

The GUAPOS project: G31.41+0.31 Unbiased ALMA sPectral Observational Survey

I. Isomers of C₂H₄O₂★

C. Mininni^{1,2}, M. T. Beltrán², V. M. Rivilla^{3,2}, A. Sánchez-Monge⁴, F. Fontani², T. Möller⁴, R. Cesaroni², P. Schilke⁴, S. Viti^{5,7}, I. Jiménez-Serra³, L. Colzi^{3,2}, A. Lorenzani², and L. Testi⁶

¹ Dipartimento di Fisica e Astronomia, Università degli Studi di Firenze, Via Sansone 1, 50019 Sesto Fiorentino, Italy

e-mail: chiara.mininni.astro@gmail.com; chiara.mininni@unifi.it

² INAF Osservatorio Astrofisico di Arcetri, Largo E. Fermi 5, 50125 Firenze, Italy

³ Centro de Astrobiología (CSIC, INTA), Ctra. de Ajalvir, km. 4, Torrejón de Ardoz, 28850 Madrid, Spain

⁴ I. Physikalisches Institut, Universität zu Köln, Zùlpicher Str. 77, 50937 Köln, Germany

⁵ Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, The Netherlands

⁶ European Southern Observatory (ESO), Karl-Schwarzschild-Str. 2, 85748 Garching, Germany

⁷ Department of Physics and Astronomy, UCL, Gower Place, London WC1E 6BT, UK

Received 18 July 2020 / Accepted 22 September 2020

ABSTRACT

Context. One of the goals of astrochemistry is to understand the degree of chemical complexity that can be reached in star-forming regions, along with the identification of precursors of the building blocks of life in the interstellar medium. To answer such questions, unbiased spectral surveys with large bandwidth and high spectral resolution are needed, in particular, to resolve line blending in chemically rich sources and identify each molecule (especially for complex organic molecules). These kinds of observations have already been successfully carried out, primarily towards the Galactic Center, a region that shows peculiar environmental conditions.

Aims. We present an unbiased spectral survey of one of the most chemically rich hot molecular cores located outside the Galactic Center, in the high-mass star-forming region G31.41+0.31. The aim of this 3mm spectral survey is to identify and characterize the physical parameters of the gas emission in different molecular species, focusing on complex organic molecules. In this first paper, we present the survey and discuss the detection and relative abundances of the three isomers of C₂H₄O₂: methyl formate, glycolaldehyde, and acetic acid.

Methods. Observations were carried out with the ALMA interferometer, covering all of band 3 from 84 to 116 GHz (~32 GHz bandwidth) with an angular resolution of 1.2'' × 1.2'' (~ 4400 au × 4400 au) and a spectral resolution of ~0.488 MHz (~1.3–1.7 km s⁻¹). The transitions of the three molecules have been analyzed with the software XCLASS to determine the physical parameters of the emitted gas.

Results. All three isomers were detected with abundances of (2 ± 0.6) × 10⁻⁷, (4.3–8) × 10⁻⁸, and (5.0 ± 1.4) × 10⁻⁹ for methyl formate, acetic acid, and glycolaldehyde, respectively. Methyl formate and acetic acid abundances are the highest detected up to now, if compared to sources in the literature. The size of the emission varies among the three isomers with acetic acid showing the most compact emission while methyl formate exhibits the most extended emission. Different chemical pathways, involving both grain-surface chemistry and cold or hot gas-phase reactions, have been proposed for the formation of these molecules, but the small number of detections, especially of acetic acid and glycolaldehyde, have made it very difficult to confirm or discard the predictions of the models. The comparison with chemical models in literature suggests the necessity of grain-surface routes for the formation of methyl formate in G31, while for glycolaldehyde both scenarios could be feasible. The proposed grain-surface reaction for acetic acid is not capable of reproducing the observed abundance in this work, while the gas-phase scenario should be further tested, given the large uncertainties involved.

Key words. astrochemistry – ISM: molecules – stars: formation – ISM: individual objects: G31.41+0.31

1. Introduction

Hot molecular cores (HMCs), the birthplaces of massive stars, are the sources that show the highest level of chemical complexity in the Galaxy (Herbst & van Dishoeck 2009). Their spectra include emission from a large variety of molecules, starting from the simplest diatomic species up through complex organic molecules (COMs), which are molecules containing carbon with

six or more atoms. In recent years, our view and comprehension of the chemistry of the interstellar medium (ISM) has improved substantially and species with up to 13 atoms have been unambiguously detected (c-C₆H₅CN, McGuire et al. 2018). The advent of more sensitive instruments with high spectral and spatial resolution, such as ALMA (Atacama Large Millimeter Array), has allowed the detection of emission from faint molecular species, particularly heavy COMs since these typically emit a large number of transitions that could be faint or blended with those of other molecules.

This chemically rich environment and the presence of several COMs is thought to be the result of the evaporation of the

* Tables C.1–C.3 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/cat/J/A+A/644/A84>

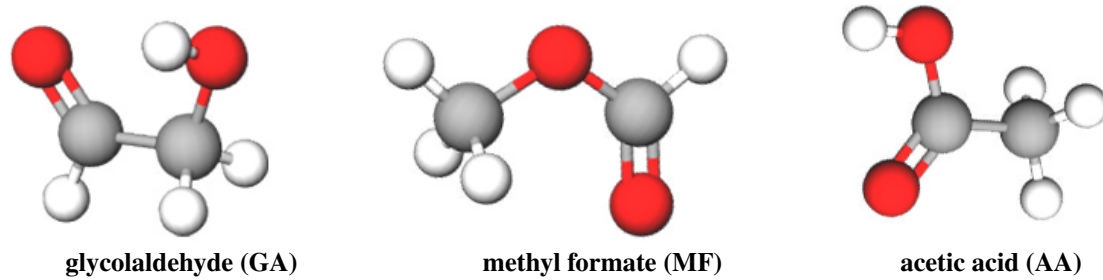


Fig. 1. Chemical structures of the three isomers of C₂H₄O₂.

products of grain-surface reactions thanks to the presence of an already formed proto-stellar object(s), which starts to heat up the surrounding medium, and of the subsequent hot gas-phase chemistry. The detection of COMs in cold environments as well (Öberg et al. 2010; Bacmann et al. 2012; Vastel et al. 2014; Jiménez-Serra et al. 2016) has opened up new scenarios and other possible chemical routes in cold gas-phase have been proposed to explain these observations (i.e., Vasyunin & Herbst 2013; Balucani et al. 2015; Vasyunin et al. 2017). Thus, it is possible that in HMCs, some of the observed COMs emission is also inherited from the cold chemistry at early stages of the star-formation process. Hence, these new routes need to be considered in chemical models.

To investigate the maximum degree of chemical complexity in star-forming regions, spectral surveys have been carried out towards a few sources, of both high-mass and low-mass cores (Kalenskii & Johansson 2010; Belloche et al. 2016; Jørgensen et al. 2016). One of the most studied sources is SgrB2 (Belloche et al. 2013, 2016, 2019; Sánchez-Monge et al. 2017), although the environmental conditions of this source, cataloged as a mini-starburst region, cannot be considered as typical of HMCs because its peculiar environment is likely to have had an impact on the chemistry.

With this in mind, we decided to perform a spectral survey covering the entire ALMA band 3 towards one of the most chemically rich HMCs outside the Galactic Center (GC), G31.41+0.31 (hereafter G31), where glycolaldehyde (CH₂OHCHO, GA) has been detected for the first time outside the GC (Beltrán et al. 2009).

