; Woods et al. 2012, 2013; Butscher et al. 2015; Fedoseev et al. 2015; Coutens et al. 2018; Skouteris et al. 2018; Rivilla et al. 2019a). While glyceraldehyde has been discovered in meteoritic samples (Cooper et al. 2001; Pizzarello et al. 2012; Furukawa et al. 2019) and in laboratory experiments of interstellar ice analogues (de Marcellus et al. 2015; Meinert et al. 2016; Fedoseev et al. 2017), its interstellar detection is elusive despite observational searches (e.g., Jiménez-Serra et al. 2020).

Recently, the experiments on interstellar ice analogues carried out by Kleimeier et al. (2021) have identified the enol

form of glycolaldehyde, (*Z*)-1,2-ethenediol $(CHOH)_2$, a species that might be a key precursor in the formation of sugars in both interstellar and prebiotic chemistry (e.g., Kitadai & Maruyama 2018). However, 1,2-ethenediol has never been searched for in the ISM because its rotational spectrum was unknown until recently. In this Letter, we use the new laboratory spectroscopic measurements performed by Melosso et al. (2022) to provide the first detection of $(CHOH)_2$ in the ISM toward the G+0.693 −0.027 molecular cloud (G+0.693, hereafter) located in the Sgr B2 complex within the Galactic Center.

2. Observations

We have used a high-sensitivity unbiased spectral survey performed with the Yebes 40 m (Guadalajara, Spain) and the IRAM 30 m (Granada, Spain) telescopes. The observations, using position-switching mode, were centered at $\alpha(J2000.0) = 17^h47^m22^s$, $\delta(J2000.0) = -28°21'27''$. The line intensity of the spectra was measured in units of T_A^* because the molecular emission toward G+0.693 is extended over the beams (Requena-Torres et al. 2006, 2008; Zeng et al. 2018, 2020). The IRAM 30 m observations cover the range from 71.770 to 116.720 GHz. The noise of the spectra (in T_A^*) is 1.3–2.8 mK in the 71–90 GHz range, 1.5–5.8 mK in the 90–115 GHz range, and ∼10 mK in the 115–116 GHz range. More detailed information on these IRAM 30 m observations is provided in Rivilla et al. (2021a, 2021b).

We have performed a new deeper spectral survey with the Yebes 40 m telescope as part of the 21A014 project (PI: Rivilla). We observed during 29 different sessions between 2021 March and 2021 June, with a total observing time of ∼125 hr. The Nanocosmos *Q*-band (7 mm) HEMT receiver was used, which provides ultra-broadband observations (18.5 GHz) in two linear polarizations (Tercero et al. 2021). The receiver was connected to 16 FFTS providing a channel width of 38 kHz. We observed two different spectral setups centered at 41.40 and 42.30 GHz, with a total frequency range of 31.07−50.42 GHz. We have performed the initial data inspection and reduction using a Python-based[14] script that uses the CLASS module of the GILDAS package.[15] For each observing day, spectra averaging was performed, and we fitted baselines using an iterative method that first masks the more visible lines using sigma clips and then applies rolling medians and rolling averages, interpolating the masked regions with splines. Spectra were then combined, averaged, and exported to MADCUBA[16] (Martín et al. 2019). The comparison of the spectra of the two frequency setups was used to identify the possible contamination of lines from the image band. Telescope pointing and focus were checked every one or two hours through pseudo-continuum observations. The final spectra were smoothed to 250 kHz. The noise achieved at this spectral resolution is ∼0.5–1.0 mK in T_A^* scale. The half-power beamwidth (HPBW) of the telescope was ∼35″–55″ in the frequency range observed.

3. Analysis and Results

To search for transitions of $(CHOH)_2$ in our spectral survey of the G+0.693 molecular cloud, we have implemented the rotational spectroscopy information from the recent laboratory work by Melosso et al. (2022) into the SLIM (Spectral Line Identification and Modeling) tool within MADCUBA (version 09/11/2021), which generates synthetic spectra under the assumption of local thermodynamic equilibrium (LTE) conditions.

Figure 1 presents the 18 unblended or slightly blended transitions of $(CHOH)_2$ detected toward G+0.693, arising from nine pairs of unresolved transitions due to the line widths of this source. Their spectroscopic information is listed in Table 1. To properly evaluate possible line contamination by other molecules, we have considered the emission from the >120 molecules already identified toward G+0.693 (e.g., Requena-Torres et al. 2006, 2008; Zeng et al. 2018; Jiménez-Serra et al. 2020; Rivilla et al. 2018, 2019b, 2020b, 2021a, 2021b; Rodríguez-Almeida et al. 2021a, 2021b; Zeng et al. 2021; Rivilla et al. 2022). Five pairs of transitions appear unblended, while the other four pairs are slightly blended with thioformic acid (*t*-HCOSH, reported in Rodríguez-Almeida et al. 2021a), propene (C_3H_6), and emission from species not yet identified (marked as U in Table 1). We note that the rest of the $(CHOH)_2$ transitions that fall in the spectral coverage of the survey are consistent with the observed spectra, but they are heavily blended with lines from other molecular species.

To derive the physical parameters of the $(CHOH)_2$ emission, we used the SLIM-AUTOFIT tool of MADCUBA that provides the best nonlinear least-squares LTE fit to the data using the Levenberg–Marquardt algorithm. The parameters used in the LTE model are molecular column density (*N*), excitation temperature (T_{ex}), velocity (v_{LSR}), and FWHM of the Gaussian line profiles. The fit of the $(CHOH)_2$ profiles was performed by also considering the contribution of the other identified molecules. Because the AUTOFIT algorithm did not converge with all four parameters adjustable, we fixed $v_{LSR} = 66$ km s^{-1} and FWHM = 20 km s^{-1}, which reproduce well the line profiles of the $(CHOH)_2$ features shown in Figure 1, and left *N* and T_{ex} as free parameters. The best LTE fit found by MADCUBA is shown by the red curve in Figure 1. In each panel, the individual contribution of the two unresolved $(CHOH)_2$ transitions close in frequency is indicated by the yellow and salmon curves. The blue curve shows the total contribution of all the molecules identified in the survey. We obtained $T_{ex} = 8.5 \pm 0.6$ K, which is in good agreement with the typical values of T_{ex} found in G+0.693, i.e., in the range of 5–20 K (see, e.g., Zeng et al. 2018). The derived column density is $(1.8 \pm 0.1) \times 10^{13}$ cm^{-2}, which translates into a molecular abundance with respect to molecular hydrogen of 1.3×10^{-10}, assuming $N_{H_2} = 1.35 \times 10^{23}$ cm^{-2} toward G+0.693 (Martín et al. 2008).

