

Formation of ethylene glycol and other complex organic molecules in star-forming regions[★]

V. M. Rivilla¹, M. T. Beltrán¹, R. Cesaroni¹, F. Fontani¹, C. Codella¹, and Q. Zhang²

¹ Osservatorio Astrofisico di Arcetri, Largo Enrico Fermi 5, 50125 Firenze, Italia
e-mail: rivilla@arcetri.astro.it

² Harvard-Smithsonian Center for Astrophysics, 60 Garden St., Cambridge, MA 02138, USA

Received 23 February 2016 / Accepted 23 August 2016

ABSTRACT

Context. The detection of complex organic molecules related with prebiotic chemistry in star-forming regions allows us to investigate how the basic building blocks of life are formed.

Aims. Ethylene glycol (CH₂OH)₂ is the simplest sugar alcohol and the reduced alcohol of the simplest sugar glycolaldehyde (CH₂OHCHO). We study the molecular abundance and spatial distribution of (CH₂OH)₂, CH₂OHCHO and other chemically related complex organic species (CH₃OCHO, CH₃OCH₃, and C₂H₅OH) towards the chemically rich massive star-forming region G31.41+0.31.

Methods. We analyzed multiple single-dish (Green Bank Telescope and IRAM 30 m) and interferometric (Submillimeter Array) spectra towards G31.41+0.31, covering a range of frequencies from 45 to 258 GHz. We fitted the observed spectra with a local thermodynamic equilibrium (LTE) synthetic spectra, and obtained excitation temperatures and column densities. We compared our findings in G31.41+0.31 with the results found in other environments, including low- and high-mass star-forming regions, quiescent clouds and comets.

Results. We report for the first time the presence of the aGg' conformer of (CH₂OH)₂ towards G31.41+0.31, detecting more than 30 unblended lines. We also detected multiple transitions of other complex organic molecules such as CH₂OHCHO, CH₃OCHO, CH₃OCH₃, and C₂H₅OH. The high angular resolution images show that the (CH₂OH)₂ emission is very compact, peaking towards the maximum of the 1.3 mm continuum. These observations suggest that low abundance complex organic molecules, like (CH₂OH)₂ or CH₂OHCHO, are good probes of the gas located closer to the forming stars. Our analysis confirms that (CH₂OH)₂ is more abundant than CH₂OHCHO in G31.41+0.31, as previously observed in other interstellar regions. Comparing different star-forming regions we find evidence of an increase of the (CH₂OH)₂/CH₂OHCHO abundance ratio with the luminosity of the source. The CH₃OCH₃/CH₃OCHO and (CH₂OH)₂/C₂H₅OH ratios are nearly constant with luminosity. We also find that the abundance ratios of pairs of isomers (CH₂OHCHO/CH₃OCHO and C₂H₅OH/CH₃OCH₃) decrease with the luminosity of the sources.

Conclusions. The most likely explanation for the behavior of the (CH₂OH)₂/CH₂OHCHO ratio is that these molecules are formed by different chemical formation routes not directly linked, although different formation and destruction efficiencies in the gas phase cannot be ruled out. The most likely formation route of (CH₂OH)₂ is by combination of two CH₂OH radicals on dust grains. We also favor that CH₂OHCHO is formed via the solid-phase dimerization of the formyl radical HCO. The interpretation of the observations also suggests a chemical link between CH₃OCHO and CH₃OCH₃, and between (CH₂OH)₂ and C₂H₅OH. The behavior of the abundance ratio C₂H₅OH/CH₃OCH₃ with luminosity may be explained by the different warm-up timescales in hot cores and hot corinos.

Key words. astrobiology – astrochemistry – line: identification – molecular data – opacity – ISM: molecules

1. Introduction

In the past few years, the improvement of the sensitivity of current radio and (sub)millimeter telescopes has allowed us to detect molecular species with an increasing number of atoms in the interstellar medium (ISM). From the astrophysical point of view, molecules with ≥ 6 atoms containing carbon are considered complex organic molecules (COMs). These molecules are expected to play an important role in prebiotic chemistry and are considered the basic ingredients to explain the origin of life. Therefore, understanding how these COMs are formed is a first and unavoidable step to ascertain how life was able to emerge in the Universe. The detection of COMs in the gas environment

where low- and high-mass stars are forming (known as hot corinos and hot cores, respectively) indicates that they are part of the material of which stars, planets, and comets are made. Did the Earth inherit the chemical composition from its parental ISM? Were the organic compounds delivered into the early Earth by a rain of meteorites? The answers to these open questions are only possible by studying the chemical compositions of comets and star-forming regions, and figuring out how they were produced.

The formation of COMs is being intensively debated in astrochemistry. Two possible routes, which can be complementary, have been proposed: gas-phase chemical reactions and reactions on the surface of interstellar dust grains. Only the detection of numerous COMs and the study of their relative abundances and spatial distributions in a large sample of star-forming regions will help us to constrain the chemical pathways leading to the formation of COMs.

* The reduced spectra (ASCII files) are only available at the CDS via anonymous ftp to cdsarc.u-strasbg.fr (130.79.128.5) or via <http://cdsarc.u-strasbg.fr/viz-bin/qcat?J/A+A/598/A59>

The basic building blocks of biochemistry – amino acids (precursors of proteins), monosaccharides (the simplest sugars), and lipids – are found among the COMs. So far, only the interstellar search for members of the sugar family has been successful. Sugars are a key ingredient in astrobiology because they are associated with both metabolism and genetic information. The simplest sugar, glycolaldehyde (CH_2OHCHO , hereafter GA), and the simplest sugar alcohol, ethylene glycol ($(\text{CH}_2\text{OH})_2$, hereafter EG), are among the largest molecules detected so far in the ISM (8 and 10 atoms, respectively). It has been proposed that GA is involved in the formation of amino acids and more complex sugars. GA can react with another sugar, propenal, to produce ribose, the central constituent of the backbone of ribonucleic acid RNA (Collins & Ferrier 1995; Weber 1998; Krishnamurthy et al. 1999).

EG is the reduced alcohol of GA, and it was first found in the ISM towards the Galactic Center (Hollis et al. 2002; Requena-Torres et al. 2008; Belloche et al. 2013). It has also been reported towards NGC 1333–IRAS 2A (Maury et al. 2014), NGC 7129 FIRS2 (Fuente et al. 2014), and towards the Orion, W51e1e2 and G34.3+0.1 hot cores (Brouillet et al. 2015; Lykke et al. 2015). Additionally, EG has also been detected in the comets Hale-Bopp (Crovisier et al. 2004), Lemmon and Lovejoy (Biver et al. 2014, 2015) and 67P/Churyumov-Gerasimenko (Goesmann et al. 2015), and in the meteorites Murchinson and Murray (Cooper et al. 2001).

Theoretical chemical modeling and experimental works have proposed different mechanisms for the formation of EG (and other COMs) in the ISM (Sorrell 2001; Charnley & Rodgers 2005; Bennett & Kaiser 2007; Garrod et al. 2008; Woods et al. 2012, 2013; Fedoseev et al. 2015; Butscher et al. 2015). In order to constrain the most likely chemical pathway, it is essential to confirm the presence of these COMs in additional interstellar sources, and to study their spatial distribution and relative abundances. In particular, the massive star-forming regions are a very suitable laboratory to study molecular complexity. The radiation from the newly formed massive stars warms up the grain mantles of the interstellar dust, triggering the desorption of molecules to the gas phase and enriching the chemical environment.

In this work we present single-dish and interferometric observations towards the G31.41+0.31 massive star-forming region (hereafter G31). G31, located at a distance of 7.9 kpc (Churchwell et al. 1990), harbors a luminous hot molecular core ($\sim 3 \times 10^5 L_\odot$, Beltrán & de Wit 2016) believed to be heated by several massive protostars. Many COMs have already been detected towards G31 (see, e.g., Beltrán et al. 2005; Fontani et al. 2007; Isokoski et al. 2013): methanol (CH_3OH), ethanol ($\text{C}_2\text{H}_5\text{OH}$), ethyl cyanide ($\text{C}_2\text{H}_5\text{CN}$), methyl formate (CH_3OCHO), and dimethyl ether (CH_3OCH_3). This chemical richness, along with the previous detection of GA (Beltrán et al. 2009), makes G31 an excellent target to search for even more complex species, and EG in particular.

We report on the detection of the aGg' conformer of EG towards G31. We note that this is the first hot core outside the GC where both EG and GA have been detected. In Sect. 2 we present the different observational campaigns. Section 3 explains the procedures and tools used to identify the lines. The results of the study, including the derivation of physical parameters and spatial distribution of the emission are presented in Sect. 4. In Sect. 5 we compare our results in G31 with other interstellar regions and comets. In Sect. 6 we discuss the implications of our findings on the formation pathways of EG and GA and other COMs. Finally, we summarize our conclusions in Sect. 7.

2. Observations

In this work we have used observations performed with the single-dish telescopes GBT and IRAM 30 m, and the SMA interferometer, which provide us with wide spectral coverage and spatial information. Technical details of the observations are given in the following sections.

2.1. Green Bank Telescope (GBT) single-dish observations

The observations were carried out in May 2011 using the Q -band receiver (7 mm). We used the maximum bandwidth provided by the GBT spectrometer (~ 800 MHz), covering the range 45.14–45.92 GHz. To avoid irregular band-pass shapes related to the wide bandwidth spectrometer mode, we used the subreflector nodding technique, in which the source is positioned first in one receiver feed and then in the other. Moving the subreflector we minimized the time spent between feeds. The advantage of this technique is that one feed is always observing the source and the calibration can be done more accurately. We observed two overlap spectral windows with dual polarization, with a spectral resolution of 0.39 MHz (2.5 km s^{-1}). The system temperature was ~ 80 K. We applied out-of-focus (OOF) holography to correct large-scale errors in the shape of the reflecting surface at the beginning of the observing run. The reduction was done with the GBT-IDL package, and the data were exported to *MADCUBAIJ*¹ for the analysis (see, e.g., Rivilla et al. 2016).

The GBT telescope provided the spectra in antenna temperature below the atmosphere, T_a . We converted this scale to antenna temperature above the atmosphere, T_a^* , using the expression $T_a = T_a^* \times \exp(-\tau/m_{\text{air}})$, where τ is the opacity at zenith and m_{air} is the air mass, which was estimated with $1/\sin(\text{elevation})$. Then, T_a^* was converted to main beam temperature with $T_{\text{mb}} = T_a^*/B_{\text{eff}}$, considering a beam efficiency of $B_{\text{eff}} = 0.83$.

2.2. IRAM 30 m single-dish observations

We used data from two different observational campaigns performed with the IRAM 30 m telescope at Pico Veleta (Spain) on July 1996 and September 2014. The phase center, frequency coverage, and beams of these observations are given in Table 1. The 1996 observations were carried out simultaneously at 3, 2, and 1.3 mm. For more details, see Fontani et al. (2007). The data were reduced using the software package CLASS of GILDAS and then exported to *MADCUBAIJ* for analysis.

The 2014 observations used the Eight Mixer Receiver (EMIR) and the Fast Fourier Transform Spectrometer (FTS) covering the frequency ranges 81.23–89.01 GHz (3 mm) and 167.93–175.71-GHz (2 mm). The spectral resolution was 0.195 kHz (0.34 km s^{-1} and 0.67 km s^{-1} for 2 and 3 mm, respectively). The system temperatures were ~ 130 K and ~ 900 K for the 3 and 2 mm receivers, respectively. The spectra were exported from CLASS to *MADCUBAIJ*, which was used for the reduction and analysis. To increase the signal-to-noise in the 2 mm spectra, we smoothed the data to a velocity resolution of 1 km s^{-1} .

The line intensity of the IRAM 30 m spectra was converted to the main beam temperature T_{mb} , which was calculated as: $T_{\text{mb}} = T_a^* F_{\text{eff}}/B_{\text{eff}}$, where T_a^* is the antenna temperature, and F_{eff} and B_{eff} are the forward and beam efficiencies, respectively. For the

¹ Madrid Data Cube Analysis on ImageJ is a software developed by the Center of Astrobiology (Madrid, INTA-CSIC) to visualize and analyze single spectra and datacubes (Martín et al., in prep.).

Table 1. Summary of observations towards hot molecular core G31 used in this work.

Telescope	Date	Phase center		Frequency coverage (GHz)	θ_{beam} (")
		RA _{J2000} (h m s)	Dec _{J2000} (° ' ")		
IRAM 30 m	1996 July	18 47 34.4	−01 12 46.0	86–258 ¹	12–22
	2014 Sep.	18 47 34.3	−01 12 45.9	81.2–89.0, 167.93–175.71	28, 14
GBT	2011 May	18 47 34.3	−01 12 45.8	45.14–45.92	16
SMA	2007 May & July	18 47 34.3	−01 12 45.9	219.45–221.45, 229.35–231.35	1.7 × 3.5 (comp.) 0.90 × 0.75 (very ext.+ comp.)

Notes. ⁽¹⁾ This range was not fully covered, only some narrower spectral windows.

Table 2. Summary of the detections of the different COMs in the different telescope datasets.

Telescope	(CH ₂ OH) ₂ (EG)	C ₂ H ₅ OH (ET)	CH ₂ OHCHO (GA)	CH ₃ OCHO (MF)	DME (CH ₃ OCH ₃)
GBT	✓	✓	– ^a	✓	– ^a
IRAM 30 m	✓	✓	✓	✓	✓
SMA	✓	✓	✓	✓	✓

Notes. ^(a) The non-detections appear because there are no bright lines of these species in the frequency range covered by the GBT observations.

2014 observations we used the ratios $F_{\text{eff}}/B_{\text{eff}}$ 1.17 and 1.37 for 3 and 2 mm, respectively. The beams of the observations can be estimated using the expression: $\theta_{\text{beam}}[\text{arcsec}] = 2460/\nu(\text{GHz})$.

2.3. Submillimeter Array (SMA) interferometric observations

The observations² were carried out in the 230 GHz band in the compact and very extended configurations in July and May 2007, respectively. The correlator was configured to a spectral resolution of 0.6 km s^{−1} over both sub-bands, from 219.3–221.3 GHz (lower sideband, LSB) and 229.3–231.3 GHz (upper sideband, USB). The visibility data were calibrated using MIR and MIRIAD and the imaging was done with MIRIAD. The resulting synthesized beam is approximately 0.90'' × 0.75'' (PA = 53°) for the combined very extended and compact data and 1.7'' × 3.5'' (PA = 59°) for the compact data. For more details see Cesaroni et al. (2011).

3. Line identification: Ethylene glycol and other COMs

A large frequency coverage is needed to robustly identify large COMs, especially in hot molecular cores where the density of lines in the spectra is high. To assure an identification at a 99%(99.8%) confidence level, at least 6(8) unblended lines are needed (Halfen et al. 2006). An additional condition is that *all* the detectable lines predicted by a local thermodynamic equilibrium (LTE) analysis that lie in the observed spectra must be revealed. Moreover, the relative intensities of the different transitions must be consistent with the LTE analysis, given the high-density environment of hot molecular cores. Furthermore, high spatial resolution data allow us to reduce the confusion limit of single-dish observations because different molecules may exhibit different spatial distributions.

The spectra of G31 exhibits a plethora of molecular lines of multiple species. Although a complete identification and

analysis of all the detected lines is planned for a forthcoming paper (Rivilla et al., in prep.), we study here the emission of selected COMs, including EG. Our motivation is to compare their physical parameters and spatial distributions. These other complex species are: glycolaldehyde (GA), ethanol (C₂H₅OH, hereafter ET), methyl formate (CH₃OCHO, hereafter MF), and dimethyl ether (CH₃OCH₃, hereafter DME). We also compare the spatial distribution of these COMs with that of methyl cyanide (CH₃CN), a typical hot molecular tracer that has been studied in detail in Cesaroni et al. (2011).