The target of the G31.41+0.31 Unbiased ALMA sPectral Observational Survey (GUAPOS) is a well-known and studied HMC located at a distance of 3.75 kpc (Immer et al. 2019) with a luminosity of $4.4 \times 10^4 L_{\odot}$ (from Osorio et al. 2009) and a mass $M \sim 70 M_{\odot}$ (Cesaroni 2019), after rescaling to the new distance estimate (the previous distance estimate was of 7.9 kpc). The source was first detected in NH₃(4, 4) (Cesaroni et al. 1994a), whose emission was co-spatial to the distribution of water masers (Hofner & Churchwell 1993), and separated by $\sim 5''$ from an already known ultra compact (UC) HII region. Cesaroni et al. (1994b) discovered a velocity gradient of $\sim 400 \text{ km s}^{-1} \text{ pc}^{-1}$ across the core in the emission of methyl cyanide (CH₃CN) at 3 mm, along the SW-NE direction. The nature of this velocity gradient was further investigated in more recent studies and is most likely associated with a rotating toroid with a spin up toward the center (Beltrán et al. 2004, 2005, 2018; Cesaroni et al. 2010, 2011, 2017). Although another interpretation in terms of expansion was proposed (Gibb et al. 2004; Araya et al. 2008), this is inconsistent with recent polarization observations (Beltrán et al. 2019). These studies revealed also the presence of several outflows, traced by CO and SiO emission (Cesaroni et al. 2011; Beltrán et al. 2018), of infalling material (firstly found by

Girart et al. 2009; see also Mayen-Gijon et al. 2014 and Beltrán et al. 2018) and of two free-free sources at 0.7 and 1.3 cm separated by $0''.19$ embedded in the core (Cesaroni et al. 2010).

G31 has also been studied in polarized emission at mm wavelengths by Girart et al. (2009) and, recently, by Beltrán et al. (2019). The reconstructed shape of the magnetic field follows an hourglass morphology, with the best model of Beltrán et al. (2019) suggesting a predominant poloidal field, oriented perpendicular to the SW-NE velocity gradient previously detected in G31, lending strong support to the rotating toroid scenario, as already mentioned.

G31 has also been studied from a chemical point of view, revealing one of the most striking features of this HMC: its chemical richness. As previously mentioned, the first detection of glycolaldehyde outside the GC was obtained towards G31 (Beltrán et al. 2009), and confirmed by Calcutt et al. (2014) and Rivilla et al. (2017). Other heavy COMs such as ethyl cyanide C₂H₅CN, dimethyl ether CH₃OCH₃, ethanol C₂H₅OH, and ethylene glycol (CH₂OH)₂ (nine, nine, nine, and ten atoms respectively) were also detected here (Beltrán et al. 2005; Fontani et al. 2007; Isokoski et al. 2013; Rivilla et al. 2017; Coletta et al. 2020). The chemical richness of this HMC has been quantified by Cesaroni et al. (2017), who reported that the fraction of channels with line emission was 0.74 (the largest in the sample of high-mass young stellar objects studied by these authors) in the 217–237 GHz range. G31 is, thus, one of the best candidates for investigating the degree of chemical complexity that can be reached in typical high-mass star-forming regions.

Among COMs, a special class of molecules is made up of isomers, which are molecules with the same chemical composition, but with a different molecular structure. These molecules can be used to explore which chemical formation pathway is more efficient, leading to a predominance of one isomer with respect to the other(s) and helping to constrain the chain of chemical reactions involved in their formation. Among isomers, those of C₂H₄O₂ (see Fig. 1), namely glycolaldehyde (CH₂OHCHO, hereafter GA), methyl formate (CH₃OCHO, hereafter MF), and acetic acid (CH₃COOH, hereafter AA), are especially interesting because of their relevance for the formation of prebiotic molecules. In fact, GA is the simplest sugar-related molecule and can react with propenal to form ribose, an essential constituent of RNA. The three molecules were all firstly detected towards SgrB2 (Brown et al. 1975; Mehringer et al. 1997; Hollis et al. 2000). MF has been detected in a large number of objects including high-mass star-forming regions (Beuther et al. 2007; Fontani et al. 2007; Favre et al. 2014; Sakai et al. 2015; Belloche et al. 2016), low-mass star-forming regions (Cazaux et al. 2003; Bottinelli et al. 2007; Jørgensen et al. 2012; Jacobsen et al. 2019), prestellar cores (Bacmann et al. 2012; Jiménez-Serra et al. 2016), cold envelopes around protostars (Öberg et al. 2010; Cernicharo et al. 2012), and outflow and shock regions

(Arce et al. 2008; Csengeri et al. 2019). On the other hand, the number of detections of GA and AA is limited compared to that of methyl formate, despite the presence of dedicated surveys (e.g., Remijan et al. 2003). In the literature, there are few star-forming regions in which all three isomers have been detected: SgrB2(N)-LMH (Hollis et al. 2001; Belloche et al. 2013; Xue et al. 2019), the two HMCs NGC 6334I MM1 & MM2 (El-Abd et al. 2019), and the low-mass sources IRAS16293-2422B and IRAS16293-2422A (Jørgensen et al. 2012, 2016; Manigand et al. 2020). To better constrain and compare the predictions of the chemical models with observations, more sources with all three isomers detected are needed.

GA was firstly detected outside the GC towards G31 (Beltrán et al. 2009), where Isokoski et al. (2013), Calcutt et al. (2014), and Rivilla et al. (2017) detected MF.

The aim of this work is to present the GUAPOS project¹ and to focus on the simultaneous analysis of the 3 C₂H₄O₂ isomers: GA, MF, and AA. In Sect. 2, we describe the observations and the data reduction process for obtaining the final spectra. In Sect. 3, we analyze the continuum emission and describe the methodology for the spectral analysis of the three isomers of C₂H₄O₂. In Sect. 4, we show the result for MF, AA, and GA. In Sect. 5, we discuss the abundances and the column density ratios among the three isomers and compare them with previous values in the literature and with the predictions of different chemical models to better understand how this COMs are formed in the ISM of star-forming regions. Finally in Sect. 6, we summarize our conclusions.

2. Observations and data reduction

2.1. Observations

The observations were carried out with ALMA during Cycle 5, between 18 January and 7 September 2018 (project 2017.1.00501.S, P.I.: M. T. Beltrán), using 43 antennas. The coordinates of phase center, together with other properties of G31, are given in Table 1. More details about the spectral setup, the baselines, and the flux and phase calibrators are given in Table 2. The survey covers the complete spectral range of ALMA band 3, between 84.05 and 115.91 GHz (~ 32 GHz bandwidth), with a spectral resolution of ~ 0.488 MHz (~ 1.3 – 1.7 km s⁻¹). We used nine correlator configurations (spec) and for each of them, four contiguous basebands of ~ 937 MHz were observed simultaneously. In order to create a single spectrum starting from the 36 spectra of the respective basebands, an overlap in frequency ranging from ~ 7.3 to ~ 29 MHz was chosen for each pair of adjacent basebands. The original angular resolution requested in the proposal ($1''$) was not achieved for two out of nine specs (spec 3 and 4) during the observations between January and March 2018. Although we were granted additional observing time for these two specs, the new observations did not reach the angular resolution required either, achieving only $\sim 1''.2$. We decided to degrade the angular resolution of all the *specs* to the lowest one ($\sim 1''.2$) to have the same spatial resolution for all of them. For all the *specs*, the source used as flux and bandpass calibrator is J1751+0939, while J1851+0035 is the phase calibrator. The uncertainties in the flux calibration are $\sim 5\%$ (from Quality Assessment 2 reports), which is in good agreement with flux uncertainties of other ALMA band 3 calibrators reported in Bonato et al. (2018).

¹ Webpage of the project: <http://www.arcetri.astro.it/~guapos/>

Table 1. Coordinates used for the observations and main parameters of the source G31.

RA(J2000)	Dec(J2000)	v_{LSR}	$d^{(a)}$	$L^{(b)}$	$M^{(c)}$
[h m s]	[° ′ ″]	[km s ⁻¹]	[kpc]	[L_{\odot}]	[M_{\odot}]
18 47 34	-01 12 45	96.5	3.75	4.4×10^4	70

Notes. ^(a)Immer et al. (2019); ^(b)from Osorio et al. (2009), rescaled to a distance of 3.75 kpc; ^(c)from Cesaroni (2019), after rescaling to a distance of 3.75 kpc.