We have estimated the detection level of the transitions of $(CHOH)_2$ by computing the signal-to-noise ratio derived with the expression:

$$S/N = \left(\int T_A^* dv\right) \Big/ \left[\mathrm{rms}\left(\frac{\Delta v}{\mathrm{FWHM}}\right)^{0.5} \mathrm{FWHM}\right],$$

where $\int T_A^* dv$ is the velocity-integrated intensity of each line derived from the LTE fit, rms is the noise of the spectrum, and Δv and FWHM are the spectral resolution and line width in velocity units, respectively. The rms was measured around each transition in a nearby line-free spectral range, obtaining values in the 0.5–2.5 mK range. The values of Δv are 1.5–2.5 km s^{-1}, while the FWHM used is the one assumed for the LTE fit, i.e., 20 km s^{-1}. Table 1 shows that the unblended $(CHOH)_2$ transitions are detected

[14] https://www.python.org
[15] https://www.iram.fr/IRAMFR/GILDAS
[16] Madrid Data Cube Analysis on ImageJ is a software developed at the Center of Astrobiology (CAB) in Madrid: https://cab.inta-csic.es/madcuba/.

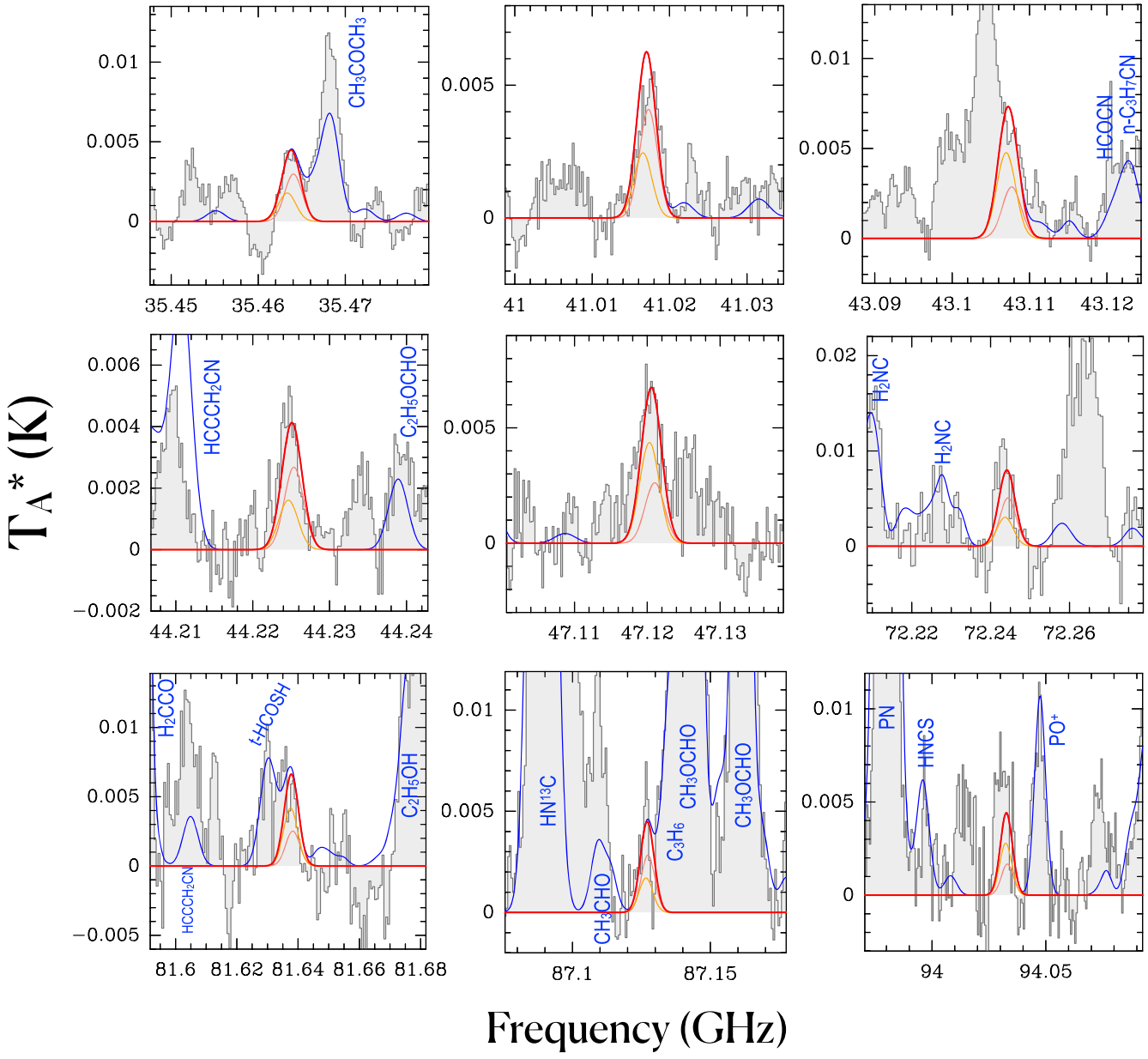

Figure 1. Selected unblended or slightly blended transitions of 1,2-ethenediol detected toward the G+0.693−0.027 molecular cloud. Note that each panel contains two 1,2-ethenediol transitions close in frequency that are not resolved (see Table 1), the individual contributions of which are shown with the yellow and salmon curves. The best LTE fit derived with MADCUBA for the 1,2-ethenediol emission is shown with a red curve. The blue curve represents the total emission considering all the species identified toward this molecular cloud.

at signal-to-noise ratio levels between 4.6 and 20.9, with at least one of the transitions of the pair >6.

We have also performed a complementary analysis using the rotational diagram method implemented in MADCUBA (see more details in Appendix A). Figure A1 shows the rotational diagram obtained using the transitions of $(CHOH)_2$ listed in Table 1. We derived physical parameters that are fully consistent with the MADCUBA−AUTOFIT analysis: $N = (2.1 \pm 0.5) \times 10^{13}$ cm^{-2}, and $T_{ex} = 8.6 \pm 1.0$ K.