We carried out the line identification of the COMs with *MADCUBAII*, which uses several molecular databases, including JPL³ and CDMS⁴. This software provides theoretical synthetic spectra of the different molecules under LTE conditions, taking into account the individual opacity of each line.

3.1. Single-dish data: GBT and IRAM 30 m

Our wide spectral coverage has allowed us to clearly identify multiple lines of all the different COMs. In Table 2 we summarize the molecules detected in each dataset. All the COMs have been detected in the different datasets, with the exception of GA and DME in the GBT data, because there are no bright lines of these species in the frequency range covered by the observations.

We confirm for the first time the detection of the lowest energy conformer (aGg' conformer) of EG towards G31 (see details of the spectroscopy of EG in Appendix A). The single-dish data (GBT+IRAM 30 m) provide us with more than 30 unblended transitions of EG (see Fig. 1 and Table 3). These transitions cover an energy range of 5–114 K. We note that all the transitions predicted by the synthetic spectrum above 3 σ are present in the observed spectra.

For identification purposes, we have overplotted in Fig. 1 the simulated LTE spectra provided by *MADCUBAII*, assuming a source size from the SMA maps (see Sect. 4.1). The relative intensities of the lines from 7 mm to 1.3 mm are well fitted with an excitation temperature of ~50 K. We note, however, that the

² The Submillimeter Array is a joint project between the Smithsonian Astrophysical Observatory and the Academia Sinica Institute of Astronomy and Astrophysics, and is funded by the Smithsonian Institution and the Academia Sinica (Ho et al. 2004).

³ <http://spec.jpl.nasa.gov/>

⁴ <http://www.astro.uni-koeln.de/cdms>

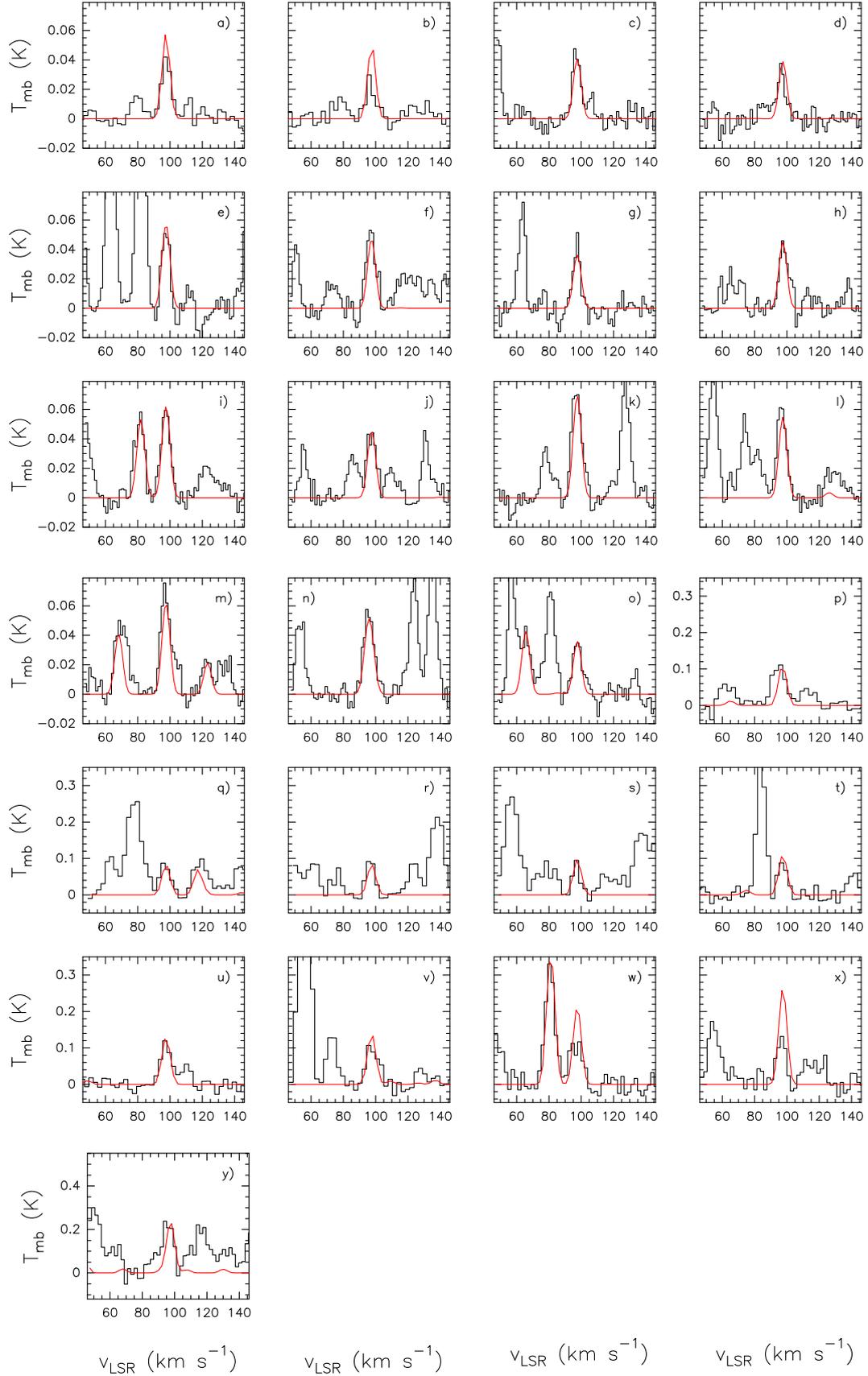


Fig. 1. Unblended transitions of EG from single-dish observations. The spectroscopic parameters (rest frequencies, quantum numbers, energies of the upper levels, and integrated intensities) from the LTE fit are listed in Table 3. For identification purposes, we have overlotted in red the LTE synthetic spectrum obtained with *MADCUBAIJ* assuming $T_{\text{ex}} = 50$ K (see text).

Table 3. Clean transitions (i.e., not blended with other molecular species) of EG identified in the single-dish spectra (GBT and IRAM 30 m) towards G31.

Frequency (GHz)	Transition J_{K_a, K_c}, v	E_{up} (K)	$\int T_{\text{mb}} \times \Delta v$ ^b (K × km s ⁻¹)	Panel Fig. 1
45.17972	$5_{2,3} (v=0) - 4_{2,2} (v=1)$	9	0.351	a)
45.54764	$4_{1,4} (v=1) - 3_{1,3} (v=0)$	5	0.305	b)
83.62492	$8_{1,8} (v=1) - 7_{1,7} (v=0)$	17	0.251	c)
83.85949	$9_{2,8} (v=0) - 8_{2,7} (v=1)$	24	0.240	d)
84.43952	$8_{0,8} (v=1) - 7_{0,7} (v=0)$	17	0.361	e)
85.21578 ^a	$9_{6,4} (v=0) - 8_{6,3} (v=1)$	40	0.135	f)
85.21582 ^a	$9_{6,3} (v=0) - 8_{6,2} (v=1)$	40	0.172	f)
85.52157	$9_{4,6} (v=0) - 8_{4,5} (v=1)$	30	0.221	g)
85.60355	$9_{4,5} (v=0) - 8_{4,4} (v=1)$	30	0.280	h)
86.65561	$9_{1,8} (v=0) - 8_{1,7} (v=1)$	23	0.393	i)
86.66012	$9_{3,6} (v=0) - 8_{3,5} (v=1)$	27	0.332	i)
87.56206	$8_{2,7} (v=1) - 7_{2,6} (v=0)$	20	0.285	j)
88.36696	$10_{1,10} (v=0) - 9_{1,9} (v=1)$	26	0.443	k)
88.74534	$10_{0,10} (v=0) - 9_{0,9} (v=1)$	26	0.342	l)
88.81207	$9_{2,7} (v=0) - 8_{2,6} (v=1)$	25	0.391	m)
88.87342 ^a	$8_{5,4} (v=1) - 7_{5,3} (v=0)$	30	0.171	n)
88.87421 ^a	$8_{5,3} (v=1) - 7_{5,2} (v=0)$	30	0.217	n)
88.98621	$8_{4,5} (v=1) - 7_{4,4} (v=0)$	26	0.221	o)
109.35771 ^a	$10_{6,5} (v=1) - 9_{6,4} (v=0)$	45	0.334	p)
109.35786 ^a	$10_{6,4} (v=1) - 9_{6,3} (v=0)$	45	0.422	p)
109.48919	$10_{5,5} (v=1) - 9_{5,4} (v=0)$	40	0.525	q)
109.51132	$10_{3,8} (v=1) - 9_{3,7} (v=0)$	32	0.536	r)
109.85770	$10_{4,6} (v=1) - 9_{4,5} (v=0)$	35	0.620	s)
111.44834	$10_{3,7} (v=1) - 9_{3,6} (v=0)$	32	0.716	t)
111.57944	$11_{1,11} (v=1) - 10_{1,10} (v=0)$	31	0.844	u)
112.34368	$12_{2,11} (v=0) - 11_{2,10} (v=1)$	40	0.909	v)
168.37712	$17_{6,12} (v=0) - 16_{6,11} (v=1)$	93	1.36	w)
168.38635	$17_{2,15} (v=0) - 16_{2,14} (v=1)$	79	2.35	w)
168.44321	$17_{6,11} (v=0) - 16_{6,10} (v=1)$	93	1.69	x)
171.41795	$16_{6,11} (v=1) - 15_{6,10} (v=0)$	85	1.48	y)

Notes. ^(a) The EG lines are autoblenched, i.e., blended with other transition of EG. ^(b) Integrated intensities derived from the LTE fit.

Table 4. Clean transitions (i.e., not blended with other molecular species) of EG identified in the interferometric SMA spectra towards G31.

Frequency (GHz)	Transition J_{K_a, K_c}, v	E_{up} (K)	$\int S \times \Delta v$ (Jy/beam × km s ⁻¹)
221.03880	$22_{6,17} (v=0) - 21_{6,16} (v=1)$	143	8.6
221.10032	$22_{5,18} (v=0) - 21_{5,17} (v=1)$	137	9.0
230.57714	$24_{3,22} (v=0) - 23_{3,21} (v=1)$	150	10.0
230.83032	$24_{2,22} (v=0) - 23_{2,21} (v=1)$	150	7.3
231.12740	$23_{7,16} (v=0) - 22_{7,15} (v=1)$	160	7.7

LTE fit of multiwavelength transitions of EG underestimates the true temperature. The reason is explained in detail in Sect. 4.2.

We also detected multiple lines of the other COMs. In Appendix B we present tables with the unblended transitions detected for each species. In Figs. B.1 and B.2 we show selected lines at 3 mm and 7 mm detected with the IRAM 30 m and GBT telescopes, respectively, along with the LTE fits obtained with MADCUBAIJ (Table C.1).

3.2. Interferometric data: SMA

For the identification of lines we used the SMA datacube obtained with the compact interferometer configuration. Figure 4

shows the full SMA spectra (containing LSB and USB sidebands) towards the continuum peak. For identification purposes, we have overplotted the LTE synthetic spectra obtained with MADCUBAIJ of the different COMs, and have indicated with colored dots the lines least affected by blending. Several EG unblended lines are clearly identified (Fig. 4 and Table 4). The energy range covered by these transitions is 114–160 K. Additionally, the interferometric data allow us to map the spatial distribution of the EG emission. To study the spatial distribution of the emission we used the datacubes obtained by combining the data of both compact and very extended configurations, which provide a better spatial resolution (0.90'' × 0.75''). Figure 3 confirms that the spatial distribution of different unblended EG lines

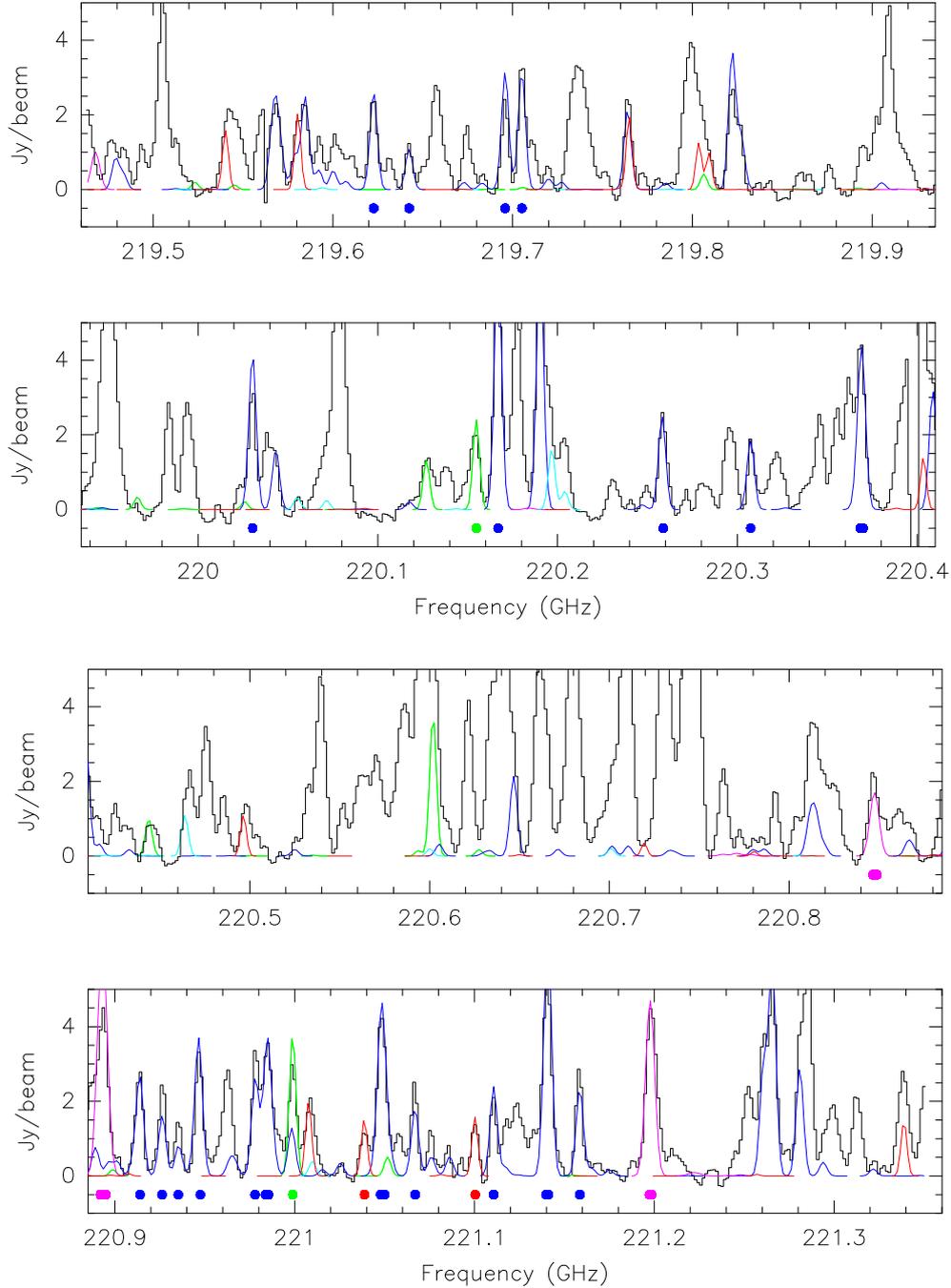


Fig. 2. Spectrum towards the peak of the continuum of the SMA compact data. We have overplotted the LTE synthetic spectrum of the different COMs obtained with *MADCUBAIJ*: EF (red), GA (light blue), MF (dark blue), DME (magenta), and ET (green). The colored circles indicate the unblended lines given in Table 4 and Tables B.2, B.4, B.6, and B.8. The physical parameters used are shown in Table C.2.

is very similar. This spatial coherence supports the identification of EG.