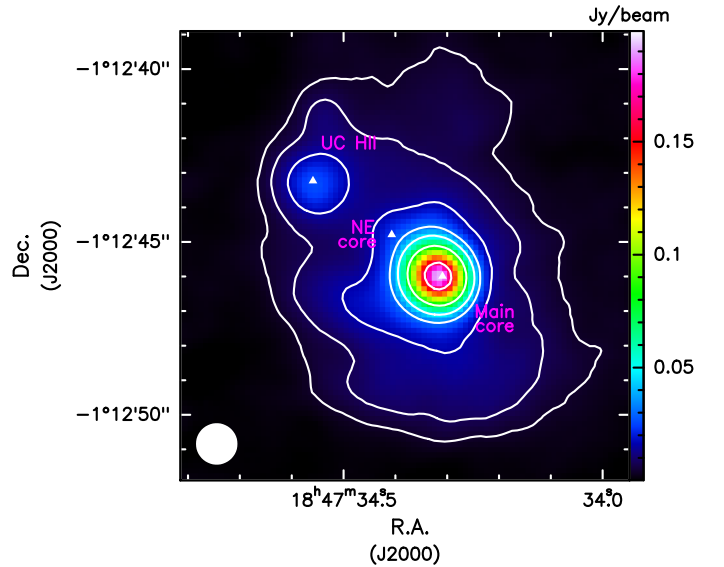


Fig. 2. Continuum map of the HMC G31.41+0.31 and the nearby UC HII region. Contour levels are at 5, 10, 20, 40, 60, 100, and 200 times the value of $\text{rms} = 0.8 \text{ mJy beam}^{-1}$. The three white triangles indicate the coordinates of the Main Core and NE core by Beltrán et al. (2018), and of the UC HII region by Cesaroni et al. (1998). The beam is shown in white in the lower left corner.

The data were calibrated and imaged with Common Astronomy Software Applications² (CASA package, McMullin et al. 2007). The maps were created using a robust parameter of Briggs (1995) set equal to 0 and a restoring synthesized beam of $1''.2 \times 1''.2$. The root mean square (rms) of the maps varies between 0.5 and $1.9 \text{ mJy beam}^{-1}$.

2.2. Continuum determination

Due to the large number of lines in the spectra, it was not possible to find a sufficient number of channels without line emission that could be used to obtain a map of the continuum emission. To overcome this problem we used STATCONT (Sánchez-Monge et al. 2018), a python-based tool designed to determine the continuum emission level in spectral data, particularly for sources with a very rich spectrum. This tool determines the continuum level by using a statistical approach on the intensity distribution of the spectrum and produces a continuum map. We obtained the continuum map (Fig. 2) from the line+continuum map of the spec5 BB1 (see Table 2, ν : 98 499.749–99 436.626 MHz) in order to have the continuum extracted from a frequency close to the center of the frequency range covered by the GUAPOS observations (~ 84 – 116 GHz).

² <https://casa.nrao.edu>

Table 2. Spectral setup of observations.

	Baseband	ν_0 [MHz]	$\Delta\nu$ [MHz]	rms [mJy beam ⁻¹]	Baselines [m]
spec 1	BB 1	84 520.801	0.48824	1.1	15–1397
	BB 2	85 378.837	0.48824	0.8	
	BB 3	86 284.458	0.48824	0.8	
	BB 4	87 190.496	0.48824	0.8	
spec 2	BB 1	88 096.667	0.48823	0.7	15–1397
	BB 2	89 002.644	0.48823	0.6	
	BB 3	89 908.312	0.48823	0.6	
	BB 4	90 814.305	0.48823	0.5	
spec 3	BB 1	91 720.492	0.48821	1.9	15–783
	BB 2	92 626.501	0.48821	1.8	
	BB 3	93 532.124	0.48821	1.9	
	BB 4	94 438.102	0.48821	1.9	
spec 4	BB 1	95 344.296	0.48829	1.2	15–783
	BB 2	96 250.320	0.48829	1.6	
	BB 3	97 155.974	0.48829	1.2	
	BB 4	98 061.998	0.48829	1.0	
spec 5	BB 1	98 968.187	0.48821	1.7	15–783
	BB 2	99 874.167	0.48821	1.7	
	BB 3	100 780.279	0.48821	1.5	
	BB 4	101 686.288	0.48821	1.4	
spec 6	BB 1	102 592.110	0.48821	1.5	15–783
	BB 2	103 498.090	0.48821	1.4	
	BB 3	104 404.324	0.48821	1.5	
	BB 4	105 310.303	0.48821	1.5	
spec 7	BB 1	106 216.143	0.48821	1.4	15–783
	BB 2	107 122.124	0.48821	1.6	
	BB 3	108 028.359	0.48821	1.5	
	BB 4	108 934.339	0.48821	1.4	
spec 8	BB 1	109 840.176	0.48821	1.5	15–783
	BB 2	110 746.157	0.48821	1.1	
	BB 3	111 652.391	0.48821	1.5	
	BB 4	112 558.371	0.48821	1.1	
spec 9	BB 1	113 363.505	0.48821	0.9	15–783
	BB 2	114 170.246	0.48821	1.0	
	BB 3	114 943.241	0.48821	1.3	
	BB 4	115 443.023	0.48821	1.3	

Notes. The table lists: number of spec, number of baseband, central observed frequency of each baseband, ν_0 , spectral resolution in observed frequency $\Delta\nu$, rms of the maps for channel and baselines range in meters.

2.3. Combination of the spectral windows

From each baseband, we extracted the mean spectrum inside an area equal to the beam, centered at the position of the continuum peak of the HMC (see coordinates in Sect. 3.1). This area is smaller than the size of the continuum core (see Sect. 3.1 and Fig. 2). In the Rayleigh-Jeans approximation, the relation between the synthesized beam brightness temperature T_{SB} [K] and the flux I_ν [mJy beam⁻¹] is:

$$T_{\text{SB}} = 1.22 \times 10^3 \frac{I_\nu}{\nu^2 \theta_a \theta_b}, \quad (1)$$

where ν is the frequency in units GHz and θ_a and θ_b are the major and minor axes of the synthesized beam in arcsec. The

discrepancy between the T_{SB} and the physical temperature of a thermalized, beam-filling source emitting the same I_ν , is of the order of $\sim 15\%$ for the channels with only continuum emission, and only of the $\sim 3\%$ for the channels with bright-line emission. Adjacent spectra have been merged to produce a single final spectrum thanks to the overlapping regions between each pair of adjacent basebands. Since jumps up to 3 K are present between partially overlapping adjacent spectra, we adopt the following procedure to create the merged spectrum. We align the spectra from the different basebands using the spectrum at longer λ (spec1 BB1, see Table 2) as reference. In practice, we forced the channels of spec1 BB2 in the overlapping region to match those of spec1 BB1 by subtracting from spec1 BB2 the mean value of the difference between spec1 BB2 and spec1 BB1. We