We have also analyzed the emission from its isomer glycolaldehyde ($HOCH_2CHO$), which is presented in detail in Appendix B. We obtained a column density of $(9.3 \pm 0.3) \times 10^{13}$ cm^{-2}. The chemically related species (see Section 4) vinyl alcohol (CH_2CHOH) has also been recently detected toward G+0.693 by Jiménez-Serra et al. (submitted). The sum of the column densities of the two conformers identified (*syn* and *anti*) gives 12.4×10^{13} cm^{-2}. Thus, the ratioratios between $HOCH_2CHO$ and CH_2CHOH with respect to $(CHOH)_2$ are ∼5.2 and ∼7.0, respectively. The upper limit of the molecular abundance of glyceraldehyde reported by Jiménez-Serra et al. (2020) is $\sim 0.93 \times 10^{-10}$, only slightly lower than the abundance of $(CHOH)_2$, which suggests that deeper spectral surveys are needed to address its detection.

4. Discussion

4.1. Formation of 1,2-ethenediol in the ISM

1,2-ethenediol is the fourth $C_2H_4O_2$ isomer detected in the ISM, after acetic acid (CH_3COOH), methyl formate (CH_3OCHO), and glycolaldehyde ($HOCH_2CHO$). 1,2-

Table 1
List of Observed Transitions of (*Z*)-1,2-ethenediol $(CHOH)_2$

Frequency (GHz)	Transition (J_{K_a,K_c}, v)	log I (nm^2 MHz)	$S_{ul} \times \mu^2$ (D^2)	log A_{ul} (s^{-1})	g_u	E_u (K)	$\int T_A^* dv$[a] (mK km s^{-1})	rms (mK)	S/N	Blending[b]
35.4628492	$3_{1,2}$, 0–2 $_{1,1}$, 1	−6.0616	10.1267	−6.9024	42	4.1	38.0	0.58	10.4	unblended
35.4635659	$3_{1,2}$, 1–2 $_{1,1}$, 0	−5.8397	10.1267	−7.1244	70	4.1	63.4	0.58	17.3	unblended
41.0160144	$4_{1,4}$, 0–3 $_{1,3}$, 1	−5.7895	14.2307	−6.6743	54	5.6	51.3	0.70	12.6	unblended
41.0167312	$4_{1,4}$, 1–3 $_{1,3}$, 0	−5.5676	14.2308	−6.8962	90	5.6	85.4	0.70	20.9	unblended
43.1064014	$4_{0,4}$, 0–3 $_{0,3}$, 1	−5.4996	15.0483	−6.8072	90	5.3	100.0	0.54	32.2	blended with U
43.1071218	$4_{0,4}$, 1–3 $_{0,3}$, 0	−5.7214	15.0482	−6.5854	54	5.3	60.2	0.54	19.4	blended with U
44.2240556	$4_{2,2}$, 1–3 $_{2,1}$, 0	−5.8240	11.3946	−6.6728	54	8.0	32.6	0.70	8.2	unblended
44.2247658	$4_{2,3}$, 1–3 $_{2,2}$, 0	−5.6021	11.3946	−6.8945	90	8.0	54.3	0.70	13.7	unblended
47.1196699	$4_{1,3}$, 0–3 $_{1,2}$, 1	−5.4482	14.2237	−6.7158	90	6.4	90.4	0.90	18.3	unblended
47.1203898	$4_{1,3}$, 1–3 $_{1,2}$, 0	−5.6701	14.2237	−6.4938	54	6.4	54.3	0.90	11.0	unblended
72.2429524	$7_{0,7}$, 0–6 $_{0,6}$, 1	−5.0451	26.1443	−5.8945	90	14.3	64.8	2.18	4.2	unblended
72.2436790	$7_{0,7}$, 1–6 $_{0,6}$, 0	−4.8233	26.1436	−6.1164	150	14.3	108.0	2.18	7.0	unblended
81.6367065	$8_{0,8}$, 0–7 $_{0,7}$, 1	−4.6642	29.8923	−5.9531	170	18.2	89.3	2.92	4.6	blended with *t*-HCOSH
81.6374347	$8_{0,8}$, 1–7 $_{0,7}$, 0	−4.8861	29.8909	−5.7314	102	18.2	53.6	2.92	2.7	blended with *t*-HCOSH
87.1257785	$8_{2,7}$, 0–7 $_{2,6}$, 1	−4.8570	28.3789	−5.6692	102	21.7	36.3	1.63	3.4	blended with C_3H_6
87.1265029	$8_{2,7}$, 1–7 $_{2,6}$, 0	−4.6351	28.3789	−5.8911	170	21.7	60.5	1.63	5.7	blended with C_3H_6
94.0314593	$8_{2,6}$, 0–7 $_{2,5}$, 1	−4.5668	28.5912	−5.7883	170	22.6	59.3	1.23	7.7	blended with U
94.0321932	$8_{2,6}$, 1–7 $_{2,5}$, 0	−4.7886	28.5912	−5.5666	102	22.6	35.6	1.23	4.6	blended with U

Notes. We provide the transition frequencies, quantum numbers, base 10 logarithm of the integrated intensity at 300 K (log I), the values of $S_{ul} \times \mu^2$, the base 10 logarithm of Einstein coefficients (log A_{ul}), and upper-state degeneracy (g_u). We also report the derived velocity-integrated intensities of each transition, the rms noise of the spectrum, and the signal-to-noise ratio of the transition (see text). The last column gives the information about the species whose transitions are partially blended with the $(CHOH)_2$ lines.

[a] The integrated intensity of the lines has been obtained from the best fit from MADCUBA (see text).