Multiple lines of the other COMs have also been identified in the SMA data. Figure 4 shows the LTE spectra of the different COMs. The cleanest transitions of each molecule are listed in Appendix B (Tables B.2, B.4, B.6, and B.8).

4. Derivation of physical properties

4.1. Spatial distribution of EG and other COMs

The combined compact and very extended SMA observations have a nearly circular beam and a good uv coverage, thus

allowing us to attain subarcsec resolution in all directions ($0.90'' \times 0.75''$; PA = 53°). Figure 4 shows the maps of the emission of different molecules obtained by integrating the emission under unblended lines. We defined the deconvolved beam size of the emitting regions as $\theta_s = \sqrt{\theta_{50}^2 - \theta_{\text{beam}}^2}$, where θ_{beam} is the half-power beam width of the interferometer and $\theta_{50} = 2\sqrt{A/\pi}$ is the diameter of the circle whose area A equals that enclosed inside the contour level corresponding to 50% of the line peak.

The EG emission has a diameter of $0.7''$, and peaks towards the position of the continuum (Fig. 4), similarly to that of GA and the vibrationally excited $\text{CH}_3\text{CN } \nu_8 = 1$. The EG emission is clearly smaller than the emission of other COMs like the ground

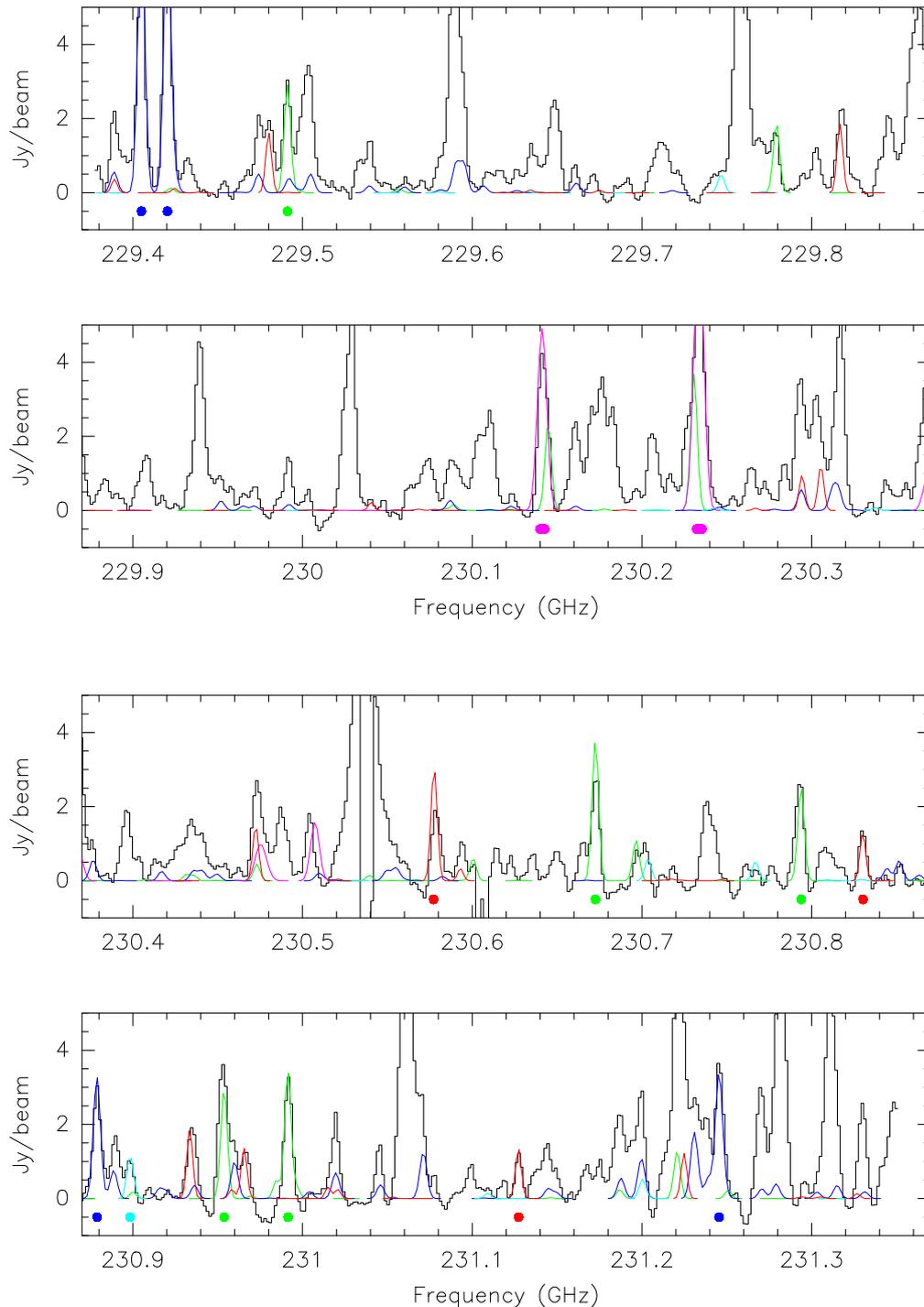


Fig. 2. continued.

vibrational level of CH_3CN , ET and MF (Fig. 4 and Table 5). This is also observed in the Orion Hot Core, where EG is more compact than MF and ET (Brouillet et al. 2015).

The EG morphology differs from the “8-shaped” structure observed in ET, MF, or the ground state of CH_3CN . Cesaroni et al. (2011) suggested that the “8-shaped” morphology is due to line opacity, which is higher at the center of the core and hence produces the observed dip. Since more complex molecules such as EG have lower abundances, they are expected to suffer less from line opacity effects and consequently allow us to trace gas located closer to the central protostar(s). This makes EG an

excellent tracer of the physical properties and kinematics of the inner regions of hot cores, which can help us to understand massive star formation itself.

4.2. Importance of dust absorption in G31

In Sect. 3 we fit the 45–219 GHz single-dish spectra of EG using an LTE analysis with a temperature of 50 K using *MADCUBAII*. This temperature is significantly lower than that usually measured in hot cores, and in particular lower than

Table 5. Diameter of the emitting regions (see Sect. 4.1).

	EG	GA	ET	MF	DME	CH ₃ CN, $K = 0, 1, 2$	CH ₃ CN, $v_8 = 1$
θ_{50} (")	1.1	1.2	1.6	1.9	1.6	2.1	1.3
θ_s (")	0.7	0.9	1.4	1.7	1.3	1.9	1.0

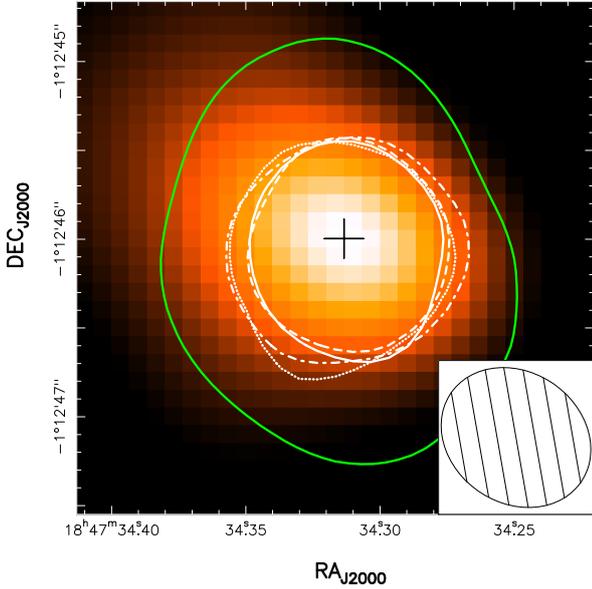


Fig. 3. Spatial distribution of EG transitions at 221.03880 GHz (white solid), 221.10032 GHz (white dashed), 230.57714 GHz (white dotted), and 230.830 GHz (white dot-dashed), overplotted on the 1.3 mm continuum observed with the SMA (color scale). The black plus sign indicates the position of the peak of the 1.3 mm continuum. For comparison, we have also added the spatial distribution of the integrated emission of $K = 0, 1, 2$ CH₃CN line (solid green line). All contours show the isocontour with 50% of the peak line intensity. The beam of the interferometer is shown in the lower right corner.

the 164 K measured towards G31 by Beltrán et al. (2005) using CH₃¹³CN and CH₃CN $v_8 = 1$. This discrepancy is striking, given that EG is tracing the innermost region of the core, as we have seen in the previous section, where a high temperature is expected. Since the low excitation transitions of EG ($E_{\text{up}} = 5\text{--}45$ K) have low frequencies (<113 GHz), and the higher excitation transitions ($E_{\text{up}} = 79\text{--}114$ K) have higher frequencies (>168 GHz), one may wonder if such a low temperature is due to the high-frequency (and high-energy) lines being somehow weakened with respect to the low-frequency (and low-energy) ones. Dust absorption provides us with a frequency-dependent mechanism of this type. In fact, dust opacity increases with frequency, $\tau_{\nu} \sim \nu^{\beta}$, and hence may absorb at high frequency more effectively than at lower frequencies. This would result in an artificially low temperature when fitting with an LTE model.

To check whether this effect is at work in the G31 core, we studied in detail the emission of MF, which unlike EG has the advantage of multiple transitions that cover wide ranges of E_{up} at similar frequencies (and hence similarly affected by dust absorption). In the upper panel of Fig. 5 we present the rotational diagram of MF (which assumes optically thin emission; see, e.g., Goldsmith & Langer 1999) obtained from unblended MF transitions detected with single-dish data. The values shown on the y -axis have been corrected for beam dilution, using the size derived for MF of 1.7" from the SMA maps (Table 5). We obtained a rotational temperature of ~ 75 K when considering all the MF

transitions, similar to the low temperature found using EG, and again lower than that expected in this hot core.

We repeated the analysis distinguishing between three different frequency ranges. The temperatures obtained when fitting transitions separately at similar frequencies are higher (150–200 K, with an average value of ~ 170 K), in much better agreement with the value of 164 K derived by Beltrán et al. (2005). The upper panel of Fig. 5 confirms that transitions at higher frequencies are systematically shifted to lower values on the y -axis of the rotational diagram. We interpret this result as an indication of dust opacity in G31: the line photons at high frequencies are more absorbed by dust grains than those at lower frequencies, which decreases the corresponding line intensities.

We note that the shift in the y -axis of the rotational diagram between transitions in different frequency ranges cannot be due to different beam-dilutions at different frequencies. Even in the extreme case of extended emission filling the beam at all frequencies, the shift is still present (middle panel of Fig. 5). However, this is not the case. We know that both EG and MF emissions are not that extended because our high angular resolution interferometric maps show compact sizes of 0.7" and 1.7", respectively.

We suggest that the dust opacity is affecting the observed molecular line intensities. This effect is expected to be particularly important towards the innermost region of the core, where EG is detected. To support this interpretation, we studied whether the physical properties of G31 can indeed produce the dust absorption needed to explain the observed line intensities at different frequencies.

We derived the excitation temperature of MF considering only 3 mm transitions, which are the least affected by dust absorption. We selected transitions that cover a wide range of energy levels (20–200 K), which allows us to constrain the excitation temperature, and with similar frequencies (in the range 83–88 GHz) so that the effect of dust absorption is nearly the same. We obtained a value of $T_{\text{ex}} = 163$ K, very similar to that derived by Beltrán et al. (2005).

It is possible to derive that the effect of dust opacity (τ_{ν}) decreases the line intensities by a factor $e^{-\tau_{\nu}}$. Assuming a dust opacity index ($\tau_{\nu} \sim \nu^{\beta}$) typical of massive star-forming cores of $\beta = 1.5$ (see, e.g., Palau et al. 2014), we then calculated the value of the opacity that would provide $T_{\text{ex}} = 163$ K when fitting *all* the transitions at different frequencies, which is $\tau(220 \text{ GHz}) = 2.6$. The lower panel of Fig. 5 shows the rotational diagram of MF after correcting the line intensities for dust opacity.

It is possible to derive the hydrogen column density of the core from the obtained value of the opacity using the expression:

$$N_{\text{H}_2} = \frac{\tau_{\nu}}{\mu m_{\text{H}} \kappa_{\nu}}, \quad (1)$$

where m_{H} is the hydrogen mass, $\mu = 2.33$ is the mean molecular mass per hydrogen atom, and κ_{ν} is the absorption coefficient per unit density. For $\tau(220 \text{ GHz}) = 2.6$, assuming $\kappa_{220 \text{ GHz}} = 0.005 \text{ cm}^{-2} \text{ g}^{-1}$ and a gas-to-dust ratio of 100, the derived hydrogen column density is $\sim 1.2 \times 10^{26} \text{ cm}^{-2}$. Using the

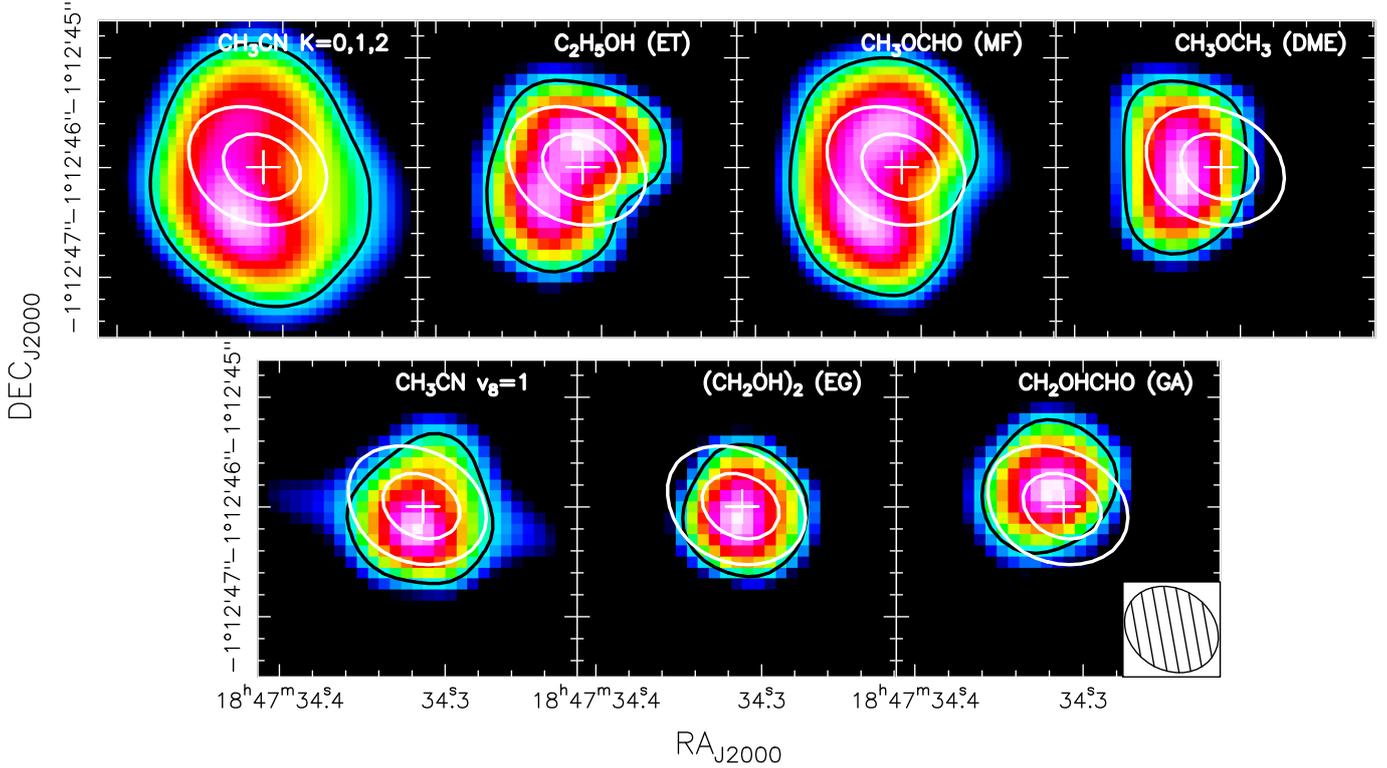


Fig. 4. Integrated emission maps of different COMs towards G31. All the maps have been obtained with the combination of the very extended and compact SMA data. The color scale in each panel spans from 30% to 100% of the line intensity peak. The frequencies (and upper energies) of the lines are: 220.747 GHz (69 K), 220.743 GHz (76 K), and 220.730 GHz (97 K) for CH_3CN ; 230.991 GHz (85 K) for ET; 229.405 GHz (111 K) and 229.420 (111 K) for MF; 230.234 (148 K) for DME; 221.228 GHz (698 K) for $\text{CH}_3\text{CN } v_8 = 1$; 221.100 (137 K) for EG; and 230.898 (131 K) for GA. The black solid line indicates the isocontour with 50% of the peak line intensity. The two white contours correspond to 50% and 80% of the peak value of the 1.3 mm continuum. The white plus sign indicates the positions of the peak of the 1.3 mm continuum. The beam of the interferometer is shown in the lower right corner of the lower right panel.