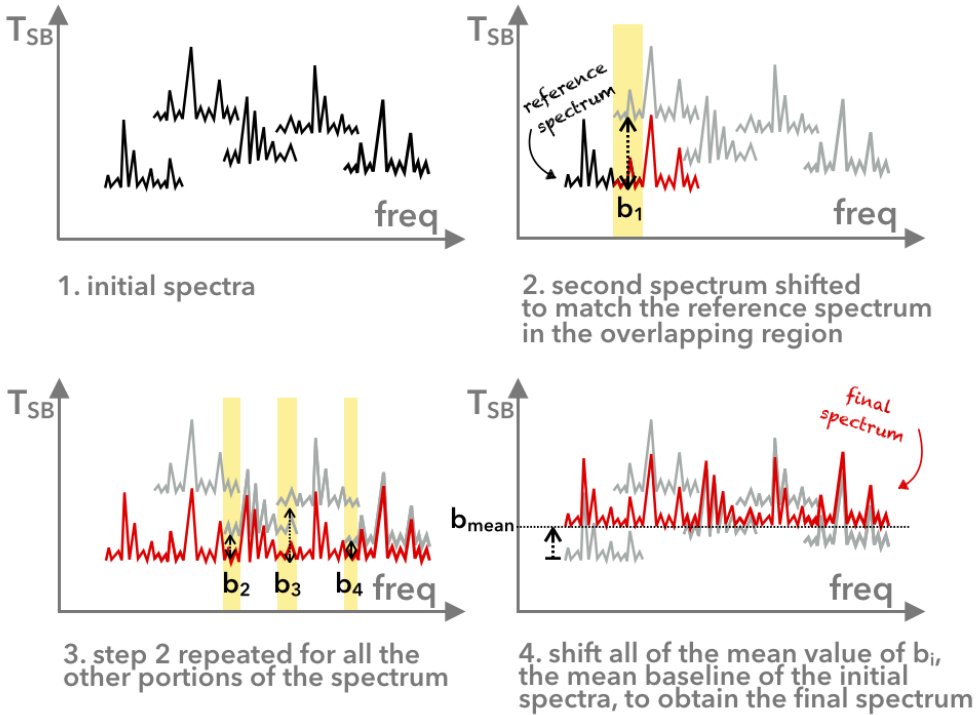


Fig. 3. Graphic scheme of the steps followed to obtain the final spectrum.

define this mean value as b_1 . Then the procedure was repeated for all the other pairs (spec1 BB2 and spec1 BB3, spec1BB3 and spec1 BB4, spec1 BB4, spec2 BB1 etc.) thus obtaining a single spectrum without artificial jumps and the values, b_i , of the baselines of the spectra of all the basebands with respect to the reference one. As a last step, we shifted the final spectrum without jumps by a value equal to the mean of b_i to obtain a final spectrum without jumps with a baseline equal to the mean value of the baselines of the original spectra. Figure 3 summarizes and graphically illustrates the main steps described above. We converted the spectra to rest frequency, assuming a velocity of the source of 96.5 km s^{-1} , and rebinned the spectra to a constant value of the bin width equal to 0.48840 MHz using *Herschel* Interactive Processing Environment³ (HIPE, Ott 2010) since the small differences in the bandwidth of different basebands lead to differences in the channel widths (see Table 2), which span between 0.48837 and 0.48845 MHz (after the conversion to rest frequency). This was done also considering the possibility of analyzing the spectrum with different software since some of the commonly used software to analyze spectra (e.g., CLASS or MADCUBA⁴) need a unique frequency width to import the data.

We remark that the procedure we used did not, in principle, affect the slope of the baseline in each baseband in any way. Considering the artificial nature of the jumps in adjacent spectra, we expect the slope of the baseline not to be shifted considerably from its true value.

The rms at the level of the baseline, calculated as the rms of the differences between the original baseline in each baseband and the baseline in the final spectrum, is 1.2 K ($\sim 12\%$ of the

continuum level, see Sect. 3.1). This 12% error on the continuum level is included as additional error on the parameters derived from the fit to the spectrum. The rms of the spectra has been derived from the rms of the maps (see Table 2) and vary from 7 to 27 mK, for the different basebands, but we consider a conservative value of 27 mK for the entire final spectrum. The fluctuations in the rms of the maps are the result of different atmospheric conditions in different days of observations and of the presence of bright lines in some of the basebands (therefore, a dynamic range effect). Figure 4 shows the total final spectrum.

3. Analysis

3.1. Continuum

The map of the continuum (see Fig. 2) shows the presence of two compact sources. The brightest source is our target, HMC G31.41+0.31, which peaks at RA $18^{\text{h}} 47^{\text{m}} 34.321^{\text{s}}$ Dec $-01^{\circ} 12' 45.977''$ (J2000). Located at a distance of $\sim 4''.5$ in the NE direction, there is a fainter continuum source that peaks at RA $18^{\text{h}} 47^{\text{m}} 34.56^{\text{s}}$ Dec $-01^{\circ} 12' 43.35''$ (J2000), which is consistent with the position of the nearby UC HII region observed by Cesaroni et al. (1998) in the continuum at 1.3 cm with the VLA. The NE core, identified by Beltrán et al. (2018) with high resolution observations, is not resolved in our data at $1''.2$ resolution.

We fitted a 2D Gaussian to the continuum emission of the main core using the task *imfit* of CASA and obtained the source size of the HMC, reported in Table 3. The source is barely resolved, with the deconvolved size of the continuum being $1''.41 \times 1''.22$.

The flux of the continuum inside the area from which we extracted the spectrum is of 0.1 Jy at the mean frequency of 99.0 GHz (see Sect. 2.2). Assuming a dust temperature of 150 K , $\beta = 2.0$, a dust opacity coefficient $k_0 = 0.8 \text{ cm}^2 \text{ g}^{-1}$ at 220 GHz (Ossenkopf & Henning 1994), a mean molecular weight of 2.33, and a gas-to-dust ratio of 100, we have a column density of H_2 is

³ <https://www.cosmos.esa.int/web/herschel/hipe-download>

⁴ Madrid Data Cube Analysis (MADCUBA) is a software developed in the Center of Astrobiology (Madrid) to visualize and analyze data cubes and single spectra (Martín et al. 2019): <https://cab.inta-csic.es/madcuba/>