[b] U refers to blending with emission from an unknown (not identified) species.

ethenediol is the highest-energy isomer (Karton & Talbi 2014), with the lowest abundance in G+0.693. This would be in agreement with the minimum energy principle (MEP; Lattelais et al. 2009), which states that the most stable isomers (those with lower energy) are the most abundant in the ISM. However, it is well known that the other three isomers do not follow this rule because the most stable $C_2H_4O_2$ isomer, acetic acid, is less abundant than methyl formate and glycolaldehyde (see, e.g., Mininni et al. 2020 and references therein). Because the difference in energy of the $C_2H_4O_2$ isomers is thousands of Kelvin (Lattelais et al. 2009; Kua et al. 2013), isomerization processes seem highly unlikely in the ISM. Therefore, the molecular abundances of the different isomers are expected to be the consequence of kinetics rather than thermodynamics, namely, they are formed through different chemical pathways that are not directly interconnected. The chemistry of acetic acid, methyl formate, and glycolaldehyde has been extensively discussed in several works (see e.g., Laas et al. 2011; Burke et al. 2015; Skouteris et al. 2018; El-Abd et al. 2019; Ahmad et al. 2020; Mininni et al. 2020; Paulive et al. 2021).

In this section, we focus our discussion on the newly discovered $C_2H_4O_2$ isomer, 1,2-ethenediol, whose formation mechanisms have not been investigated so far. A number of viable routes can be hypothesized, which are depicted in Figure 2. For example, neutral–radical reactions involving formaldehyde and vinyl alcohol:

$$H_2CO + CH_2OH \longrightarrow (CHOH)_2 + H, \quad (1)$$

$$CH_2CHOH + OH \longrightarrow (CHOH)_2 + H, \quad (2)$$

are likely to be barrierless and involve the elimination of a H atom, which removes the excess energy, stabilizing the products. Hence, they can take place both in the gas phase and on grain surfaces. In these routes, two of the precursors (H_2CO and CH_2CHOH) have been detected toward the G+0.693 molecular cloud at relatively high abundances: $\geqslant 3 \times 10^{-9}$ and 9.2×10^{-10} (Jiménez-Serra et al. 2020; I. Jiménez-Serra et al., in preparation, respectively). The hydroxymethyl radical (CH_2OH) has not been detected yet in the ISM, although it is expected to be an important intermediate in the formation of numerous interstellar molecules on the surface of dust grains (Bermudez et al. 2017 and references therein), thus making the surface route (1) a viable formation pathway to $(CHOH)_2$. The hydroxyl radical (OH) is very abundant in the whole Sgr B2 region (Goicoechea & Cernicharo 2002), which makes route (2) a plausible formation pathway of $(CHOH)_2$.

Other reactions may occur on grain surfaces only (see Figure 2), like the dimerization of hydroxymethylene:

$$2CHOH \longrightarrow (CHOH)_2, \quad (3)$$

or the neutral–neutral reaction between hydroxymethylene and formaldehyde:

$$CHOH + H_2CO \longrightarrow (CHOH)_2, \quad (4)$$

Unfortunately, we cannot search for CHOH toward G+0.693 because the required spectroscopic information is not available.

We have also considered another route based on the dimerization of the formyl radical (HCO) to form glyoxal (HCOCHO; Woods et al. 2013) followed by hydrogenation (Figure 2):

$$2HCO \longrightarrow HCOCHO \xrightarrow{2H} (CHOH)_2. \quad (5)$$

HCOCHO has never been detected in the ISM so far. Its lowest-energy conformer (*trans*), which is the one expected to be detected in interstellar conditions (the *cis* form is 2237 K

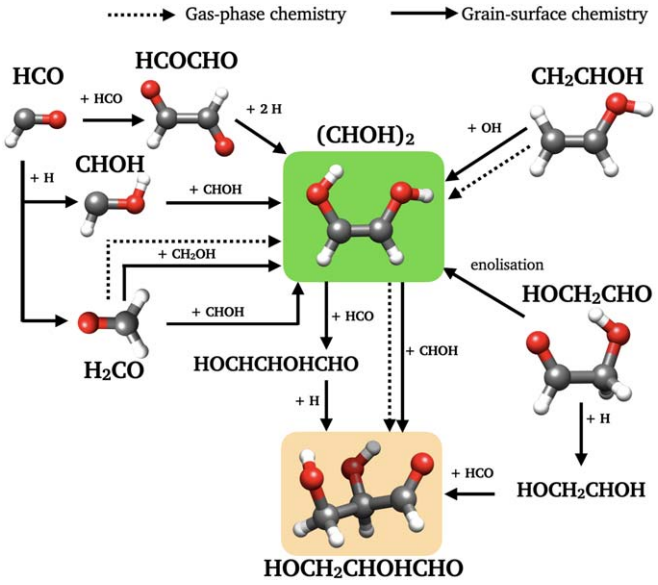

Figure 2. Possible formation routes of 1,2-ethenediol (green box) and proposed chemical pathways for the formation of glyceraldehyde (orange box).

higher in energy), has a vanishing dipole moment because of symmetry, so it cannot be detected by rotational emission. Route (4) was favored by the experimental work by Kleimeier et al. (2021), in which HCOCHO was identified after electron irradiation of methanol-bearing interstellar ice analogues. However, we note that the experiments by Leroux et al. (2020) showed that the hydrogenation of HCOCHO ices does not lead to the formation of more complex species (e.g., glycolaldehyde or 1,2-ethenediol; L. Krim 2022, private communication), but to lighter molecules such as CO and H_2CO. Additional laboratory works, complemented by new quantum mechanical calculations, will help to discern the contribution of the chemical routes proposed (1–5) to the formation of $(CHOH)_2$.

An alternative formation pathway of $(CHOH)_2$ could be the enolization of glycolaldehyde (Figure 2). However, the calculations reported by Kua et al. (2013) pointed out that this enolization requires 39 kJ mol^{-1} (4690 K), which increases to 88 kJ mol^{-1} (10,600 K) in water-assisted isomerization. This indicates that $(CHOH)_2$ cannot be formed through this process in the conditions of the ISM.

4.2. Formation of Complex Sugars in the ISM

Given its prebiotic relevance, the growth of the chemical complexity of sugars in the ISM is widely debated in astrochemistry. In the ISM, both $HOCH_2CHO$ and its enolic form $(CHOH)_2$ can be regarded as the closest precursors of glyceraldehyde (see Figure 2). The laboratory experiments on H bombardment of interstellar ice analogues by Fedoseev et al. (2017) showed that glyceraldehyde formation can take place on grain surfaces via successive neutral–radical reactions:

$$HOCH_2CHO \xrightarrow{H} HOCH2CHOH$$
$$\xrightarrow{HCO} HOCH_2CHOHCHO.$$

We here propose an alternative and likely more favorable route on grains, which involves the highly reactive species $(CHOH)_2$ and CHOH (see Figure 2):

$$(CHOH)_2 + CHOH \longrightarrow HOCH_2CHOHCHO. \quad (6)$$

The advantage of this pathway is that both reactants are less stable than their corresponding keto forms (Schreiner et al. 2008; Karton & Talbi 2014), which favors overcoming the reaction barrier and enhances the overall process exothermicity.