deconvolved angular diameter of the continuum core of $\sim 0.76''$ (Cesaroni et al. 2011), this implies a total mass of $\sim 2300 M_\odot$, which is consistent with the value of $1700 M_\odot$ obtained from the flux continuum emission by Cesaroni et al. (2011). The continuum flux density can be derived with the expression:

$$F_\nu = (1 - e^{-\tau}) \Omega_s B_\nu(T_d), \quad (2)$$

where Ω_s is the beam solid angle covered by the source, and $B(T_d)$ is the Planck function. Using $\Omega_s = \pi\theta_c^2/4\ln 2$, where θ_c is the deconvolved angular diameter of the continuum core ($\sim 0.76''$, from Cesaroni et al. 2011) and assuming $\tau(220 \text{ GHz}) = 2.6$ and $T_d = 163 \text{ K}$, we obtain a continuum flux density of 3.8 Jy, consistent with the observed value (4.6 Jy, Cesaroni et al. 2011). Therefore, this confirms our interpretation that dust absorption is affecting the line intensities in G31. Obviously, the effect of dust absorption is expected to be important only if the level of clumpiness of the dense core is very low, i.e., if the dusty core is not fragmented into smaller condensations. Interestingly, recent high angular resolution ($\sim 0.2''$) ALMA observations confirm that there is no hint of fragmentation in the continuum maps of G31 (Cesaroni, priv. comm.).

In summary, a LTE analysis of multifrequency transitions without considering the dust absorption underestimates the excitation temperatures, explaining the intriguingly low temperature of 50 K derived for EG when fitting multiwavelength transitions. In the next section we explain how to derive reliable values of the excitation temperatures and hence of molecular column densities.

4.3. Molecular column densities

Keeping the effect of dust absorption in mind, to obtain the physical parameters of the different COMs, we restricted our analysis to the IRAM 30 m 3 mm transitions, the least affected by the dust opacity. We used *MADCUBAIJ* to fit the observed 3 mm spectra with the LTE model⁵. *MADCUBAIJ* considers five different parameters to create the simulated LTE spectra: column density of the molecule (N), excitation temperature (T_{ex}), linewidth (Δv), velocity (v), and source size (θ_s). The procedure used to derive the main physical parameters is the following:

i) we fix the velocity of the emission and the linewidth to the values obtained from Gaussian fits of unblended lines; ii) we fix the gas temperature to the value estimated using the 3 mm transition of MF (163 K), which is a good temperature tracer (Favre et al. 2011); iii) to obtain the source-average column density N_s , we fix the size of the molecular emission to the value obtained from the integrated maps of clean lines detected with the SMA, and apply the beam dilution factor accordingly; instead we do not apply the beam dilution factor to derive the beam-averaged column density N_b ; iv) leaving N as free parameter, we use the *AUTOFIT* tool of *MADCUBAIJ* to find the value that fits the unblended molecular transitions of each COM better.

The results for the different molecular species are summarized in Table 6. To account for the dust absorption discussed above, the values for the column densities have been corrected

⁵ The details of the LTE synthetic spectra used to identify the different COMs in the SMA and GBT data (Figs. 2 and B.2, respectively) are presented in Appendix C.

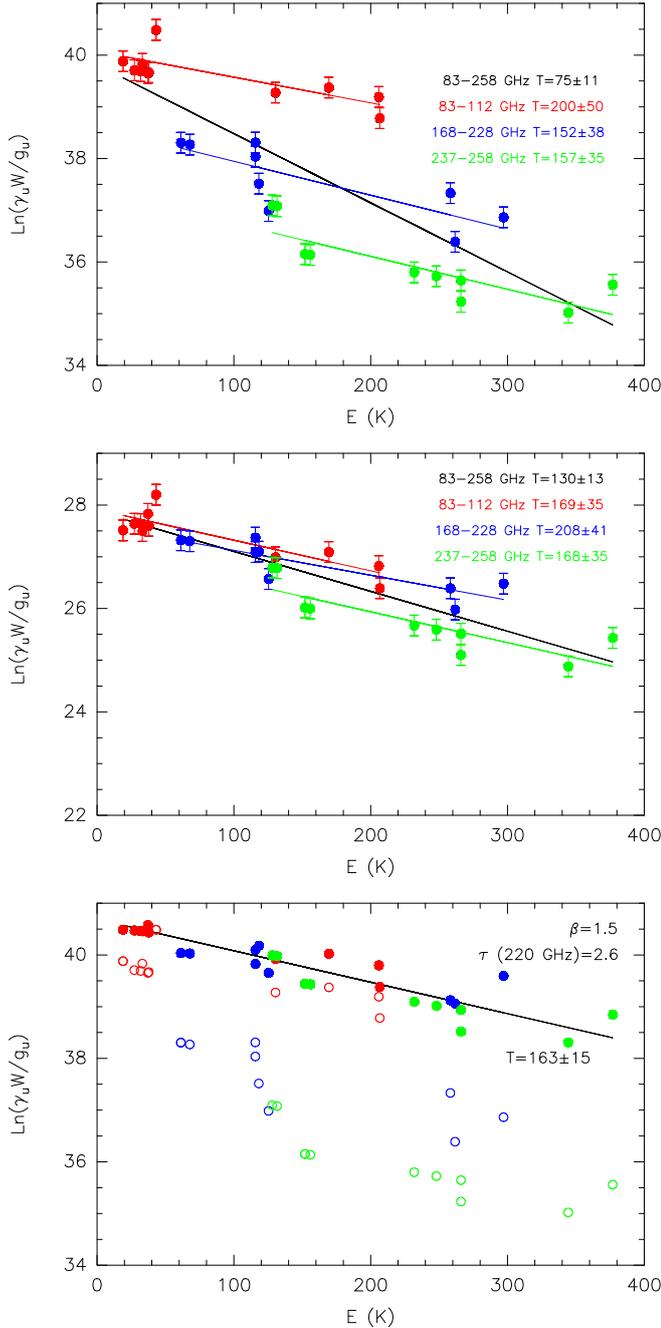


Fig. 5. *Upper panel:* rotational diagram of MF considering the beam dilution factor. The different colors indicate transitions in the different frequency ranges indicated by the labels. The black line is the best fit considering all transitions, while the colored lines are the best fits considering only the transitions in the same frequency range. *Middle panel:* same as upper panel, without considering beam dilution, i.e., assuming that the molecular emission completely fills the telescope beam. *Lower panel:* same as upper panel after correcting the line intensities for the dust opacity (solid circles) and without correction (open circles).

by a factor $e^{\tau(3 \text{ mm})} \sim 2$, where $\tau(3 \text{ mm})$ has been derived assuming the derived value of $\tau(220 \text{ GHz}) = 2.6$ and $\beta = 1.5$.

The main source of uncertainty in the estimated column densities arises from the assumption of the LTE approximation using a single temperature. However, even in the extreme case that the temperature varies by a factor of 2, the variation in the derived column densities would be less than a factor of 2. We therefore

Table 6. Physical parameters obtained from the 3 mm IRAM 30 m data for the different molecules.

	EG	ET	GA	MF	DME
$\log N_b \text{ (cm}^{-2}\text{)}$	15.0	15.9	14.7	16.2	15.9
$\log N_s \text{ (cm}^{-2}\text{)}$	18.5	18.5	17.5	18.7	19.0
$X_s \text{ (10}^{-8}\text{)}^a$	2.6	0.84	0.26	4.2	8.4
$v_{\text{LSR}} \text{ (km s}^{-1}\text{)}$	97.6	97.6	96.1	97.6	97.0
$\Delta v \text{ (km s}^{-1}\text{)}$	5.6	5.4	5.3	4.7	4.4

Notes. ^(a) Assuming $N_{\text{H}_2} = 1.2 \times 10^{26} \text{ cm}^{-2}$ (see Sect. 4.2).

consider that the uncertainty in the estimated column densities is less than a factor of 2.

All the molecules have velocities very close to the source systemic velocity, $\sim 97 \text{ km s}^{-1}$, and linewidths (FWHM) $\sim 5 \text{ km s}^{-1}$, typical of hot molecular cores. The derived source-averaged molecular abundances with respect to molecular hydrogen are in the range $\sim (0.3\text{--}8) \times 10^{-8}$.

5. Comparison with other interstellar sources

5.1. Comparison with other star-forming regions

We compared the relative molecular abundances obtained in G31 with those already reported in other star-forming regions: the massive star-forming regions in Orion KL (Brouillet et al. 2015), W51e2, and G34.3+0.2 (Lykke et al. 2015); the intermediate-mass star-forming region NGC 7129 FIRS2 (Fuente et al. 2014); and the low-mass protostars IRAS 16293-2422 (Jørgensen et al. 2012), NGC 1333 IRAS 2A (Maury et al. 2014; Coutens et al. 2015; Taquet et al. 2015), and NGC 133 IRAS 4A (Taquet et al. 2015). The different relative molecular abundances are given in Table 7. The uncertainty of the abundance ratios can be estimated by considering propagation of the uncertainty of the column densities of each molecule. We expect then that the uncertainty of relative molecular ratios is less than a factor of 3. In Fig. 6 we show the behavior of the different abundance ratios with respect to the luminosity of the sources.

Table 7 shows that EG is always more abundant than GA in all star-forming regions. This has also been noted by Requena-Torres et al. (2008) in the Galactic Center, where saturated species, such as EG, are always more abundant than unsaturated species, such as GA (see Table 7); and in star-forming regions (e.g. Ikeda et al. 2001; Hollis et al. 2002; Bisschop et al. 2007). Therefore, there is evidence that chemistry favors saturated versus unsaturated species. Requena-Torres et al. (2008) suggested that the double bond C=O (present in the unsaturated species, e.g. GA) can be easily broken.

In star-forming regions, the [EG/GA] ratio changes with luminosity, spanning from 1 to >15 . The lower limits of the higher luminosity sources (hot cores) suggest that the [EG/GA] ratio increases with luminosity.

The [DME/MF] and [EG/ET] ratios are nearly constant with luminosity, with values of the order of ~ 1 and ~ 0.1 , respectively. The [GA/MF] and [ET/DME] ratios clearly decrease with increasing luminosity. In Sect. 6 we discuss how these results can help us to understand the formation of EG and other COMs.

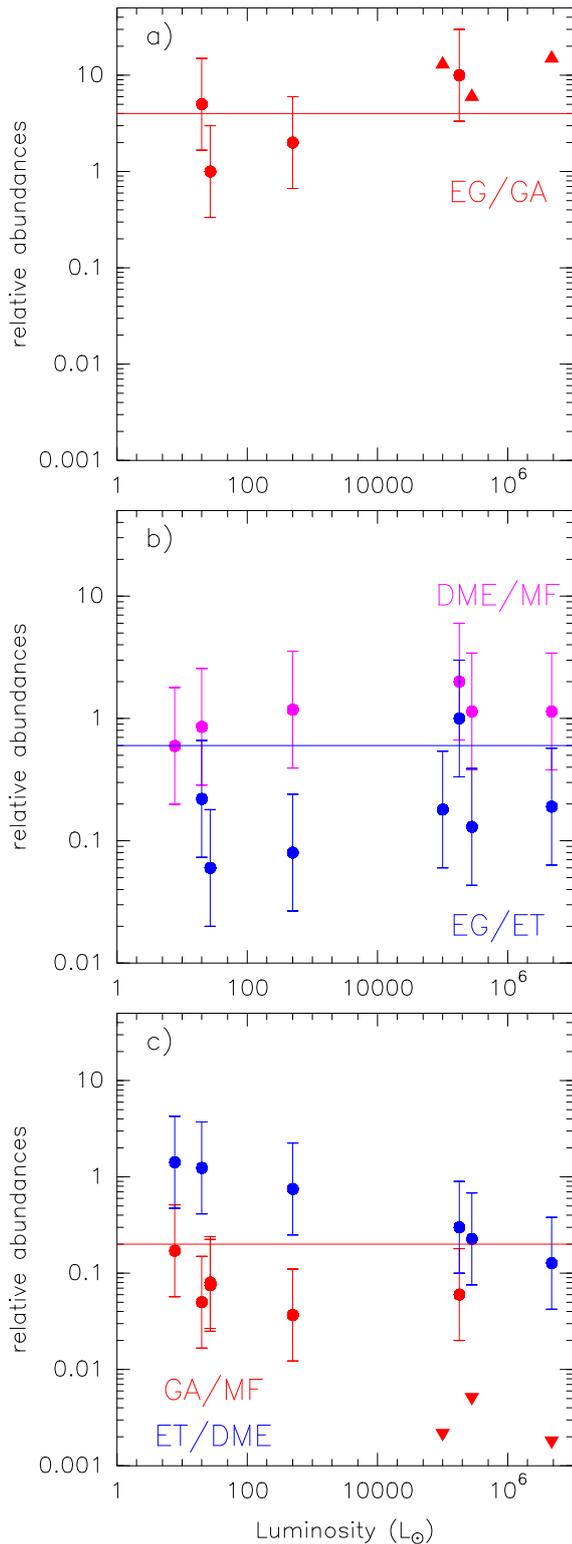


Fig. 6. Relative abundances of pairs of COMs with respect to the luminosity of the star-forming regions. The conservative uncertainty of the relative abundances correspond to a factor of 3 (see text). The triangles in the *upper*(*lower*) panels are lower(*upper*) limits. The solid horizontal lines correspond to the relative abundance value found in the comet Lovejoy, from [Biver et al. \(2015\)](#).

5.2. Comparison with quiescent clouds and comets

The formation of stars and planets undergoes different evolutionary stages: i) cold and dense prestellar cores are able to

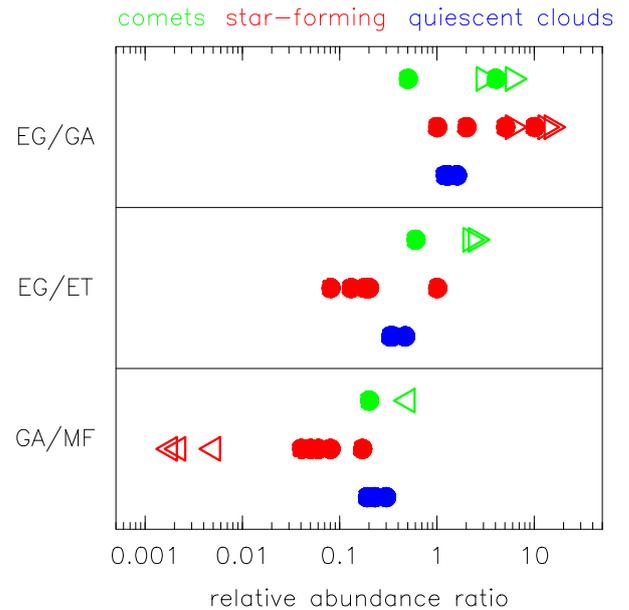


Fig. 7. Comparison of different ratios of molecular abundances of COMs between comets, star-forming regions, and quiescent clouds. The color circles indicate measured ratios, while right and left-oriented triangles denote lower and upper limits.

collapse, triggering star formation; ii) the gravitational energy is converted into thermal energy and radiation, and the forming star warms up the dense neighboring envelope; and iii) the envelope dissipates and a circumstellar disk remains, where eventually planets and other bodies such as comets are formed. We investigate here whether there is any hint of a chemical trend between different evolutionary phases of the ISM: quiescent clouds, star-forming regions and comets. To this end we have also included in Table 7 the relative abundance ratios towards quiescent clouds located in the Galactic Center, G-002, G-001, G+0.693 ([Requena-Torres et al. 2008](#)), and the comets Hale-Bopp ([Crovisier et al. 2004](#)), Lemmon ([Biver et al. 2014](#)), Lovejoy ([Biver et al. 2015](#)), and 67P/Churyumov-Gerasimenko ([Goesmann et al. 2015](#)). In Fig. 7 we compare different molecular abundance ratios between these three groups of sources (i.e., comets, star-forming regions, and quiescent clouds).