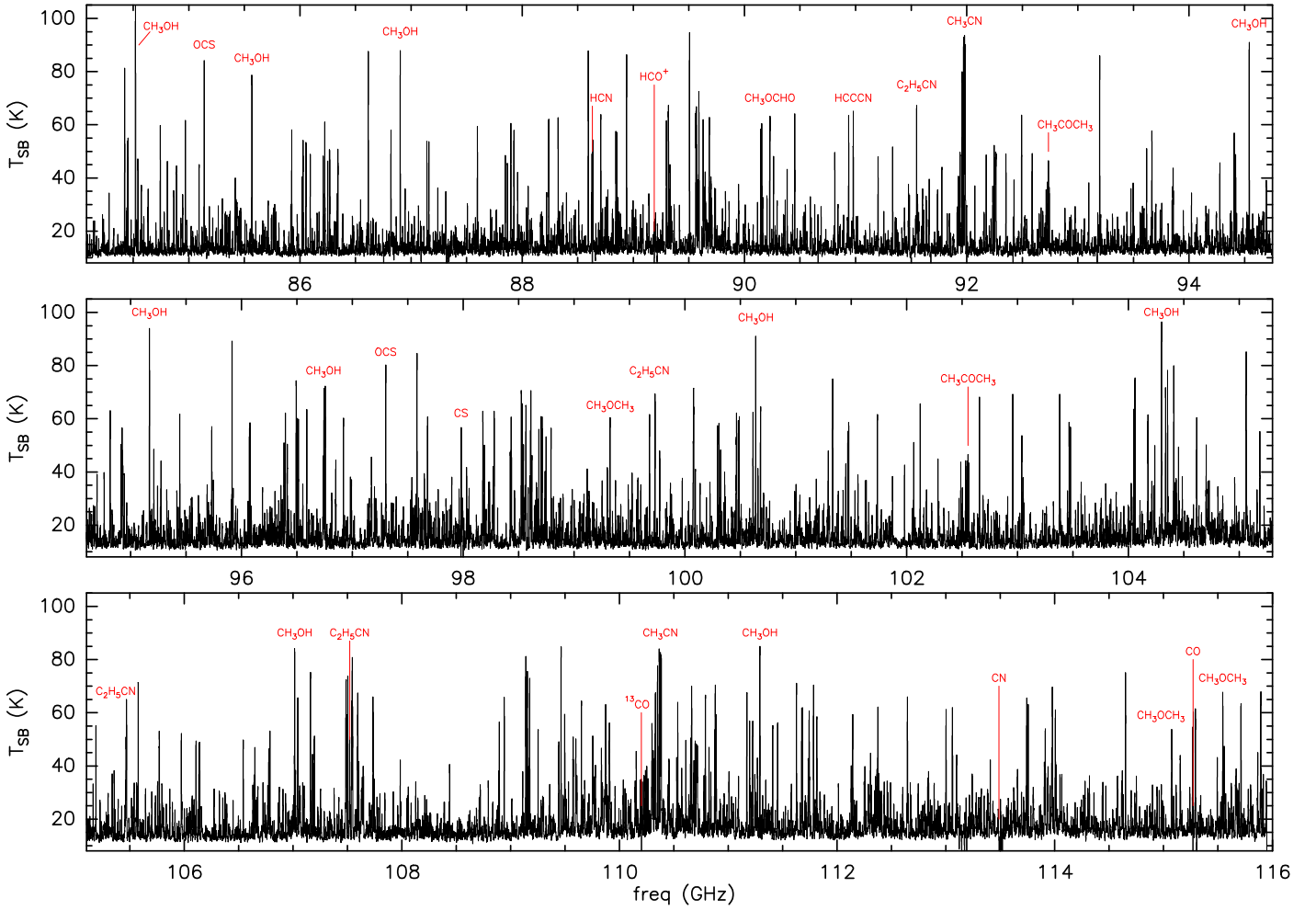


Fig. 4. Full final spectrum from 84 to 116 GHz. In red is reported the name of the molecular species associated to some of the most common or bright lines.

$N(\text{H}_2) = 1.0 \times 10^{25} \text{ cm}^{-2}$. This value is consistent within a factor ~ 4 with the previous estimate by Rivilla et al. (2017).

We have fitted by eye the continuum level of the spectrum assuming a power law,

$$T_{\text{SB}}(\nu) = T_{\text{SB}}(\nu_0)(\nu/\nu_0)^\beta, \quad (2)$$

where ν_0 is the reference frequency, $T_{\text{SB}}(\nu_0)$ is the level of the continuum at the arbitrary reference frequency, and β is the dust emissivity spectral index. The parameters of the continuum are given in Table 4. The dust emissivity spectral index β is related to α , the flux spectral index ($S_\nu \propto \nu^\alpha$), by

$$\alpha = 2 + \beta. \quad (3)$$

The value of $\beta = 0.4$ ($\alpha = 2.4$) is smaller than the expected value for dust thermal emission from small grains ($\beta \in [1; 2]$). However, it is consistent with the thermal emission by dust in the case of the presence of larger grains together with possibly different grain chemical composition (see Beckwith & Sargent 1991; Galametz et al. 2019; Ysard et al. 2019 and references therein). Based on the fluxes at 7 mm and 1.3 cm of the two unresolved radio sources found within the HMC by Cesaroni et al. (2010), the contribution of free-free emission to the continuum flux should vary in a range of a few percent to a maximum value of 8%, assuming a spectral index for the

free-free emission of 0.6 and 1.3 (maximum spectral index found by Cesaroni et al. 2010), respectively. The vast majority of the flux of the continuum is thus due to thermal emission from dust grains, with possibly larger grains already present.

3.2. Spectral analysis

The line identification and the fit to the $\text{C}_2\text{H}_4\text{O}_2$ isomers have been obtained with the software XCLASS⁵ (eXtended CASA Line Analysis Software Suite – Möller et al. 2017), which makes use of the Cologne Database for Molecular Spectroscopy⁶ (CDMS, Müller et al. 2001, 2005) and Jet Propulsion Laboratory⁷ (JPL, Pickett et al. 1998) catalogs, via the Virtual Atomic and Molecular Data Centre (VAMDC, Endres et al. 2016). The spectroscopic data of acetic acid, from the work of Ilyushin et al. (2013), were introduced into the CDMS in May 2019 to perform the analysis presented in this work. A complete documentation of the spectroscopic works the entries of the catalogs for MF, AA, and GA are based on, is given in Appendix A. For each molecular species the software computes a synthetic spectrum assuming LTE (the software also allows non-LTE analysis) using as possible free parameters the excitation temperature, T_{ex} , the

⁵ <https://xclass.astro.uni-koeln.de>

⁶ <https://cdms.astro.uni-koeln.de>

⁷ <https://spec.jpl.nasa.gov>

Table 3. Parameter of the best fit for a 2D Gaussian model to the map of the continuum and the maps of MF, AA and GA in Figs. 8 and 9, as well as the difference between the mean diameter of each map with the mean diameter of the mean map.

	ν_0 [GHz]	E_u [K]	RA (J2000) [18h 47m s]	Dec (J2000) [−01° 12′ ″]	$\theta_{\text{conv}} \times \theta_{\text{conv}}$ [″ × ″]	PA [°]	$\bar{\theta} - \bar{\theta}_{\text{mean}}$ [″]
Continuum			34.321	45.976	1.85×1.71	29	
CH ₃ OCHO	88.2207	205	34.316	46.0414	1.61×1.56	44	−0.14
	88.7707	208	34.316	46.030	1.63×1.60	27	−0.11
	98.8153	215	34.320	45.949	1.65×1.59	75	−0.10
	103.4787	25	34.318	46.069	1.84×1.81	60	+0.10
	110.7765	216	34.318	46.053	1.60×1.58	63	−0.13
	111.4084	37	34.318	46.087	1.81×1.79	68	+0.08
	111.4533	37	34.317	46.094	1.83×1.82	83	+0.11
	111.6822	28	34.317	46.101	1.82×1.81	65	+0.10
Mean	–	–	34.318	46.054	1.73×1.71	55	
CH ₃ COOH	85.319/.322	181/207	34.320	46.098	1.53×1.39	6	+0.06
	85.6331	229	34.320	46.078	1.48×1.31	8	−0.00
	90.2462	20	34.313	46.107	1.46×1.43	19	+0.05
	94.4995	58	34.311	46.018	1.37×1.35	109	−0.03
	100.8554	26	34.314	46.017	1.43×1.36	58	−0.00
	100.8975	25	34.314	46.028	1.45×1.37	50	+0.01
	104.0780/.0786	80/266	34.317	46.049	1.33×1.31	92	−0.07
	114.6380	75	34.313	46.068	1.39×1.35	179	−0.02
Mean	–	–	34.315	46.057	1.42×1.37	19	
CH ₂ OHCHO	88.5304	30	34.304	46.163	1.43×1.37	19	−0.05
	88.6912	49	34.306	46.124	1.52×1.46	167	+0.04
	93.0485	192	34.310	46.054	1.44×1.36	45	−0.05
	93.0527	23	34.309	46.066	1.50×1.45	42	+0.03
	95.7562	178	34.305	46.103	1.44×1.37	73	−0.04
	102.5729	64	34.314	46.049	1.51×1.44	79	+0.03
	102.6144	104	34.310	46.098	1.41×1.35	77	−0.07
	103.6680	32	34.310	46.100	1.56×1.51	71	+0.09
Mean	–	–	34.309	46.094	1.47×1.42	56	

Notes. $\bar{\theta} - \bar{\theta}_{\text{mean}}$ is calculated as $\sqrt{\theta_1 \theta_2} - \sqrt{\theta_1^{\text{mean}} \theta_2^{\text{mean}}}$.

Table 4. Parameters of the continuum baseline.