Therefore, the detection of the 2-carbon sugar glycolaldehyde and its enol form in the ISM opens the possibility of detecting more complex sugars. Several interstellar searches of 3-carbons sugars, already identified in meteorites (Cooper et al. 2001), including the aldose glyceraldehyde[17] (Hollis et al. 2004), and the ketose[18] dihydroxyacetone (Widicus Weaver & Blake 2005; Apponi et al. 2006), indicated that they are at least a factor of 5 less abundant than glycolaldehyde. We also searched for these 3-carbon sugars toward G + 0.693 (Jiménez-Serra et al. 2020) and obtained that glyceraldehyde and dihydroxyacetone are less abundant than glycolaldehyde by factors >7 and >13, respectively. This indicates that a 1-carbon increase in chemical complexity of sugars implies a molecular abundance drop of at least one order of magnitude, as previously observed in other chemical families like alcohols, thiols, isocyanates (Rodríguez-Almeida et al. 2021a, 2021b; Jiménez-Serra et al., submitted), and aldehydes (Sanz-Novo et al. 2022).

4.3. Implications for the RNA World

The backbone of RNA is composed of molecules of ribose (c-$C_5H_{10}O_5$) linked by phosphate groups. Initially, it was proposed that ribose, along with other sugars, might be formed through the *formose* reaction (Butlerow 1861), starting from aqueous formaldehyde (H_2CO) and glycolaldehyde ($HOCH_2CHO$). However, stability and stereochemical purity problems (Larralde et al. 1995; Toxvaerd 2013) cast doubts about the efficiency of formation of sugars in a prebiotic context. Moreover, the formation of RNA nucleotides by adding ribose to nucleobases is highly inefficient (Fuller et al. 1972; Orgel 2004). As a consequence, alternative synthesis routes are needed. Several prebiotic experiments (Powner et al. 2009; Becker et al. 2019) have found viable synthesis of RNA nucleotides starting from simple precursors such as glycolaldehyde and glyceraldehyde. In this context, the new detection of (*Z*)-1,2-ethenediol in the ISM reported here is especially relevant for both extraterrestrial and terrestrial environments. First, this species can be a direct precursor of glyceraldehyde in interstellar ices, as discussed in Section 4.2. Second, it has been proposed that $(CHOH)_2$ is the key intermediate in the formation of glyceraldehyde starting from glycolaldehyde under plausible early-Earth conditions (Kitadai & Maruyama 2018):

$$HOCH_2CHO \longrightarrow (CHOH)_2 \longrightarrow HOCH_2CHOHCHO$$

Therefore, the detection of 1,2-ethenediol ($(CHOH)_2$) toward the G+0.693 molecular cloud, along with those of its tautomer glycolaldehyde ($HOCH_2CHO$) in several other sources, indicates that immediate molecular precursors of more complex species like glyceraldehyde are available in the ISM. These compounds, if delivered to the surface of planets and moons through meteoritic and cometary impacts, could have triggered

[17] An aldose is a sugar consisting of a carbon backbone and a carbonyl group at the end of the carbon chain, resulting in an aldehyde group.
[18] A ketose is a sugar consisting of a carbon backbone and a carbonyl group within the backbone.

prebiotic chemistry by providing key precursors of RNA nucleotides.

We acknowledge the two anonymous reviewers for their careful reading of the manuscript and their useful comments. We are grateful to the IRAM 30 m and Yebes 40 m telescope staff for their help during the different observing runs. The 40 m radio telescope at Yebes Observatory is operated by the Spanish Geographic Institute (IGN, Ministerio de Transportes, Movilidad y Agenda Urbana). I.R.A.M. is supported by INSU/CNRS (France), MPG (Germany) and IGN (Spain). V.M.R., L.C., and A.L.-G. have received funding from the Comunidad de Madrid through the Atracción de Talento Investigador (Doctores con experiencia) Grant (COOL: Cosmic Origins Of Life; 2019-T1/TIC-15379). I.J.-S., J.M.-P., L.C., and A.M. have received partial support from the Spanish project numbers PID2019-105552RB-C41 and MDM-2017-0737 (Unidad de Excelencia María de Maeztu−Centro de Astrobiología, INTA-CSIC). A.M. has received support from the Spanish project number MDM-2017-0737-19-2 and grant PRE2019-091471, funded by MCIN/AEI/10.13039/501100011033 and by "ESF Investing in your future". P.d.V. and B.T. thank the support from the European Research Council through Synergy Grant ERC-2013-SyG, G.A. 610256 (NANOCOSMOS) and from the Spanish Ministerio de Ciencia e Innovación (MICIU) through project PID2019-107115GB-C21. B.T. also thanks the Spanish MICIU for funding support from grant PID2019-106235GB-I00. J.-C.G. thanks the Centre National d'Etudes Spatiales (CNES) for a grant.

Appendix A
Rotational Diagram of (*Z*)-1,2-ethenediol, $(CHOH)_2$

We have performed a rotational diagram following the standard procedure (Goldsmith & Langer 1999) implemented in MADCUBA and described in detail in Rivilla et al. (2021a). We have used FWHM = 20 km s^{-1} and v_{LSR} = 66 km s^{-1}, as in the AUTOFIT analysis. The resulting diagram is shown in Figure A1. We have obtained $N = (2.1 \pm 0.5) \times 10^{13}$ cm^{-2} and $T_{ex} = 8.6 \pm 1.0$ K, which are in agreement within the uncertainties with those derived with AUTOFIT (Section 3).