We find that for all kind of sources $[EG/GA] > [EG/ET] > [GA/MF]$. This suggests that the relative abundances of different COMs, despite the observed large dispersions, share a general trend in interstellar sources at different evolutionary stages. This might mean that the chemistry at different stages of star formation is not significantly different, and there is a chemical thread across the entire star formation process. However, the low number of sources with positive detections of these COMs is still low, and prevents a statistically significant conclusion.

6. Discussion: formation of ethylene glycol and other COMs in star-forming regions

We detected 30 unblended transitions of EG with single-dish observations and 5 unblended transitions with interferometric observations. We note that the best detection so far of EG (in SgrB2 N, the prototypical interstellar source for the search of large COMs) relies on the detection of only ~ 13 unblended lines ([Belloche et al. 2013](#)), despite the much wider spectral coverage of these observations. This clearly shows that although the extreme conditions of the Galactic Center create a unique

Table 7. Relative abundances of COMs in comets, star-forming regions and Galactic Center quiescent clouds.

Source	[EG/GA]	[EG/ET]	[GA/MF]	[ET/DME]	[DME/MF]	Luminosity (L_{\odot})	Reference
Comets							
Hale-Bopp	>6	>2.5	<0.5	–	–	–	1
Lemmon	>3	>2.2	^a	–	–	–	6
Lovejoy	4	0.6	0.2	–	–	–	14
67P	0.5	–	–	–	–	–	13
Low-mass star-forming regions (hot corinos)							
IRAS 16293-2422	1	–	0.08	–	–	2.7	5
NGC 1333 IRAS 2A	5	0.2	0.05	1.3	0.9	20	7, 9, 10
NGC 1333 IRAS 4A	–	–	0.17	1.4	0.6	7.7	10
Intermediate-mass star-forming regions							
NGC 7129 FIRS2	2	0.08	0.04	0.75	1.2	5×10^2	8
High-mass star forming regions (hot cores)							
G31.41+0.31	10	1.0	0.06	0.3	2.0	1.8×10^5	3, 15
W51 e2	>15	0.19	<0.0018	0.13	1.1	4.7×10^6	12
G34.3+0.2	>6	0.13	<0.005	0.23	1.1	2.8×10^5	12
Orion KL	>13	0.18	<0.0022	–	–	1×10^5	4, 11
Galactic Center quiescent clouds							
G-002	1.3	0.33	0.3	–	–	–	2
G-001	1.6	0.47	0.23	–	–	–	2
G+0.693	1.2	0.35	0.19	–	–	–	2

Notes. ^(a) Only upper limits were reported for both species.

References. (1) Crovisier et al. (2004); (2) Requena-Torres et al. (2008); (3) Beltrán et al. (2009); (4) Favre et al. (2011); (5) Jørgensen et al. (2012); (6) Biver et al. (2014); (7) Maury et al. (2014); (8) Fuente et al. (2014); (9) Coutens et al. (2015); (10) Taquet et al. (2015); (11) Brouillet et al. (2015); (12) Lykke et al. (2015); (13) Goesmann et al. (2015); (14) Biver et al. (2015); (15) this work.

chemically rich environment, paradoxically the presence of so many molecules produces blending problems that make it difficult to identify molecules with faint emission. Moreover, in SgrB2 N multiple velocity components and strong absorption also complicate a correct identification of the lines. Therefore, sources with relatively simple structure and kinematics, but also chemically rich, such as G31, are a good alternative to carry out astrochemical studies. Furthermore, our work proves that single-dish observations are suitable for the detection of large COMs in hot molecular cores, despite the beam dilution. Obviously, interferometric observations at high resolution are also needed to measure the size of the emitting regions and hence obtain proper values of the molecular abundances.

The formation of COMs is still an open debate in astrophysics. Theoretical chemical models based only on gas-phase reactions have failed by orders of magnitude to reproduce the observed abundances of COMs (e.g., Geppert et al. 2006). Moreover, in the particular case of EG and GA there are no known gas-phase reactions to form GA or EG (see the chemical databases KIDA⁶ and UMIST⁷; and also Woods et al. 2012). For this reason, chemistry on the surface of dust grains has been invoked to form COMs (e.g., Garrod et al. 2008), which would be subsequently desorbed as a result of heating from the central protostar or from shocks, enriching the molecular environment. However, Balucani et al. (2015) have recently pointed out that some gas-phase reactions are feasible, and that gas-phase chemistry should also be considered as an alternative to form COMs.

Based on the derived molecular abundances in G31 and other star-forming regions, we discuss in this section how observations can give us clues about the formation of COMs.

We have found that the [EG/GA] ratio seems to increase with the luminosity of the source, from the low values of ~ 1 found towards the low-mass protostar IRAS 16293-2422 (Jørgensen et al. 2012; and Jørgensen et al., in prep.) to the higher value of ~ 10 towards the hot core G31 and the lower limits of the hot cores in Orion, W51e2, and G34.3+0.2, which are as high as >15 . We have also found that while the [DME/MF] and [EG/ET] are nearly constant with luminosity, the ratios between the pair of isomers [GA/MF] and [ET/DME] decrease with luminosity. In the following sections we propose different mechanisms that are able to explain these observational trends, and that give us hints about the formation of COMs in star-forming regions: i) different chemical formation routes; ii) different warm-up timescales; iii) different initial gas compositions of the dust grains; iv) different hydrogen densities of the sources; v) different formation/destruction efficiency in the gas phase.

6.1. Chemical link

Table 8 shows different chemical routes proposed in the literature to form the COMs studied in this work: EG, GA, MF, DME, and ET. These chemical routes rely on theoretical chemical models (Bennett & Kaiser 2007; Garrod et al. 2008; Woods et al. 2013; Balucani et al. 2015) and laboratory experiments (Fedoseev et al. 2015; Butscher et al. 2015).

If two molecular species are chemically linked, i.e., if they have a common precursor and/or one is formed from the other, their relative abundance ratio should be nearly constant regardless of the luminosity of the source. This occurs, for instance, in the [DME/MF] ratio (middle panel of Fig. 6), which is ~ 1 for all star-forming regions. This chemical link suggested by the observations is indeed predicted by theoretical models. Garrod et al. (2008) proposed that both MF and DME form from

⁶ <http://kida.obs.u-bordeaux1.fr/>

⁷ <http://udfa.ajmarkwick.net/>

Table 8. Chemical routes proposed in the literature for the formation of various COMs studied in this work.

COM		Chemical pathway	Reference
(CH ₂ OH) ₂ (EG)	[1]	CH ₂ OHCHO + 2H → (CH ₂ OH) ₂	4
	[2]	HCO + 2H → CH ₂ OH; CH ₂ OH + CH ₂ OH → (CH ₂ OH) ₂	1, 2, 6
CH ₂ OHCHO (GA)	[3]	2HCO → CO + H ₂ CO → HOCCOH; HOCCOH + H → CH ₂ OCHO; CH ₂ OCHO + H → CH ₂ OHCHO	3
	[4]	HCO + CH ₂ OH → CH ₂ OHCHO	1, 2, 6
CH ₃ OCHO (MF)	[5]	CH ₃ O + HCO → CH ₃ OCHO	2
	[6]	CH ₃ OCH ₂ + O → CH ₃ OCHO + H (gas phase)	5
CH ₃ OCH ₃ (DME)	[7]	CH ₃ O + CH ₃ → CH ₃ OCH ₃	2
CH ₃ CH ₂ OH (ET)	[8]	CH ₂ OH + CH ₃ → CH ₃ CH ₂ OH	2

References. (1) Bennett & Kaiser (2007); (2) Garrod et al. (2008); (3) Woods et al. 2013; (4) Fedoseev et al. (2015); (5) Balucani et al. (2015); (6) Butscher et al. (2015).

the common precursor methoxy (CH₃O) by surface chemistry on dust grains (route 5 and 7 in Table 8, respectively). In addition, Balucani et al. (2015) have proposed a new gas-phase route that can efficiently contribute to the formation of MF directly from DME (route 6 in Table 8). In any case, it seems that MF and DME are chemically linked, in agreement with the observed constant abundance ratio.

Likewise, if EG were formed by sequential hydrogenation of GA as suggested by the route 1, a relatively constant [EG/GA] ratio would be expected. Instead, the values of the [EG/GA] ratio differ between sources by more than an order of magnitude, from 1 to >15 (upper panel of Fig. 6). This suggests that EG and GA are not directly linked.

Therefore, it seems more plausible that EG is formed by the combination of two CH₂OH radicals (route 2 in Table 8). Although Walsh et al. (2014) have raised some doubts about the efficiency of this process owing to the low mobility produced by the OH-bonds, the recent experimental work by Butscher et al. (2015) has shown that this mechanism seems to be a very efficient pathway to form EG.

The large variation of the [EG/GA] ratio suggests not only that GA is not the direct precursor of EG, but also that they do not share a common precursor. Since we favor the formation of EG by the combination of two CH₂OH radicals (route 2), this would imply that CH₂OH is not a precursor of GA, and route 4 can be ruled out. Then, the chemical pathway based on the dimerization of HCO (route 3) proposed by Woods et al. (2013) appears to be the most plausible scenario. This route has also been recently supported by the laboratory experiments of Fedoseev et al. (2015).

The nearly constant behavior of the [EG/ET] ratio with luminosity (middle panel of Fig. 6) suggests that these two species might be chemically linked. Indeed, this agrees with the chemical model proposed by Garrod et al. (2008), which produces EG and ET from the common precursor CH₂OH (routes 2 and 8, respectively).

In conclusion, the observed abundance ratios can all be justified by routes 2 (for EG), route 3 (for GA), routes 5 and/or 6 (for MF), route 7 (for DME), and route 8 (for ET), while routes 1 and 4 are excluded.

6.2. Timescales

The different warm-up timescales associated with objects of different luminosities (hot cores and hot corinos) are expected to affect the chemistry. We discuss here whether this effect is able to explain the behavior of the relative molecular abundances with luminosity observed in Fig. 6. According to Garrod et al. (2008), the [EG/GA] and [ET/DME] ratios are higher at longer warm-up timescales, while the ratio [GA/MF] decreases with the timescale. Assuming that the warm phase in hot cores is longer than in hot corinos, as indicated by the chemical models by Awad et al. (2014), then the heating timescale would be proportional to the luminosity. Hence, the timescale effect could contribute to explain the observed trend of [ET/DME], but not those of [EG/GA] and [GA/MF].

6.3. Chemical composition of dust grains

A possible explanation for the different [EG/GA] ratios is different chemical compositions of the dust grains in different star-forming regions. The experiments of Öberg et al. (2009) indicate that different initial compositions of the ices could produce very different values of the [EG/GA] ratio: pure CH₃OH ices produce a ratio >10, while ices with a composition CH₃OH:CO 1:10 produce a ratio <0.25. If this were the explanation for the different [EG/GA] ratios, it would imply that the initial composition of grains is different for low- and high-mass star-forming regions.

6.4. Atomic hydrogen density

The experiments and chemical modeling by Fedoseev et al. (2015) show that different initial atomic H densities provide different [EG/GA] ratios. Coutens et al. (2015) argue that the density could explain the difference observed towards IRAS 16293-2422 and NGC 1333 IRAS 2A, with [EG/GA] ratios of 1 and 5, respectively. They suggest that the lower H₂ density of IRAS 2A would lead to a higher atomic H density and then to a higher [EG/GA] ratio, as observed. However, this scenario does not apply to G31, which is denser and exhibits an even higher [EG/GA] ratio. Therefore, we believe that atomic hydrogen density cannot explain the observed behavior of the [EG/GA] ratio.

6.5. Formation and destruction on gas phase

So far we have assumed in our discussion that COMs are mainly formed by surface chemistry on dust grains and subsequently desorbed into the gas phase. However, the situation can be more complex if we also consider gas phase formation routes. Recently, [Balucani et al. \(2015\)](#) have suggested that some gas reactions may be important for the formation of MF in cold environments. If this were also the case for GA and EG, for which no efficient gas phase reactions are available to date (see [Woods et al. 2012](#)), the interpretation of the observational results should be revisited. More theoretical work is needed to make this hypothesis plausible.

Finally, another possible cause for the variation of the [EG/GA] ratio could be different destruction mechanisms in the gas phase. [Hudson et al. \(2005\)](#) show that GA is more sensitive to irradiation than EG, which may explain its lower abundance in luminous sources. Higher levels of irradiation by the higher luminosity sources would produce an increase in the [EG/GA] ratio with luminosity, as observed. However, the destruction pathways of these COMs have not yet been explored by the theoretical models and laboratory experiments.

7. Summary and conclusions

We have reported for the first time the detection of the aGg' conformer of the ten-atoms organic molecule ethylene glycol, (CH₂OH)₂, towards the hot molecular core G31.41+0.31. This and previous detections of COMs (e.g., CH₂OHCHO) confirm that G31.41+0.31 is an excellent laboratory for studying the astrochemistry related with the building blocks of primordial life. We have detected multiple unblended transitions of (CH₂OH)₂ with single-dish and interferometric observations. Our SMA interferometric maps show that the (CH₂OH)₂ emission is very compact, with a diameter of 0.7'', peaking towards the maximum of the continuum. This morphology is similar to that observed in vibrationally excited CH₃CN or CH₂OHCHO, and differs from that observed in other more abundant COMs such as the ground state of CH₃CN or CH₃OCHO, which exhibit a dip in the central region. We propose that this difference is likely due to line opacity. Lower abundance molecules such as (CH₂OH)₂ and CH₂OHCHO are expected to be more optically thin and allow us to trace the gas closest to the forming protostar(s).

Our detailed LTE analysis of (CH₂OH)₂ and CH₃OCHO towards G31.41+0.31 suggests that dust opacity plays an important role in the line intensities. We have found evidence that high-frequency transitions are more efficiently absorbed than low-frequency transitions owing to higher dust opacity. This effect should be taken into account to derive a correct value of the excitation temperature and column density of molecules in dense hot molecular cores. We have derived the molecular abundances of several COMs towards G31.41+0.31 (CH₂OHCHO, (CH₂OH)₂, C₂H₅OH, and CH₃OCH₃), finding abundances with respect to molecular hydrogen in the range $(0.3\text{--}8) \times 10^{-8}$.

Observations in different star-forming regions indicate that the ratio (CH₂OH)₂/CH₂OHCHO spans from 1 to >15. We have found evidence of an increase in the ratio with the luminosity of the source. We have discussed different possible explanations for this trend, concluding that the most likely scenario is that both species are formed through different chemical routes that are not directly linked. We cannot exclude, however, a contribution of different formation and destruction efficiencies in the gas phase of both species, but more laboratory work and theoretical chemical models are needed to better constrain this hypothesis.