ν_0 [GHz]	$T_{\text{SB}}(\nu_0)$ [K]	β
84.079	10.4	0.4

Notes. The table lists: reference frequency ν_0 , level of the continuum at the reference frequency $T_{\text{SB}}(\nu_0)$ and spectral index β .

total column density, N_{tot} , of the molecule, the line FWHM, $\Delta\nu$, the velocity of the source, V_{LSR} , and the angular diameter of the source θ_s . The LTE assumption is justified by the high density of hot cores; in fact, from the column density of H₂ derived in Sect. 3.1, a rough estimate of the volume density of the core is $n \sim 10^8 \text{ cm}^{-3}$. The line analysis can be done without subtracting the continuum, which can be simulated simultaneously with the line emission. The fitting procedure varies the free parameters to minimize the χ^2 between the observed spectrum and the synthetic one. It is possible to select the frequency ranges for the fit, which allows us to limit the fit to a selected number of transitions. In order to estimate the degree of contamination in the observed transitions of AA, GA, and MF, due to emission

by other species, we performed a preliminary line identification across the whole spectrum. Starting from the molecules responsible for the brightest lines, we carried out an investigation to identify as many species that can potentially contribute to the observed spectrum as possible. For the species identified in this way, we adjusted the parameters of the synthetic spectra until a reasonable fit (by visual inspection) was obtained. This preliminary identification was used to identify the transitions of the isomers of C₂H₄O₂, and among them, the most unblended, with the intention to use them in the XCLASS fitting procedure.

We followed three rules to identify a transition of one of the 3 C₂H₄O₂ isomers: (i) the intensity of the line should be higher than $3 \times \text{rms}$; (ii) blending of a line of another species, with a separated peak, should start at a distance larger than $FWHM/2$ from the peak; (iii) in the case of blending with separation $< FWHM/2$, the intensity of the line of the other species should be $< 15\%$ of the peak. In the case of AA, whose lines are more affected by blending, we included also lines with one of the two wings blended up to 60% of the peak (with respect to the baseline). The MF spectrum presents 196 identified lines, with brightness temperatures up to ~ 70 K. These include both ground state and vibrationally excited $\nu_{18} = 1$ transitions, with upper state energies (E_u) between ~ 10 and ~ 400 K. For GA we

Table 5. Parameter of the best fit and abundances for MF, AA, and GA.

	$FWHM$ [km s ⁻¹]	V_{LRS} [km s ⁻¹]	T_{ex} [K]	N_{tot} [10 ¹⁷ cm ⁻²]	X
CH ₃ OCHO	6.8	1.1	221 ± 27	20 ± 4	(2.0 ± 0.6) × 10 ⁻⁷
CH ₃ COOH	7.8	0.0	299 ± 42	8.4 ± 1.4	(8 ± 3) × 10 ⁻⁸
CH ₂ OHCHO	8.8	0.0	128 ± 17	0.50 ± 0.09	(5.0 ± 1.4) × 10 ⁻⁹

Notes. N_{tot} : column density inside 1''2 beam; X : abundances inside a beam calculated as N_{tot}/N_{H_2} , where $N_{H_2} = (1.0 \pm 0.2) \times 10^{25}$ cm⁻². FWHM and V_{LRS} parameters were kept fixed in the fitting procedure.

detected 20 lines, with brightness temperatures up to ~28 K. All the transitions are in the ground vibrational state and have energies E_u between ~25 and ~190 K. Lastly, we identified 26 lines of AA, with brightness temperatures up to ~23 K. These are ground state and vibrationally excited $v_{18} = 1$ transitions, with energies E_u between ~25 and ~400 K. To compare the best fit physical parameters of the three species in a consistent way, we performed the fit by selecting the most unblended transitions with energies $E_u < 250$ K. This energy constraint was due to GA not presenting identified unblended transitions of higher upper state energy. Beltrán et al. (2018) found a decrease in T_{ex} of methyl formate with distance from the core center from $T_{ex} \sim 497$ K at the inner radius of 0''22 to $T_{ex} \sim 110$ K at a radius of 1''32. From these data, the average value of T_{ex} of MF inside the main core is ~250 K. Despite the presence of this gradient in T_{ex} , we decided to perform a single temperature component analysis, to reduce the number of free parameters as much as possible. For each molecule we selected the lines for the fitting procedure choosing the most unblended transitions and balancing the presence of low E_u and high E_u transition, to avoid possible biases. For MF we selected 25 lines, for GA 12 and for AA 14. We calculated the optical depth τ_0 at the center of all the identified lines (see Appendix B) using the values listed in Table 5 and we find that all transitions are optically thin ($\tau_0 \ll 1$). The identified transitions of the three molecules are listed in Tables C.1–C.3, together with τ_0 . The transitions selected for the XCLASS fit are listed in boldface. As only exception to the $E_u < 250$ K constraint, we have included in the fit the transition 24(14, 11)A2 → 24(13, 12)A1 of AA, which has $E_u = 266$ K, because it is blended with other transitions of the same species with $E_u = 80$ K. During the fit, the continuum, from Eq. (2), the FWHM and V_{LRS} of the selected species were kept fixed to the values found by a preliminary fit (in which single transitions were selected and all the parameters were free to vary), while T_{ex} and N_{tot} were free to vary.

4. Results

In the following, we analyze the best fit physical parameters obtained from the line fitting performed with XCLASS for the three different isomers. The results of the fit are given in Table 5, while in Figs. 5–7 the transitions used to constrain the fit for MF, AA, and GA are reported, with the synthetic spectra obtained using the best fit parameters overlaid on the data. The total spectrum with the synthetic spectra of the three isomers obtained using the best fit parameters is given in Appendix D. The errors on T_{ex} and N_{tot} take into account a 12% error, from the rms of the continuum baseline determination (see Sect. 2.3). To determine the abundances of the molecules we used the value of the column densities inside the 1''2 beam, divided by the value

of column density of H₂ found in Sect. 3.1, extracted from the same area, considering a 20% uncertainty.

Figure 8 shows the integrated maps of 8 of the most unblended transitions of MF, AA, and GA with different values of upper state energy E_u (from ~25 to ~200 K). As seen in this figure, the emission of MF, AA, and GA comes only from the HMC and has a nearly circular shape in all the emission maps. To obtain a mean value of diameter of the emitting region of each molecule, the eight maps of a given molecule are normalized with respect to the peak intensity and averaged together. The result is shown in Fig. 9. The *imfit* task from CASA was used to fit a 2D Gaussian to these maps and thus obtain an estimate of the angular diameters of the emission, reported in Table 3. The column densities of the three isomers have been derived assuming a beam filling factor of 1.

The decrease in the normalized flux (F/F_{peak}) in the emission maps, as a function of the radius from the peak R , is shown in Fig. 10 for the mean maps (bold lines) and for the eight selected transitions (indicated in Table 3) at different values of E_u . These data were obtained by averaging the normalized flux inside circular rings centered on the mean map peak. As shown in Fig. 10, the emission of MF is the most extended among the three isomers, while AA is the most compact, even if with a small difference with respect to GA. In addition, we performed the fit of AA and MF including also the most unblended transitions of high-energy ($E_u > 250$ K) and discuss the results in Sect. 4.4.