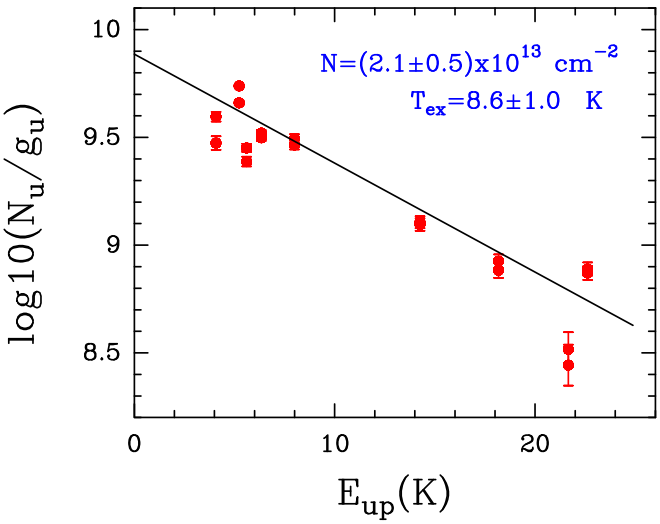

Figure A1. Rotational diagram of (*Z*)-1,2-ethenediol. The red dots correspond to the transitions listed in Table 1. The black line is the best linear fit to the data points. The derived values for the molecular column density (N) and the excitation temperature (T_{ex}), along with their uncertainties, are indicated in blue in the upper-right corner.

Appendix B
LTE Fit of Glycolaldehyde ($HOCH_2CHO$)

Multiple transitions of glycolaldehyde ($HOCH_2CHO$) are detected within the spectral survey toward G+0.693. We have used the Cologne Database for Molecular Spectroscopy (CDMS; Endres et al. 2016) entry 060501 (June 2021). To perform the LTE fit, we have selected completely unblended transitions that cover a broad range of E_{up} from 4 to 51 K, so that T_{ex} can be directly constrained. To evaluate that the

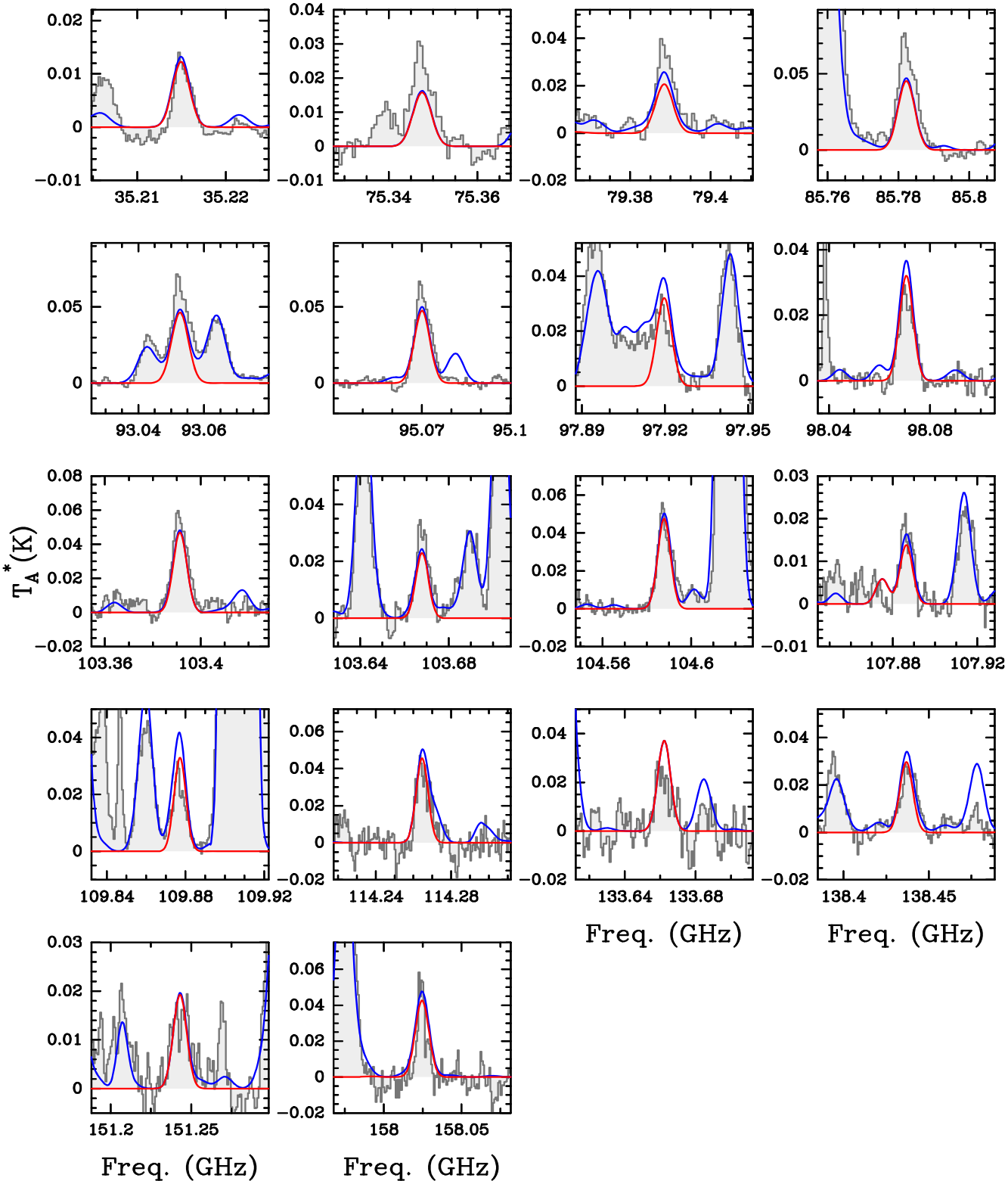

Figure B1. Selected unblended transitions of $HOCH_2CHO$ identified toward the G+0.693−0.027 molecular cloud and employed to perform the LTE fit. The corresponding spectroscopic information is provided in Table B1. The best LTE fit derived with MADCUBA is shown with a red curve, while the blue curve depicts the total emission considering all the species identified toward this molecular cloud.