We have discussed the molecular abundance ratios of COMs in different star-forming regions in the context of the different chemical routes proposed for their formation. We conclude that observations favor the formation of (CH₂OH)₂ by the combination of two CH₂OH radicals on dust grains, in agreement with the models of [Bennett & Kaiser \(2007\)](#) and [Garrod et al. \(2008\)](#), and the laboratory experiments by [Butscher et al. \(2015\)](#). The most likely formation route of CH₂OHCHO is via solid-phase dimerization of the formyl radical HCO, a mechanism proposed by [Woods et al. \(2013\)](#) and supported by the laboratory experiments of [Fedoseev et al. \(2015\)](#). The observational data also suggests that CH₃OCHO and CH₃OCH₃ are chemically linked, which may mean that both share a common precursor, CH₃O, as proposed by [Garrod et al. \(2008\)](#), or that the former is directly formed from the latter through viable gas-phase reactions, as suggested by [Balucani et al. \(2015\)](#). Similarly, EG and ET might share the common precursor CH₂OH.

The behavior of the abundance ratio C₂H₅OH/CH₃OCH₃ with luminosity may be explained by the different warm-up timescales in hot cores and hot corinos.

Acknowledgements. This work was partly supported by the Italian Ministero dell'Istruzione, Università e Ricerca through the grant Progetti Premiali 2012 – iALMA.

References

- Awad, Z., Viti, S., Bayet, E., & Caselli, P. 2014, *MNRAS*, **443**, 275
 Balucani, N., Ceccarelli C., & Taquet, V. 2015, *MNRAS*, **449**, L16
 Belloche, A., Müller, H. S. P., Menten, K. M., Schilke, P., & Comito, C. 2013, *A&A*, **559**, A47
 Beltrán, M. T., & de Wit, W. J. 2016, *A&ARv*, **24**, 6
 Beltrán, M. T., Cesaroni, R., Neri, R., et al. 2005, *A&A*, **435**, 901
 Beltrán, M. T., Codella, C., Viti, S., Neri, R., & Cesaroni, R. 2009, *ApJ*, **690**, L93
 Bennett, C. J., & Kaiser, R. I. 2007, *ApJ*, **661**, 899
 Bisschop, S. E., Jørgensen, J. K., van Dishoeck, E. F., & de Wachter, E. B. M. 2007, *A&A*, **465**, 913
 Biver, N., Bockelée-Morvan, D., Debout, V., et al. 2014, *A&A*, **566**, L5
 Biver, N., Bockelée-Morvan, D., Moreno, R., & Crovisier, J. A. A. 2015, *Sci. Adv.*, **1**, e1500863
 Brouillet, N., Despois, D., Lu, X.-H., et al. 2015, *A&A*, **576**, A129
 Butscher, T., Duvernay, F., Theule, P., et al. 2015, *MNRAS*, **453**, 1587
 Cesaroni, R., Beltrán, M. T., Zhang, Q., Beuther, H., & Fallscheer, C. 2011, *A&A*, **533**, A73
 Charnley, S. B., & Rodgers, S. D. 2005, in *Astrochemistry: Recent Successes and Current Challenges*, eds. D. C. Lis, G. A. Blake, & E. Herbst, 237
 Christen, D., & Müller, H. S. P. 2003, *Phys. Chem. Chem. Phys.*, **5**, 3600
 Christen, D., Coudert, L. H., Suenram, R., & Lovas, F. J. 1995, *J. Mol. Spectr.*, **172**, 57
 Christen, D., Coudert, L. H., Larsson, J. A., & Cremer, D. 2001, *J. Mol. Spectr.*, **205**, 185
 Churchwell, E., Walmsley, C. M., & Cesaroni, R. 1990, *A&AS*, **83**, 119
 Collins, P., & Ferrier, R. 1995, *Monosaccharides, Their Chemistry and Their Roles in Natural Products* (Wiley)
 Cooper, G., Kimmich, N., Belisle, W., et al. 2001, *Nature*, **414**, 879
 Coutens, A., Persson, M. V., Jørgensen, J. K., Wampfler, S. F., & Lykke, J. M. 2015, *A&A*, **576**, A5
 Crovisier, J., Bockelée-Morvan, D., Biver, N., et al. 2004, *A&A*, **418**, L35
 Favre, C., Despois, D., Brouillet, N., et al. 2011, *A&A*, **532**, A32
 Fedoseev, G., Cuppen, H. M., Ioppolo, S., Lamberts, T., & Linnartz, H. 2015, *MNRAS*, **448**, 1288
 Fontani, F., Pascucci, I., Caselli, P., et al. 2007, *A&A*, **470**, 639
 Fuente, A., Cernicharo, J., Caselli, P., et al. 2014, *A&A*, **568**, A65
 Garrod, R. T., Weaver, S. L. W., & Herbst, E. 2008, *ApJ*, **682**, 283
 Geppert, W. D., Hamberg, M., Thomas, R. D., et al. 2006, *Faraday Discussions*, **133**, 177
 Goesmann, F., Rosenbauer, H., Bredehöft, J. H., et al. 2015, *Science*, **349**, 020689
 Goldsmith, P. F., & Langer, W. D. 1999, *ApJ*, **517**, 209

- Halfen, D. T., Apponi, A. J., Woolf, N., Polt, R., & Ziurys, L. M. 2006, [ApJ](#), **639**, 237
- Ho, P. T. P., Moran, J. M., & Lo, K. Y., 2004, [ApJ](#), **616**, L1
- Hollis, J. M., Lovas, F. J., Jewell, P. R., & Coudert, L. H. 2002, [ApJ](#), **571**, L59
- Hudson, R. L., Moore, M. H., & Cook, A. M. 2005, [Adv. Space Res.](#), **36**, 184
- Ikeda, M., Ohishi, M., Nummelin, A., et al. 2001, [ApJ](#), **560**, 792
- Isokoski, K., Bottinelli, S., & van Dishoeck, E. F. 2013, [A&A](#), **554**, A100
- Jørgensen, J. K., Favre, C., Bisschop, S. E., et al. 2012, [ApJ](#), **757**, L4
- Krishnamurthy, R., Arrhenius, G., & Eschenmoser, A. 1999, [Origins of Life and Evolution of the Biosphere](#), **29**, 333
- Lykke, J. M., Favre, C., Bergin, E. A., & Jørgensen, J. K. 2015, [A&A](#), **582**, A64
- Maury, A. J., Belloche, A., André, Ph., et al. 2014, [A&A](#), **563**, L2
- Öberg, K. I., Garrod, R. T., van Dishoeck, E. F., & Linnartz, H. 2009, [A&A](#), **504**, 891
- Palau, A., Estalella, R., Girart, J. M., et al. 2014, [ApJ](#), **785**, 42
- Requena-Torres, M. A., Martín-Pintado, J., Martín, S., & Morris, M. R. 2008, [ApJ](#), **672**, 352
- Rivilla, V. M., Fontani, F., Beltrán, M. T., et al. 2016, [ApJ](#), **826**, 161
- Sorrell, W. H. 2001, [ApJ](#), **555**, L129
- Taquet, V., López-Sepulcre, A., Ceccarelli, C., et al. 2015, [ApJ](#), **804**, 81
- Walsh, C., Millar, T. J., Nomura, H., et al. 2014, [A&A](#), **563**, A33
- Weber, A. L. 1998, [Origins of Life and Evolution of the Biosphere](#), **28**, 259
- 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

Appendix A: Spectroscopy of ethylene glycol

Ethylene glycol is a triple rotor molecule. Torsions around its C–C bond and its two C–O bonds lead to ten different conformers, four forms with an *anti* (A) arrangement of the OH groups and six with a *gauche* (G) arrangement (Fig. 1 from Christen et al. 2001). Two of the G conformers are capable of establishing intramolecular hydrogen bonds, known as g'Ga and g'Gg⁸, which are expected to be energetically favorable in the gas phase. The aGg' conformer of EG is the lowest energy state conformer. The torsion of both hydroxyl groups (OH) around the two CO bonds produces a tunneling splitting. This translates into two sublevels, $v = 0$ and $v = 1$, separated by ~ 7 GHz. Within each tunneling sublevel, the different rotational energy levels are characterized by three quantum numbers, J , K_a , K_c . To label the different energy levels we use in this paper the following notation J_{K_a, K_c} , $v = 0$ or 1 . The transitions occur for changes from a rotational level in the $v = 1$ state to a rotational level in the $v = 0$ state, or viceversa. Figure 2 of Christen et al. (1995) shows a fraction of the energy

level diagram of the g'Ga conformer of EG. In our analysis we used the experimental spectrum of the aGg' conformer obtained by Christen et al. (1995) and Christen & Müller (2003).

Appendix B: Identified unblended transitions of COMs

We present in this appendix several tables with the unblended transitions of the different COMs found in the single-dish and interferometric data: ET (Tables B.1 and B.2), MF (Tables B.3 and B.4), GA (Tables B.5 and B.6), and DME (Tables B.7 and B.8).

In Fig. B.1 we also present some selected IRAM 30 m spectra at 3 mm of MF, ET, DME, and GA. We have overplotted in red the LTE fit using *MADCUBAIJ* with the physical parameters shown in Table 6.

In Fig. B.2 we present the unblended transitions of the COMs detected with the GBT telescope.

⁸ The lowercase letters a and g correspond to rotation at the CO bond where clockwise/counterclockwise rotation is indicated by a positive/negative dihedral angle and the symbols g/g'. The rotational directions are determined by looking along the C–C and C–O axes from the first to the second atom.

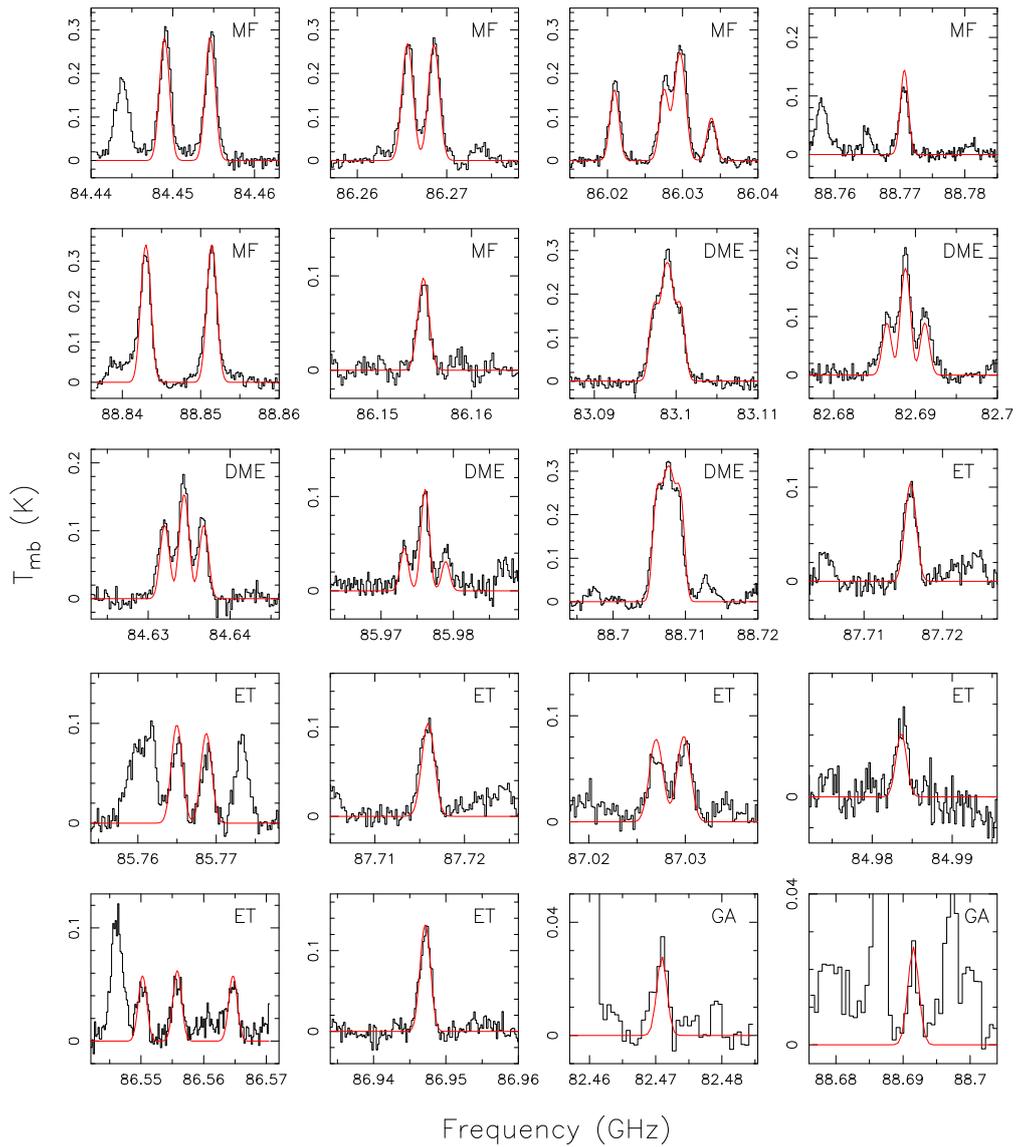


Fig. B.1. Selected transitions at 3 mm of COMs detected towards the G31.41+0.31 hot molecular core with the IRAM 30 m telescope: methyl formate (MF), dimethyl ether (DME), ethanol (ET) and glycolaldehyde (GA). Overplotted in red are the LTE synthetic spectra obtained with *MADCUBAII* using the physical parameters given in Table 6.

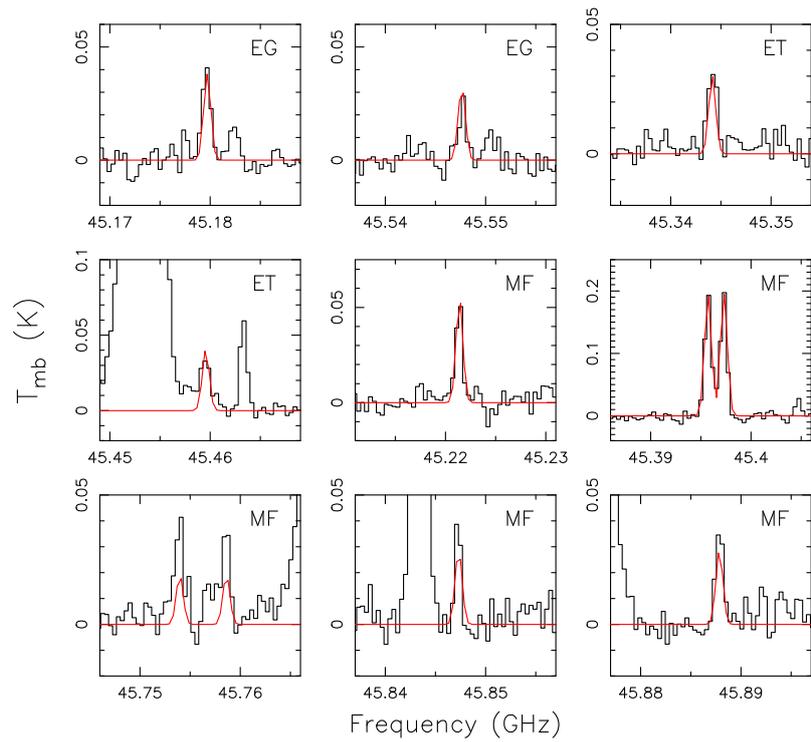


Fig. B.2. Unblended transitions of the COMs detected towards the G31.41+0.31 hot molecular core with the GBT telescope: ethylene glycol (EG), ethanol (ET) and methyl formate (MF). Overplotted in red are the LTE synthetic spectra obtained with *MADCUBAII* using the physical parameters given in Table C.1.

Table B.1. Unblended transitions of ET identified in the single-dish spectra (GBT and IRAM 30 m) towards G31.