4.1. Methyl formate

For methyl formate, we found an excitation temperature of $T_{ex} = 221 \pm 27$ K, a column density of $N_{tot} = (2.0 \pm 0.4) \times 10^{18}$ cm⁻², and an abundance $X = (2.0 \pm 0.6) \times 10^{-7}$. The T_{ex} is consistent with 250 K, the mean temperature of MF in the main core calculated using the temperature gradient found by Beltrán et al. (2018). From the total spectrum in Appendix D, we see that there are transitions of MF (not used for the fit) that are under-reproduced by more than 20% by the best fit synthetic spectrum. About a third of these transitions are indeed blended with other molecular species, such as NS, CH₃CN, ¹³CH₃CN, C₂H₅CN, and more. The others are mostly low-energy transitions, which would be better reproduced by a lower temperature component of the fit. However, a two-temperature component fit does not properly converge and does not significantly improve the fit, either in terms of χ^2 or of visual comparison with the observed spectrum. These transitions will likely be better reproduced by a fit that takes into account the temperature gradient observed by the previous study of Beltrán et al. (2018). However, this is beyond the scope of this paper. In addition, we cannot exclude that some of these transitions may also be blended with

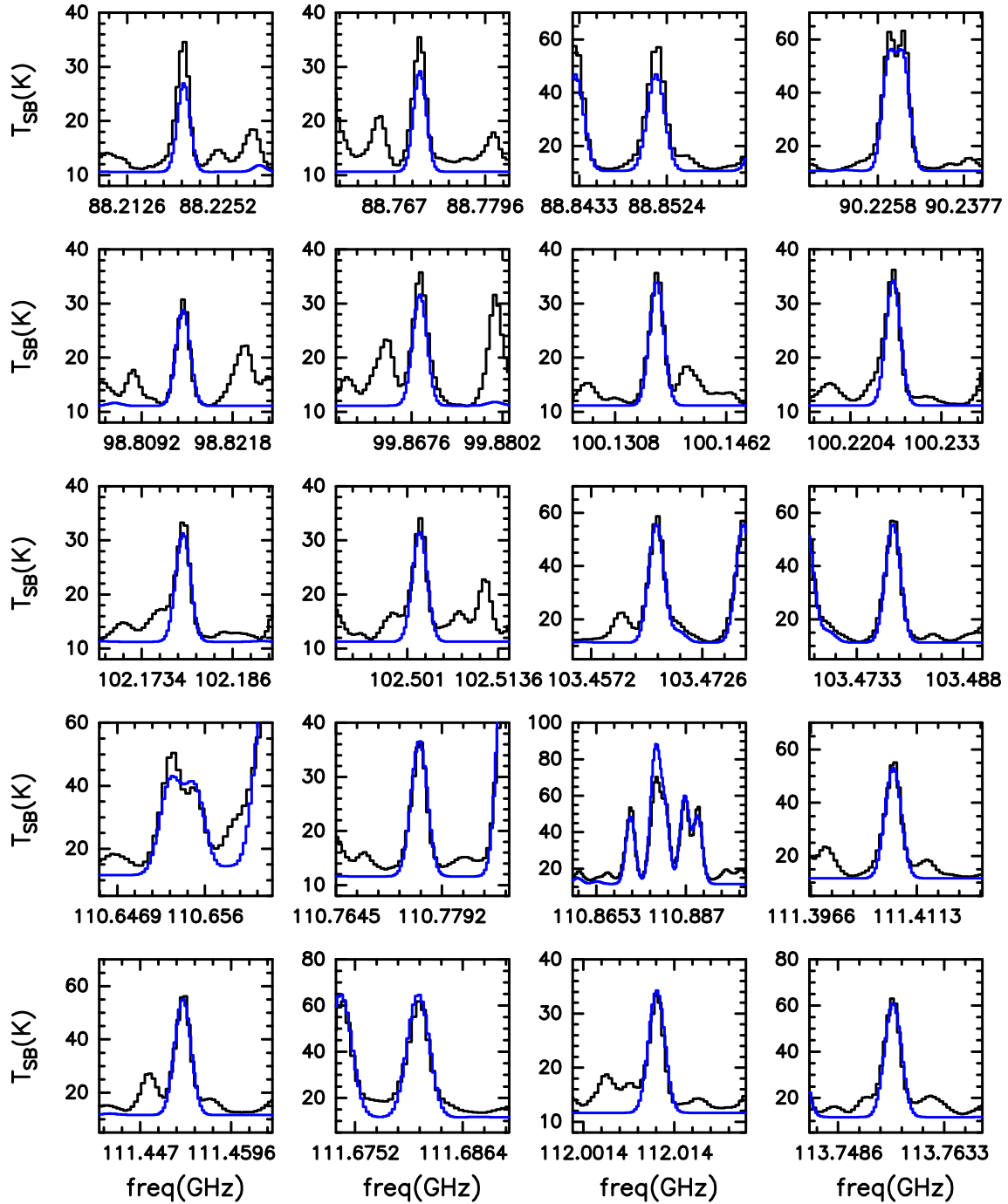


Fig. 5. Transitions used to fit methyl formate, with the LTE synthetic spectra simulated using the best fit parameters given in Table 5 in blue.

still unidentified molecular species since the preliminary line identification does not explain the totality of the transitions in the spectrum.

The mean value of column density of MF inside $1''.1$ from the data of Beltrán et al. (2018) is of $\sim 1.5 \times 10^{18} \text{ cm}^{-2}$, which is consistent with our results. The column density derived by us differs by only a factor 2.5 from the value of $5 \times 10^{18} \text{ cm}^{-2}$ found by Rivilla et al. (2017) from IRAM 30m data, considering a source size of $1''.7$, and is also consistent with the estimate of $3.4 \times 10^{18} \text{ cm}^{-2}$ of Calcutt et al. (2014), when rescaled to the same source size of this work. The distance between the peak of the continuum emission and the MF emission is only $0''.09$, which is smaller than the pixel size ($0''.15$).

4.2. Acetic acid

The best-fit parameters obtained with XCLASS for acetic acid are $T_{\text{ex}} = 299 \pm 42 \text{ K}$, $N_{\text{tot}} = (8.4 \pm 1.4) \times 10^{17} \text{ cm}^{-2}$ and an abundance of $X = (8 \pm 3) \times 10^{-8}$. The temperature is consistent within the errors with the mean temperature in the main core of $\sim 250 \text{ K}$ found from the relationship of $T_{\text{ex}}(\text{MF})$ given in Beltrán et al. (2018). Nevertheless, the values of T_{ex} for the three isomers cover a large range from ~ 130 to $\sim 300 \text{ K}$ (see Table 5). For this reason, to understand if the value of T_{ex} of AA might be overestimated, we performed the fit to AA for a grid of fixed temperatures ranging from 100 K to 300 K . The only free parameter was the column density, N_{tot} . The results are given in Table 6 and shown

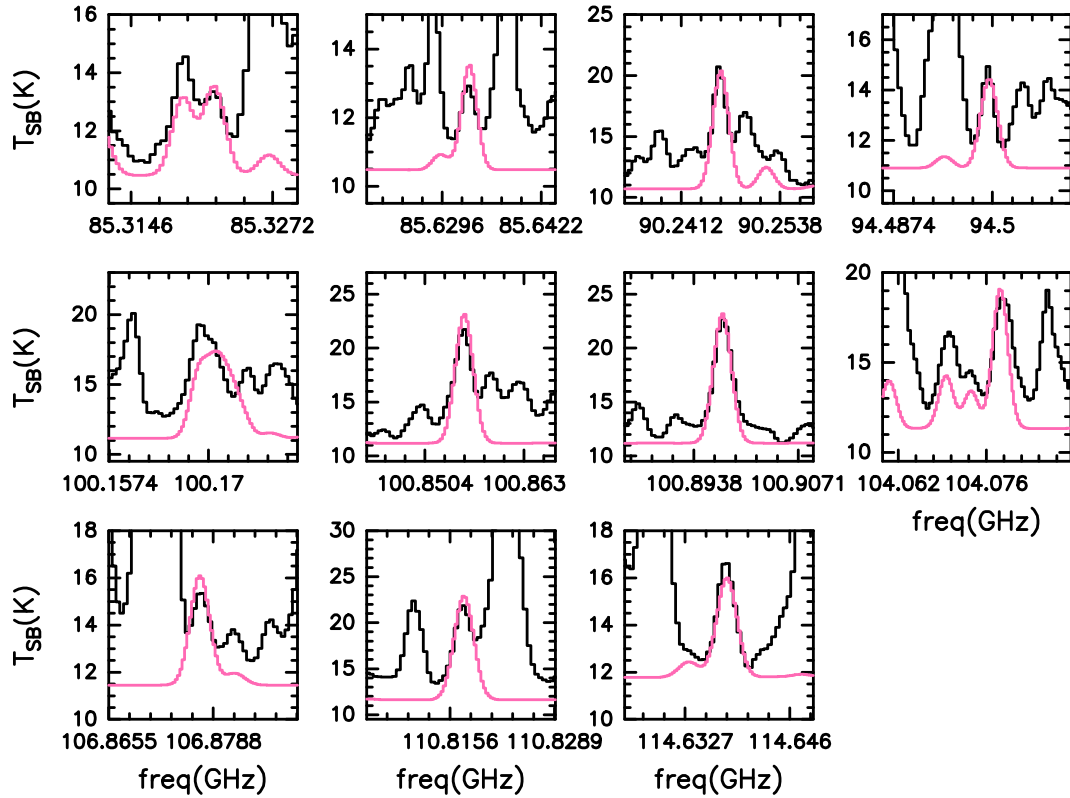


Fig. 6. Transitions used to fit acetic acid, with the LTE synthetic spectrum simulated using the best fit parameters given in Table 5 in pink.