Table B1
Selected Transitions of $HOCH_2CHO$ Used to Perform the LTE Fit

Frequency (GHz)	Transition ($J_{Ka,Kc}$)	log I (nm^2 MHz)	$S_{ul} \times \mu^2$ (D^2)	log A_{ul} (s^{-1})	g_u	E_u (K)
35.2148100	$6_{1,5}-6_{0,6}$	−5.4016	17.948	−6.1538	13	12.9
75.3474860	$8_{1,7}-7_{2,6}$	−4.8156	15.498	−5.3430	17	21.4
79.3884359	$4_{2,3}-3_{1,2}$	−4.9263	10.349	−5.1741	9	7.9
85.7822074	$8_{1,8}-7_{0,7}$	−4.3514	34.509	−4.8265	17	18.9
93.0526650	$9_{0,9}-8_{1,8}$	−4.2245	39.850	−4.7062	19	23.4
95.0700722	$9_{1,9}-8_{0,8}$	−4.2039	40.025	−4.6763	19	23.4
97.9196993	$3_{3,1}-2_{2,0}$	−4.6336	13.362	−4.6807	7	8.8
98.0705529	$3_{3,0}-2_{2,1}$	−4.6328	13.343	−4.6792	7	8.8
103.3913370	$10_{0,10}-9_{1,9}$	−4.0826	45.462	−4.5552	21	28.4
103.6679627	$10_{1,9}-9_{2,8}$	−4.3248	26.204	−4.7910	21	32.0
104.5877362	$10_{1,10}-9_{0,9}$	−4.0718	45.551	−4.5394	21	28.3
107.8863063	$5_{2,3}-4_{1,4}$	−4.9176	5.759	−5.1162	11	10.9
109.8771836	$4_{3,1}-3_{2,2}$	−4.5319	13.495	−4.6354	9	11.0
114.2645516	$11_{1,11}-10_{0,10}$	−3.9527	51.066	−4.4140	23	33.8
133.6622379	$13_{0,13}-12_{1,12}$	−3.7493	62.041	−4.1947	27	46.2
138.4367820	$7_{3,5}-6_{2,4}$	−4.2955	15.114	−4.5070	15	21.0
151.2430164	$14_{1,13}-13_{2,12}$	−3.7535	49.939	−4.1590	29	58.6
158.0246068	$6_{4,2}-5_{3,3}$	−4.0767	19.193	−4.1687	13	21.2

Note. We provide the transition frequencies, quantum numbers, base 10 logarithm of the integrated intensity at 300 K (log I), the values of $S_{ul} \times \mu^2$, the base 10 logarithm of the Einstein coefficients (log A_{ul}), and upper-state degeneracy (g_u).

transitions are not contaminated, we have used, as explained in the text (Section 3), the line emissions from all the molecules identified in the source. We have run MADCUBA using the transitions shown in Table B1. We have left free to vary the four model parameters. We obtain $N = (9.3 \pm 0.3) \times 10^{13}$ cm^{-2}, $T_{ex} = 21.8 \pm 0.8$ K, $v_{LSR} = 68.9 \pm 0.2$ km s^{-1}, and FWHM $= 19.4 \pm 0.5$ km s^{-1}. The resulting best LTE fit is shown in red in Figure B1.

ORCID iDs

Víctor M. Rivilla https://orcid.org/0000-0002-2887-5859
Laura Colzi https://orcid.org/0000-0001-8064-6394
Izaskun Jiménez-Serra https://orcid.org/0000-0003-4493-8714
Jesús Martín-Pintado https://orcid.org/0000-0003-4561-3508
Andrés Megías https://orcid.org/0000-0002-6389-7172
Mattia Melosso https://orcid.org/0000-0002-6492-5921
Luca Bizzocchi https://orcid.org/0000-0002-9953-8593
Álvaro López-Gallifa https://orcid.org/0000-0001-6049-9366
Antonio Martínez-Henares https://orcid.org/0000-0001-5191-2075
Sarah Massalkhi https://orcid.org/0000-0002-7387-9787
Belén Tercero https://orcid.org/0000-0002-4782-5259
Pablo de Vicente https://orcid.org/0000-0002-5902-5005
Jean-Claude Guillemin https://orcid.org/0000-0002-2929-057X
Juan García de la Concepción https://orcid.org/0000-0001-6484-9546
Fernando Rico-Villas https://orcid.org/0000-0002-5351-3497
Shaoshan Zeng https://orcid.org/0000-0003-3721-374X
Sergio Martín https://orcid.org/0000-0001-9281-2919
Francesca Tonolo https://orcid.org/0000-0002-9555-7834
Silvia Alessandrini https://orcid.org/0000-0003-3152-3261
Luca Dore https://orcid.org/0000-0002-1009-7286
Vincenzo Barone https://orcid.org/0000-0001-6420-4107
Cristina Puzzarini https://orcid.org/0000-0002-2395-8532

References

Ahmad, A., Misra, S. A., & Tandon, P. 2020, RAA, 20, 014
Apponi, A. J., Halfen, D. T., Ziurys, L. M., et al. 2006, ApJL, 643, L29
Becker, S., Feldmann, J., Wiedemann, S., et al. 2019, Sci, 366, 76
Becker, S., Thoma, I., Deutsch, A., et al. 2016, Sci, 352, 833
Beltrán, M. T., Codella, C., Viti, S., Neri, R., & Cesaroni, R. 2009, ApJL, 690, L93
Benner, S. A., Kim, H.-J., & Biondi, E. 2019, Life, 9, 84
Bennett, C. J., & Kaiser, R. I. 2007, ApJ, 661, 899
Bermudez, C., Bailleux, S., & Cernicharo, J. 2017, A&A, 598, A9
Burke, D. J., Puletti, F., Brown, W. A., et al. 2015, MNRAS, 447, 1444
Butlerow, A. 1861, Justus Liebigs Ann. Chem., 120, 295
Butscher, T., Duvernay, F., Theule, P., et al. 2015, MNRAS, 453, 1587
Chyba, C., & Sagan, C. 1992, Natur, 355, 125
Cooper, G., Kimmich, N., Belisle, W., et al. 2001, Natur, 414, 879
Coutens, A., Viti, S., Rawlings, J. M. C., et al. 2018, MNRAS, 475, 2016
de Marcellus, P., Meinert, C., Myrgorodska, I., et al. 2015, PNAS, 112, 965
El-Abd, S. J., Brogan, C. L., Hunter, T. R., et al. 2019, ApJ, 883, 129
Endres, C. P., Schlemmer, S., Schilke, P., Stutzki, J., & Müller, H. S. P. 2016, JMoSp, 327, 95
Fedoseev, G., Chuang, K. J., Ioppolo, S., et al. 2017, ApJ, 842, 52
Fedoseev, G., Cuppen, H. M., Ioppolo, S., Lamberts, T., & Linnartz, H. 2015, MNRAS, 448, 1288
Fuller, W. D., Sanchez, R. A., & Orgel, L. E. 1972, J. Mol. Biol., 67, 25
Furukawa, Y., Chikaraishi, Y., Ohkouchi, N., et al. 2019, PNAS, 116, 24440
Gilbert, W. 1986, Natur, 319, 618
Goicoechea, J. R., & Cernicharo, J. 2002, ApJL, 576, L77
Goldsmith, P. F., & Langer, W. D. 1999, ApJ, 517, 209
Hollis, J. M., Jewell, P. R., Lovas, F. J., Remijan, A., & Møllendal, H. 2004, ApJL, 610, L21
Hollis, J. M., Lovas, F. J., & Jewell, P. R. 2000, ApJL, 540, L107
Jiménez-Serra, I., Martí-Pintado, J., Rivilla, V. M., et al. 2020, AsBio, 20, 1048
Jorgensen, J. K., Favre, C., Bisschop, S. E., et al. 2012, ApJ, 757, L4
Karton, A., & Talbi, D. 2014, CP, 436, 22
Kitadai, N., & Maruyama, S. 2018, Geosci. Front., 9, 1117
Kleimeier, N. F., Eckhardt, A. K., & Kaiser, R. I. 2021, JAChS, 143, 14009
Kua, J., Avila, J. E., Lee, C. G., & Smith, W. D. 2013, JPCA, 117, 12658