Frequency (GHz)	Transition ^a	E_{up} (K)
45.34445	4(2, 2)–3(1, 2), $v_t = 0-1$	70
45.45987	13(3, 11)–12(4, 8)	88
81.68343	8(1, 7)–7(2, 6)	32
83.56877	5(1, 5)–4(1, 4), $v_t = 0-0$	70
83.66196	5(1, 5)–4(1, 4), $v_t = 1-1$	75
84.59587	4(2, 3)–4(1, 4)	13
85.26550	6(0, 6)–5(1, 5)	17
85.74718	5(0, 5)–4(0, 4), $v_t = 0-0$	69
85.76515	8(0, 8)–8(1, 8), $v_t = 1-0$	91
85.76890	5(0, 5)–4(0, 4), $v_t = 1-1$	74
86.31130	5(2, 4)–4(2, 3), $v_t = 0-0$	74
86.51654	5(4, 2)–4(4, 1), $v_t = 1-1$	94
86.51654	5(4, 1)–4(4, 0), $v_t = 1-1$	94
86.55042	5(3, 3)–4(3, 2), $v_t = 1-1$	85
86.55596	5(4, 2)–4(4, 1), $v_t = 0-0$	89
86.55596	5(4, 1)–4(4, 0), $v_t = 0-0$	89
86.56484	5(3, 2)–4(3, 1), $v_t = 1-1$	85
86.62171	5(3, 2)–4(3, 1), $v_t = 0-0$	80
87.02713	5(2, 3)–4(2, 2), $v_t = 1-1$	79
87.03001	5(2, 3)–4(2, 2), $v_t = 0-0$	74
87.71610	5(2, 4)–5(1, 5)	18
88.69737	8(1, 7)–7(0, 7), $v_t = 0-1$	89
98.58509	15(1, 15)–15(0, 15), $v_t = 1-0$	158
112.12954	12(3, 9)–12(2, 10)	77
140.50681	8(2, 6)–7(2, 5), $v_t = 0-0$	92
147.42746	8(1, 8)–7(0, 7)	30
154.93019	9(2, 8)–8(2, 7), $v_t = 1-1$	104
168.12339	10(0, 10)–9(0, 9), $v_t = 0-0$	102
168.24791	10(0, 10)–9(0, 9), $v_t = 1-1$	106
169.70219	18(3, 16)–18(2, 17)	155
171.96575	10(2, 9)–9(2, 8), $v_t = 1-1$	112
173.07117	10(6, 4)–9(6, 3), $v_t = 1-1$	152
173.07117	10(6, 5)–9(6, 4), $v_t = 1-1$	152
173.24010	10(5, 6)–9(5, 5), $v_t = 0-0$	133
173.24010	10(5, 5)–9(5, 4), $v_t = 0-0$	133
173.39133	5(2, 3)–4(1, 4)	18
173.39339	10(4, 6)–9(4, 5), $v_t = 0-0$	122
174.23291	6(2, 5)–5(1, 4)	23
224.44306	12(7, 6)–12(6, 6), $v_t = 0-1$	181
224.44311	12(7, 5)–12(6, 7), $v_t = 0-1$	181
224.59698	16(7, 10)–16(6, 10), $v_t = 0-1$	230
224.59850	16(7, 9)–16(6, 11), $v_t = 0-1$	230
224.66995	3(2, 1)–2(1, 1), $v_t = 1-0$	71
224.73929	11(3, 9)–10(2, 9), $v_t = 0-1$	123
224.82313	13(1, 12)–12(1, 11), $v_t = 0-0$	135
228.02905	13(3, 10)–13(3, 9), $v_t = 0-0$	144

Notes. ^(a) The transitions of the *gauche* state of ethanol are designated with $v_t = 0$ (*gauche* +) and $v_t = 1$ (*gauche* –), while the transitions without v number correspond to the *trans* state.

Table B.2. Unblended transitions of ET identified in the interferometric SMA spectra towards G31.

Frequency (GHz)	Transition ^a	E_{up} (K)
220.15478	24(3, 22)–24(2, 23)	263
220.99889	13(0, 13)–12(0, 12), $v_t = 1-1$	135.5
229.49113	17(5, 12)–17(4, 13)	160.1
230.67255	13(2, 11)–12(2, 10), $v_t = 0-0$	138.6
230.79386	6(5, 2)–5(4, 2), $v_t = 0-1$	104.8
230.95378	16(5, 11)–16(4, 12)	145.8
230.99137	14(0, 14)–13(1, 13)	85.5

Notes. ^(a) The transitions of the *gauche* state of ethanol are designated with $v_t = 0$ (*gauche* +) and $v_t = 1$ (*gauche* –), while the transitions without v number correspond to the *trans* state.

Table B.3. Unblended transitions (i.e. non blended with other molecular species) of MF identified in the single-dish spectra (GBT and IRAM 30 m) towards G31.

Frequency (GHz)	Transition	E_{up} (K)
45.22148	$v=1, 4(1, 4)-3(1, 3)$ E	193
45.39579	$v=0, 4(1, 4)-3(1, 3)$ E	6
45.39738	$v=0, 4(1, 4)-3(1, 3)$ A	6
45.75404	$v=0, 3(1, 3)-2(0, 2)$ E	4
45.75870	$v=0, 3(1, 3)-2(0, 2)$ A	4
45.84739	$v=0, 14(3, 11)-14(3, 12)$ E	70
45.88793	$v=0, 14(3, 11)-14(3, 12)$ A	70
81.31421	$v=0, 16(3, 13)-16(2, 14)$ E	89
81.36235	$v=0, 16(3, 13)-16(2, 14)$ A	89
81.38058	$v=0, 3(2, 1)-2(1, 2)$ E	6
82.24298	$v=0, 7(1, 7)-2(0, 6)$ E	16
82.24447	$v=0, 7(1, 7)-2(0, 6)$ A	16
82.52349	$v=0, 19(4, 15)-19(3, 16)$ E	126
82.56197	$v=0, 19(4, 15)-19(3, 16)$ A	126
83.60516	$v=0, 10(3, 8)-10(2, 9)$ E	38
83.63843	$v=0, 10(3, 8)-10(2, 9)$ A	38
84.22466	$v=0, 11(4, 7)-11(3, 8)$ E	50
84.23334	$v=0, 11(4, 7)-11(3, 8)$ A	50
84.28311	$v=1, 7(2, 6)-6(2, 5)$ E	206
84.44917	$v=0, 7(2, 6)-6(2, 5)$ E	19
84.45475	$v=0, 7(2, 6)-6(2, 5)$ A	19
85.32703	$v=1, 7(4, 4)-6(4, 3)$ A	215
85.37173	$v=1, 7(3, 5)-6(3, 4)$ A	210
85.50622	$v=1, 7(4, 3)-6(4, 2)$ E	215
85.55338	$v=1, 7(5, 3)-6(5, 2)$ E	220
85.63833	$v=0, 4(2, 3)-3(1, 2)$ E	9
85.74398	$v=1, 7(4, 4)-6(4, 3)$ E	214
85.77340	$v=0, 21(5, 16)-21(4, 17)$ A	156
85.78067	$v=0, 20(5, 15)-20(4, 16)$ E	143
85.78534	$v=0, 20(5, 15)-20(4, 16)$ A	143
85.91921	$v=0, 7(6, 1)-6(6, 0)$ E	40
86.02112	$v=0, 7(5, 2)-6(5, 1)$ E	33
86.02772	$v=0, 7(5, 3)-6(5, 2)$ E	33
86.02944	$v=0, 7(5, 3)-6(5, 2)$ A	33
86.03019	$v=0, 7(5, 2)-6(5, 1)$ A	33
86.03401	$v=1, 7(3, 4)-6(3, 3)$ E	210
86.15508	$v=1, 7(3, 4)-6(3, 3)$ A	210
86.17271	$v=1, 7(3, 5)-6(3, 4)$ E	209
86.21006	$v=0, 7(4, 4)-6(4, 3)$ A	27
86.22365	$v=0, 7(4, 3)-6(4, 2)$ E	27
86.22416	$v=0, 7(4, 4)-6(4, 3)$ E	27
86.25055	$v=0, 7(4, 3)-6(4, 2)$ A	27
86.26580	$v=0, 7(3, 5)-6(3, 4)$ A	23
86.26874	$v=0, 7(3, 5)-6(3, 4)$ E	23
87.14328	$v=0, 7(3, 4)-6(3, 3)$ E	23
87.16129	$v=0, 7(3, 4)-6(3, 3)$ A	23
87.76638	$v=0, 8(0, 8)-7(1, 7)$ E	20
87.76904	$v=0, 8(0, 8)-7(1, 7)$ A	20
88.05397	$v=0, 19(5, 14)-19(4, 15)$ E	130
88.05446	$v=0, 19(5, 14)-19(4, 15)$ A	130
88.17551	$v=0, 10(4, 6)-10(3, 7)$ E	43
88.18042	$v=0, 10(4, 6)-10(3, 7)$ A	43
88.22075	$v=1, 7(1, 6)-6(1, 5)$ E	205
88.35849	$v=0, 22(5, 17)-22(4, 18)$ A	170
88.68689	$v=0, 11(3, 9)-11(2, 10)$ E	45
88.72327	$v=0, 11(3, 9)-11(2, 10)$ A	45
88.77087	$v=1, 8(1, 8)-7(1, 7)$ A	208
88.85161	$v=0, 7(1, 6)-6(1, 5)$ A	18
88.86241	$v=1, 8(1, 8)-7(1, 7)$ E	207
98.42421	$v=0, 8(5, 3)-7(5, 2)$ E	38
98.43180	$v=0, 8(5, 4)-7(5, 3)$ E	38

Table B.3. continued.

Frequency (GHz)	Transition	E_{up} (K)
98.43276	$v=0, 8(5, 4)-7(5, 3)$ A	38
98.43580	$v=0, 8(5, 3)-7(5, 2)$ A	38
98.60686	$v=0, 8(3, 6)-7(3, 5)$ E	27
98.61116	$v=0, 8(3, 6)-7(3, 5)$ A	27
98.68242	$v=1, 8(3, 6)-7(3, 5)$ E	214
98.68261	$v=0, 8(4, 5)-7(4, 4)$ A	32
98.71200	$v=0, 8(4, 5)-7(4, 4)$ E	32
98.74791	$v=0, 8(4, 4)-7(4, 3)$ E	32
98.79229	$v=0, 8(4, 4)-7(4, 3)$ E	32
110.05033	$v=1, 9(6, 4)-8(6, 3)$ E	237
110.15365	$v=1, 10(1, 10)-9(1, 9)$ A	218
110.23871	$v=1, 10(1, 10)-9(1, 9)$ E	217
111.40841	$v=0, 9(4, 5)-8(4, 4)$ E	37
111.45330	$v=0, 9(4, 5)-8(4, 4)$ A	37
111.67413	$v=0, 9(1, 8)-8(1, 7)$ E	28
111.68219	$v=0, 9(1, 8)-8(1, 7)$ A	28
111.73400	$v=0, 10(1, 10)-9(0, 9)$ E	30
111.73531	$v=0, 10(1, 10)-9(0, 9)$ A	30
140.16666	$v=1, 11(2, 9)-10(2, 8)$ A	231
147.24796	$v=0, 18(6, 13)-18(5, 14)$ E	125
147.24796	$v=1, 12(5, 7)-11(5, 6)$ E	250
147.25078	$v=0, 12(11, 1)-11(11, 0)$ E	126
147.25568	$v=0, 12(11, 1)-11(11, 0)$ A	126
147.25568	$v=0, 12(11, 2)-11(11, 1)$ A	126
147.26531	$v=0, 12(11, 2)-11(11, 1)$ E	126
147.30479	$v=0, 19(6, 14)-19(5, 15)$ E	137
147.31057	$v=0, 12(10, 2)-11(10, 1)$ E	112
147.31775	$v=0, 12(10, 2)-11(10, 1)$ A	112
147.31775	$v=0, 12(10, 3)-11(10, 2)$ A	112
147.32539	$v=0, 12(10, 3)-11(10, 2)$ E	112
147.33163	$v=0, 19(6, 14)-19(5, 15)$ A	137
147.39707	$v=0, 12(9, 3)-11(9, 2)$ E	100
147.40637	$v=0, 12(9, 3)-11(9, 2)$ A	100
147.40637	$v=0, 12(9, 4)-11(9, 3)$ A	100
147.41182	$v=0, 12(9, 4)-11(9, 3)$ E	100
147.52431	$v=0, 12(8, 4)-11(8, 3)$ E	89
147.53554	$v=0, 12(8, 5)-11(8, 4)$ A	88
147.53554	$v=0, 12(8, 4)-11(8, 3)$ A	88
147.53864	$v=0, 12(8, 5)-11(8, 4)$ E	88
147.53917	$v=0, 17(6, 12)-17(5, 13)$ E	115
154.98454	$v=0, 12(3, 9)-11(3, 8)$ E	53
155.00232	$v=0, 12(3, 9)-11(3, 8)$ A	53
161.41614	$v=0, 13(5, 8)-12(5, 7)$ E	71
161.45743	$v=0, 20(2, 18)-20(1, 19)$ A	128
161.45822	$v=0, 13(5, 8)-12(5, 7)$ A	70
168.49507	$v=0, 13(3, 10)-12(3, 9)$ E	61
168.51375	$v=0, 13(3, 10)-12(3, 9)$ A	61
168.91461	$v=0, 15(2, 13)-14(3, 12)$ E	76
168.93457	$v=0, 15(2, 13)-14(3, 12)$ A	76
169.57068	$v=0, 20(7, 13)-20(6, 14)$ A	157
170.23327	$v=0, 14(3, 12)-13(3, 11)$ E	68
170.59390	$v=0, 6(4, 2)-5(3, 3)$ A	23
171.06714	$v=0, 23(7, 17)-23(6, 18)$ A	196
171.79421	$v=0, 14(13, 1)-13(13, 0)$ E	174
171.79451	$v=0, 14(13, 1)-13(13, 0)$ A	174
171.79451	$v=0, 14(13, 2)-13(13, 1)$ A	174
171.84471	$v=0, 14(12, 2)-13(12, 1)$ E	157
171.84664	$v=0, 19(1, 18)-19(1, 19)$ E	109
171.84789	$v=0, 14(12, 3)-13(12, 2)$ A	157
171.84792	$v=0, 14(12, 2)-13(12, 1)$ A	157
171.84992	$v=0, 19(1, 18)-19(0, 19)$ E	109
171.86084	$v=0, 14(12, 3)-13(12, 2)$ E	157
171.88085	$v=0, 25(5, 21)-25(4, 22)$ E	210
171.91579	$v=0, 14(11, 3)-13(11, 2)$ E	142

Notes. Integrated intensities derived from the LTE fit.

Table B.3. continued.