Laas, J. C., Garrod, R. T., Herbst, E., & Widicus Weaver, S. L. 2011, ApJ, 728, 71
Larralde, R., Robertson, M. P., & Miller, S. L. 1995, PNAS, 92, 8158
Lattelais, M., Pauzat, F., Ellinger, Y., & Ceccarelli, C. 2009, ApJL, 696, L133
Leroux, K., Guillemin, J.-C., & Krim, L. 2020, MNRAS, 491, 289
Marchi, S., Bottke, W., Elkins-Tanton, L., et al. 2014, Natur, 511, 578
Martín, S., Martín-Pintado, J., Blanco-Sánchez, C., et al. 2019, A&A, 631, A159
Martín, S., Requena-Torres, M. A., Martín-Pintado, J., & Mauersberger, R. 2008, ApJ, 678, 245
Meinert, C., Myrgorodska, I., De Marcellus, P., et al. 2016, Sci, 352, 208
Melosso, M., Bizzocchi, L., Gazzeh, H., et al. 2022, ChemComm, 58, 2750
Mininni, C., Beltrán, M. T., Rivilla, V. M., et al. 2020, A&A, 644, A84
Orgel, L. E. 2004, Crit. Rev. Biochem. Mol. Biol., 39, 99
Oró, J. 1961, Natur, 191, 1193
Pasek, M. A. 2008, PNAS, 105, 853
Patel, B. H., Percivalle, C., Ritson, D. J., Duffy, C. D., & Sutherland, J. D. 2015, NatCh, 7, 301
Paulive, A., Shingledecker, C. N., & Herbst, E. 2021, MNRAS, 500, 3414
Pearce, B. K., Pudritz, R. E., Semenov, D. A., & Henning, T. K. 2017, PNAS, 114, 11327
Pizzarello, S., Schrader, D. L., Monroe, A. A., & Lauretta, D. S. 2012, PNAS, 109, 11949
Powner, M. W., Gerland, B., & Sutherland, J. D. 2009, Natur, 459, 239
Requena-Torres, M. A., Martín-Pintado, J., Martín, S., & Morris, M. R. 2008, ApJ, 672, 352
Requena-Torres, M. A., Martín-Pintado, J., Rodríguez-Franco, A., et al. 2006, A&A, 455, 971
Rivilla, V. M., Beltrán, M. T., Vasyunin, A., et al. 2019a, MNRAS, 483, 806
Rivilla, V. M., Drozdovskaya, M. N., Altwegg, K., et al. 2020a, MNRAS, 492, 1180
Rivilla, V. M., García de la Concepción, J., Jiménez-Serra, I., et al. 2022, arXiv:2202.13928
Rivilla, V. M., Jiménez-Serra, I., García de la Concepción, J., et al. 2021b, MNRAS, 506, L79
Rivilla, V. M., Jiménez-Serra, I., Martín-Pintado, J., et al. 2021a, PNAS, 118, 2101314118
Rivilla, V. M., Jiménez-Serra, I., Zeng, S., et al. 2018, MNRAS, 474, L30
Rivilla, V. M., Martín-Pintado, J., Jiménez-Serra, I., et al. 2019b, MNRAS, 483, L114
Rivilla, V. M., Martín-Pintado, J., Jiménez-Serra, I., et al. 2020b, ApJL, 899, L28
Rodríguez-Almeida, L. F., Jiménez-Serra, I., Rivilla, V. M., et al. 2021a, ApJL, 912, L11
Rodríguez-Almeida, L. F., Rivilla, V. M., Jiménez-Serra, I., et al. 2021b, A&A, 654, L1
Rubin, M., Bekaert, D. V., Broadley, M. W., Drozdovskaya, M. N., & Wampfler, S. F. 2019, ESC, 3, 1792
Ruiz-Mirazo, K., Briones, C., & de la Escosura, A. 2014, ChRv, 114, 285
Sanz-Novo, M., Belloche, A., Rivilla, V. M., et al. 2022, arXiv:2203.07334
Schreiner, P. R., Reisenauer, H. P., Pickard Iv, F. C., et al. 2008, Natur, 453, 906
Skouteris, D., Balucani, N., Ceccarelli, C., et al. 2018, ApJ, 854, 135
Tercero, F., López-Pérez, J. A., Gallego, J. D., et al. 2021, A&A, 645, A37
Toxvaerd, S. 2013, OLEB, 43, 391
Widicus Weaver, S. L., & Blake, G. A. 2005, ApJL, 624, L33
Woods, P. M., Kelly, G., Viti, S., et al. 2012, ApJ, 750, 19
Woods, P. M., Slater, B., Raza, Z., et al. 2013, ApJ, 777, 90
Zeng, S., Jiménez-Serra, I., Rivilla, V. M., et al. 2018, MNRAS, 478, 2962
Zeng, S., Jiménez-Serra, I., Rivilla, V. M., et al. 2021, ApJL, 920, L27
Zeng, S., Zhang, Q., Jiménez-Serra, I., et al. 2020, MNRAS, 497, 4896