Frequency (GHz)	Transition	E_{up} (K)
171.92174	$v=0, 14(11, 3)-13(11, 2)$ A	142
171.92174	$v=0, 14(11, 4)-13(11, 3)$ A	142
171.93258	$v=0, 14(11, 4)-13(11, 3)$ E	142
171.93339	$v=0, 19(1, 18)-19(1, 19)$ A	109
171.93528	$v=0, 25(5, 21)-25(4, 22)$ A	210
171.93658	$v=0, 19(1, 18)-19(0, 19)$ A	109
171.99364	$v=0, 19(2, 18)-19(0, 19)$ E	109
172.01564	$v=0, 14(10, 4)-13(10, 3)$ E	128
172.02428	$v=0, 14(10, 4)-13(10, 3)$ A	128
172.02428	$v=0, 14(10, 5)-13(10, 4)$ A	128
172.15764	$v=0, 14(9, 5)-13(9, 4)$ E	116
172.16873	$v=0, 14(9, 6)-13(9, 5)$ A	116
172.16873	$v=0, 14(9, 5)-13(9, 4)$ A	116
172.36437	$v=0, 14(8, 6)-13(8, 5)$ E	104
172.37775	$v=0, 14(8, 7)-13(8, 6)$ A	104
172.37775	$v=0, 14(8, 6)-13(8, 5)$ A	104
172.38095	$v=0, 14(8, 7)-13(8, 6)$ E	104
172.38578	$v=0, 15(1, 14)-14(2, 13)$ E	71
172.39349	$v=0, 15(1, 14)-14(2, 13)$ A	71
172.69216	$v=0, 14(7, 8)-13(7, 7)$ A	95
172.69337	$v=0, 14(7, 8)-13(7, 7)$ E	95
172.69337	$v=0, 14(7, 7)-13(7, 6)$ A	95
173.19427	$v=0, 14(6, 9)-13(6, 8)$ E	86
173.21868	$v=0, 14(6, 8)-13(6, 7)$ A	86
173.51541	$v=0, 15(2, 14)-14(2, 13)$ E	71
173.52168	$v=0, 15(2, 14)-14(2, 13)$ A	71
173.65011	$v=0, 14(4, 11)-13(4, 10)$ A	73
173.81936	$v=0, 14(5, 10)-13(5, 9)$ A	79
173.82245	$v=0, 14(5, 10)-13(5, 9)$ E	79
174.20980	$v=0, 15(1, 14)-14(1, 13)$ E	71
174.21556	$v=0, 15(1, 14)-14(1, 13)$ A	71
174.21802	$v=0, 27(6, 22)-27(5, 23)$ A	249
174.37741	$v=0, 14(5, 9)-13(5, 8)$ E	79
174.40618	$v=0, 14(5, 9)-13(5, 8)$ A	79
174.54656	$v=0, 16(0, 16)-15(1, 15)$ E	73
174.54801	$v=0, 16(0, 16)-15(1, 15)$ A	73
174.56580	$v=0, 16(1, 16)-15(1, 15)$ E	73
174.56721	$v=0, 16(1, 16)-15(1, 15)$ A	73
174.58112	$v=0, 16(0, 16)-15(0, 15)$ E	73
174.58251	$v=0, 16(0, 16)-15(0, 15)$ A	73
174.60035	$v=0, 16(1, 16)-15(0, 15)$ E	73
174.60170	$v=0, 16(1, 16)-15(0, 15)$ A	73
224.31315	$v=0, 18(5, 14)-17(5, 13)$ E	118
224.60938	$v=0, 18(6, 12)-17(6, 11)$ A	125
237.29691	$v=0, 13(4, 10)-12(3, 9)$ E	65
237.29748	$v=0, 20(2, 18)-19(2, 17)$ E	128
237.30597	$v=0, 20(2, 18)-19(2, 17)$ A	128
237.30954	$v=0, 21(2, 20)-20(2, 19)$ E	132
237.31508	$v=0, 21(2, 20)-20(2, 19)$ A	132
237.31705	$v=0, 13(4, 10)-12(3, 9)$ A	65
237.34487	$v=0, 21(1, 20)-20(1, 19)$ E	132
237.35039	$v=0, 21(1, 20)-20(1, 19)$ A	132
258.08104	$v=0, 22(2, 20)-21(2, 19)$ E	152
258.08949	$v=0, 22(2, 20)-21(2, 19)$ A	152
258.30628	$v=0, 11(5, 7)-10(4, 6)$ A	56
258.49087	$v=0, 23(2, 22)-22(2, 21)$ E	156
258.49624	$v=0, 23(2, 22)-22(2, 21)$ A	156
258.49933	$v=0, 21(12, 10)-20(12, 9)$ E	232
258.50273	$v=0, 23(1, 22)-22(1, 21)$ E	156
258.50818	$v=0, 23(1, 22)-22(1, 21)$ A	156

Table B.4. Unblended transitions (i.e. non blended with other molecular species) of MF identified in the interferometric SMA spectra towards G31.

Frequency (GHz)	Transition	E_{up} (K)
219.62269	$v=1, 18(12, 6)-17(12, 5)$ A	384
219.62269	$v=1, 18(12, 7)-17(12, 6)$ A	384
219.64240	$v=1, 18(13, 6)-17(13, 5)$ E	401
219.69583	$v=1, 18(11, 8)-17(11, 7)$ A	369
219.69583	$v=1, 18(11, 7)-17(11, 6)$ A	369
219.70513	$v=1, 18(4, 15)-17(4, 14)$ A	299
220.03034	$v=1, 18(9, 10)-17(9, 9)$ A	342
220.03034	$v=1, 18(9, 9)-17(9, 8)$ A	342
220.16689	$v=0, 17(4, 13)-16(4, 12)$ E	103
220.25866	$v=1, 18(8, 10)-17(8, 9)$ E	331
220.30738	$v=1, 18(10, 9)-17(10, 8)$ E	354
220.36833	$v=1, 18(8, 11)-17(8, 10)$ A	331
220.36988	$v=1, 18(8, 10)-17(8, 9)$ A	331
220.91395	$v=1, 18(7, 12)-17(7, 11)$ A	321
220.92617	$v=0, 18(16, 2)-17(16, 1)$ A	271
220.92617	$v=1, 18(16, 3)-17(16, 2)$ A	271
220.93518	$v=0, 18(16, 2)-17(16, 1)$ E	271
220.94742	$v=0, 18(16, 3)-17(16, 2)$ E	271
220.97782	$v=1, 18(15, 4)-17(15, 3)$ A	250
220.97782	$v=1, 18(15, 3)-17(15, 2)$ A	250
220.98367	$v=0, 18(15, 3)-17(15, 2)$ E	250
220.98533	$v=1, 18(7, 11)-17(7, 10)$ E	321
221.04766	$v=1, 18(14, 5)-17(14, 4)$ A	231
221.04766	$v=1, 18(14, 4)-17(14, 3)$ A	231
221.04999	$v=0, 18(14, 4)-17(14, 3)$ E	231
221.06693	$v=0, 18(14, 5)-17(14, 4)$ E	231
221.11067	$v=1, 18(8, 11)-17(8, 10)$ E	330
221.13972	$v=0, 18(13, 5)-17(13, 4)$ E	213
221.14113	$v=0, 18(13, 5)-17(13, 4)$ A	213
221.14113	$v=0, 18(14, 4)-17(14, 3)$ A	213
221.15854	$v=0, 18(13, 6)-17(13, 5)$ E	213
229.40502	$v=0, 18(3, 15)-17(3, 14)$ E	111
229.42034	$v=0, 18(3, 15)-17(3, 14)$ A	111
230.87881	$v=1, 18(4, 14)-17(4, 13)$ A	301
231.24542	$v=1, 19(4, 16)-18(4, 15)$ A	310

Table B.5. Unblended transitions (i.e. non blended with other molecular species) of GA identified in the single IRAM 30 m single-dish spectra towards G31.

Frequency (GHz)	Transition	E_{up} (K)
82.47067	$8(0, 8)-7(1, 7)$	19
88.69126	$12(3, 10)-12(2, 11)$	49

Table B.6. Unblended transition of GA identified in the interferometric SMA spectra towards G31.

Frequency (GHz)	Transition	E_{up} (K)
230.89847	$21(2, 19)-20(3, 18)$	131

Table B.7. Unblended transitions of DME identified in the IRAM 30 m single-dish spectra towards G31. **Table B.7.** continued.

Frequency (GHz)	Transition	E_{up} (K)
82.45696	11(1, 10)–11(0, 11) AE	63
82.45696	11(1, 10)–11(0, 11) EA	63
82.45862	11(1, 10)–11(0, 11) EE	63
82.46038	11(1, 10)–11(0, 11) AA	63
82.64930	3(1, 3)–2(0, 2) EA	7
82.64930	3(1, 3)–2(0, 2) AE	7
82.65018	3(1, 3)–2(0, 2) EE	7
82.65108	3(1, 3)–2(0, 2) AA	7
82.68650	4(2, 2)–4(1, 3) AE	15
82.68650	4(2, 2)–4(1, 3) EA	15
82.68877	4(2, 2)–4(1, 3) EE	15
82.69114	4(2, 2)–4(1, 3) AA	15
83.09738	14(2, 12)–14(1, 13) EA	103
83.09738	14(2, 12)–14(1, 13) AE	103
83.09892	14(2, 12)–14(1, 13) EE	103
83.10044	14(2, 12)–14(1, 13) AA	103
84.63202	3(2, 1)–3(1, 2) AE	11
84.63202	3(2, 1)–3(1, 2) EA	11
84.63440	3(2, 1)–3(1, 2) EE	11
84.63680	3(2, 1)–3(1, 2) AA	11
85.97322	13(2, 12)–12(3, 9) AA	88
85.97610	13(2, 12)–12(3, 9) EE	88
85.97894	13(2, 12)–12(3, 9) EA	88
85.97894	13(2, 12)–12(3, 9) AE	88
88.70622	15(2, 13)–15(1, 14) EA	117
88.70622	15(2, 13)–15(1, 14) AE	117
88.70764	15(2, 13)–15(1, 14) EE	117
88.70907	15(2, 13)–15(1, 14) AA	117
109.57140	8(2, 7)–8(1, 8) EA	38
109.57140	8(2, 7)–8(1, 8) AE	38
109.57409	8(2, 7)–8(1, 8) EE	38
109.57678	8(2, 7)–8(1, 8) AA	38
111.74135	19(3, 16)–19(2, 17) AE	187
111.74135	19(3, 16)–19(2, 17) EA	187
111.74279	19(3, 16)–19(2, 17) EE	187
111.74424	19(3, 16)–19(2, 17) AA	187
111.78256	7(0, 7)–6(1, 6) AA	25
111.78365	7(0, 7)–6(1, 6) EE	25
111.78365	7(0, 7)–6(1, 6) AE	25
111.78403	7(0, 7)–6(1, 6) EA	25
111.81205	18(3, 15)–18(2, 16) AE	170
111.81205	18(3, 15)–18(2, 16) EA	170
111.81367	18(3, 15)–18(2, 16) EE	170
111.81529	18(3, 15)–18(2, 16) AA	170
147.20209	6(3, 4)–6(2, 5) EA	32
147.20376	6(3, 4)–6(2, 5) AE	32
147.20681	6(3, 4)–6(2, 5) EE	32
147.21071	6(3, 4)–6(2, 5) AA	32
154.45366	11(1, 10)–10(2, 9) AA	63
154.45508	11(1, 10)–10(2, 9) EE	63
154.45650	11(1, 10)–10(2, 9) AE	63
154.45651	11(1, 10)–10(2, 9) EA	63
169.74067	16(3, 14)–16(2, 15) EA	137
169.74067	16(3, 14)–16(2, 15) AE	137
169.74349	16(3, 14)–16(2, 15) EE	137
169.74631	16(3, 14)–16(2, 15) AA	137
169.90032	16(2, 15)–16(1, 16) EA	128
169.90032	16(2, 15)–16(1, 16) AE	128
169.90369	16(2, 15)–16(1, 16) EE	128
169.90705	16(2, 15)–16(1, 16) AA	128

Frequency (GHz)	Transition	E_{up} (K)
173.29268	10(0, 10)–9(1, 9) AA	49
173.29297	10(0, 10)–9(1, 9) EE	49
173.29327	10(0, 10)–9(1, 9) AE	49
173.29327	10(0, 10)–9(1, 9) EA	49
174.89053	17(3, 15)–17(2, 16) EA	152
174.89053	17(3, 15)–17(2, 16) AE	152
174.89329	17(3, 15)–17(2, 16) EE	152
174.89605	17(3, 15)–17(2, 16) AA	152
209.51519	11(1, 11)–10(0, 10) EA	59
209.51519	11(1, 11)–10(0, 10) AE	59
209.51561	11(1, 11)–10(0, 10) EE	59
209.51603	11(1, 11)–10(0, 10) AA	59
237.61880	9(2, 8)–8(1, 7) EA	46
237.61881	9(2, 8)–8(1, 7) AE	46
237.62089	9(2, 8)–8(1, 7) EE	46
237.62297	9(2, 8)–8(1, 7) AA	46

Table B.8. Unblended transitions (i.e. non blended with other molecular species) of DME identified in the interferometric SMA spectra towards G31.

Frequency (GHz)	Transition	E_{up} (K)
220.84652	24(4, 20)–23(5, 19) AA	298
220.84764	24(4, 20)–23(5, 19) EE	298
220.84876	24(4, 20)–23(5, 19) AE	298
220.84876	24(4, 20)–23(5, 19) EA	298
220.89182	23(4, 20)–23(3, 21) EA	274
220.89182	23(4, 20)–23(3, 21) AE	274
220.89311	23(4, 20)–23(3, 21) EE	274
220.89500	23(4, 20)–23(3, 21) AA	274
221.19722	27(5, 22)–27(4, 23) AA	381
221.19782	27(5, 22)–27(4, 23) EE	381
221.19842	27(5, 22)–27(4, 23) AE	381
221.19842	27(5, 22)–27(4, 23) EA	381
230.14003	25(4, 22)–25(3, 22) EA	319
230.14003	25(4, 22)–25(3, 22) AE	319
230.14137	25(4, 22)–25(3, 22) EE	319
230.14271	25(4, 22)–25(3, 22) AA	319
230.23215	17(2, 15)–16(3, 14) AA	148
230.23376	17(2, 15)–16(3, 14) EE	148
230.23536	17(2, 15)–16(3, 14) AE	148
230.23536	17(2, 15)–16(3, 14) EA	148

Appendix C: MADCUBAIJ LTE fits of the GBT and SMA data

In this appendix we present the physical parameters used to model with *MADCUBAIJ* the transitions of the different COMs in the GBT and SMA spectra (Figs. B.2 and 2, respectively). The procedure used is: i) we fix the velocity of the emission and the linewidth to the values obtained from Gaussian fits of unblended lines; ii) to obtain the source-average column density N_s , we fix the size of the molecular emission to the value obtained from the integrated maps of unblended lines detected with the SMA, and apply the beam dilution factor accordingly; while we do not apply the beam dilution factor to derive the beam-average column density N_b ; iii) leaving N and T_{ex} as free parameters, we use the *AUTOFIT* tool of *MADCUBAIJ* to find the pair of values that fits the unblended molecular transitions of each COM better.

Table C.1. MADCUBAIJ best fits of the COMs detected in the GBT spectra.

	EG	ET	MF
$\log N_b$ (cm ⁻²)	15.1	15.5	16.1
$\log N_s$ (cm ⁻²)	17.9	17.7	18.1
T_{ex} (K) ^a	147 ^b /142 ^b	89/76	147/142
v_{LSR} (km s ⁻¹)	97.0	99.0	97.5
Δv (km s ⁻¹)	5.5	5.4	5.3

Notes. The parameters of this table have been used to produce the synthetic spectra of Fig. B.2. ^(a) The first temperature corresponds to the beam-average value, while the second is the source-average value obtained assuming the source diameter given in Table 5. ^(b) Since the energy range covered by the transitions is very narrow, we have fixed the temperature to the values obtained for MF.

Table C.2. MADCUBAIJ best fits of the SMA compact spectra towards the peak of the continuum for the different molecules.

	EG	ET	GA	MF	DME
$\log N_b$ (cm ⁻²)	16.3	17.2	16.0	17.5	17.7
$\log N_s$ (cm ⁻²)	17.7	17.7	17.4	17.9	18.7
T_{ex} (K) ^a	88 / 75	187 / 184	165 ^a	240/243	113/105
v_{LSR} (km s ⁻¹)	97.2	97.0	97.0	97.0	97.2
Δv (km s ⁻¹)	5.2	5.9	6.0	6.4	7.6

Notes. The parameters of this table have been used to produce the synthetic spectra of Fig. 2. ^(a) The first temperature corresponds to the beam-average value, while the second is the source-average value obtained assuming the source diameter given in Table 5.