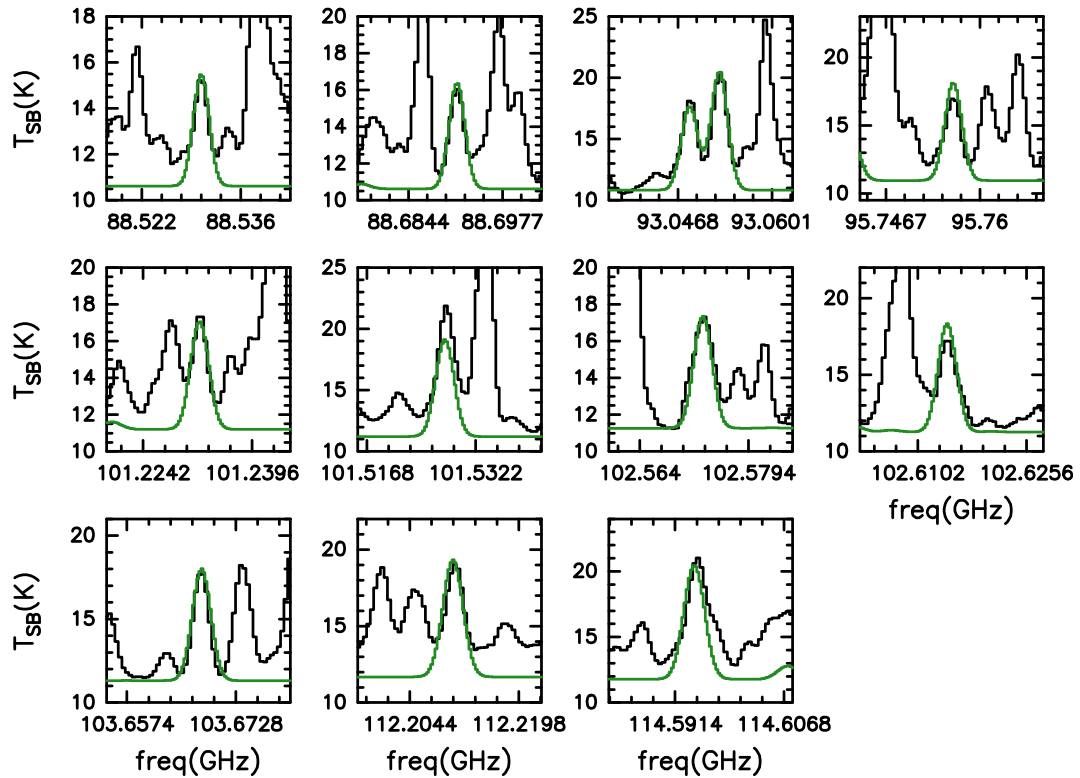


Fig. 7. Transitions used to fit glycolaldehyde, with the LTE synthetic spectra simulated using the best fit parameters given in Table 5 in green.

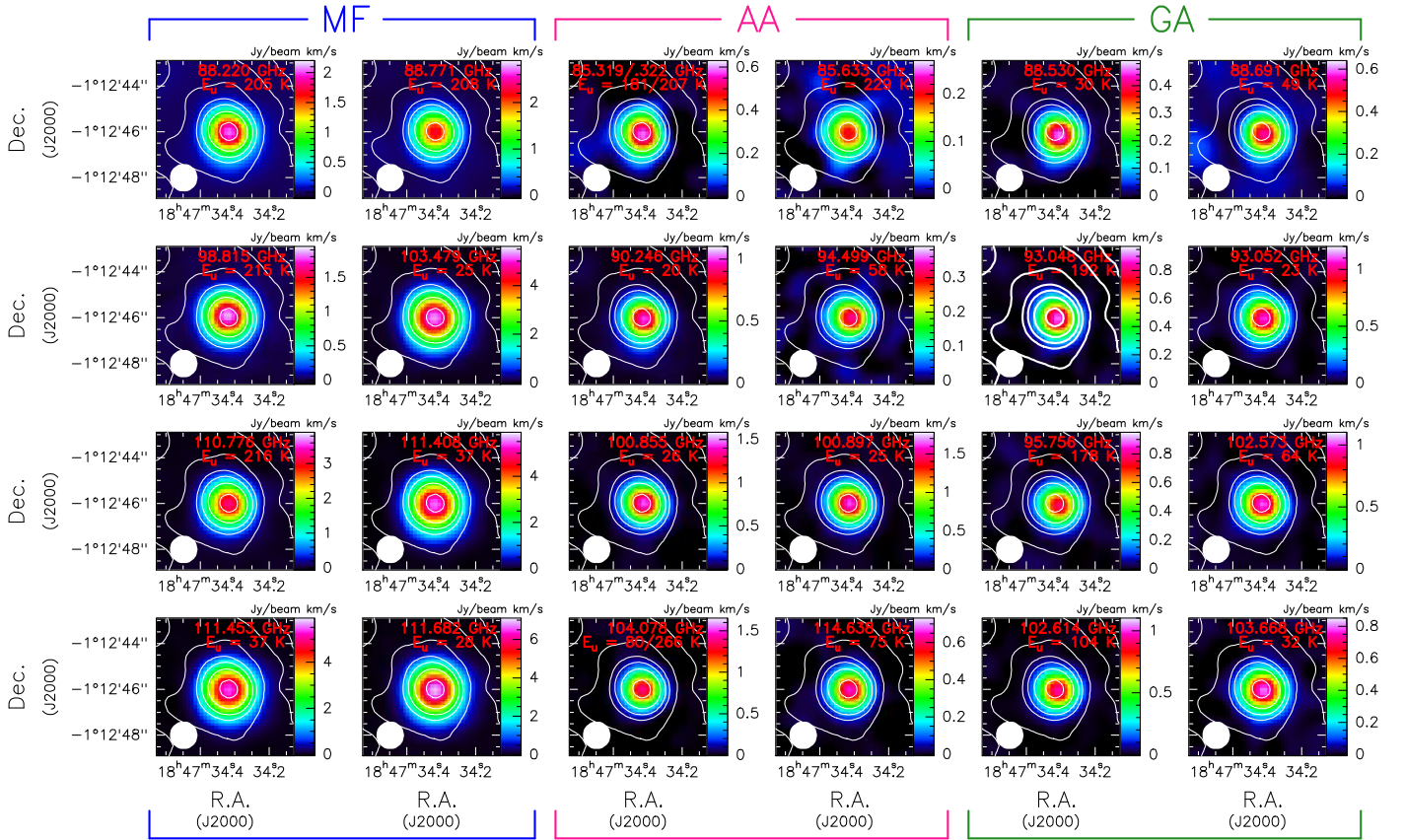


Fig. 8. Maps of integrated emission of 8 transitions with different E_u (from ~ 25 K to ~ 200 K) for MF (left panels), AA (middle panels) and GA (right panels) in color scale. Continuum emission is shown in white contours, as in Fig. 2, and the beam is indicated in the left-bottom corner of the maps.

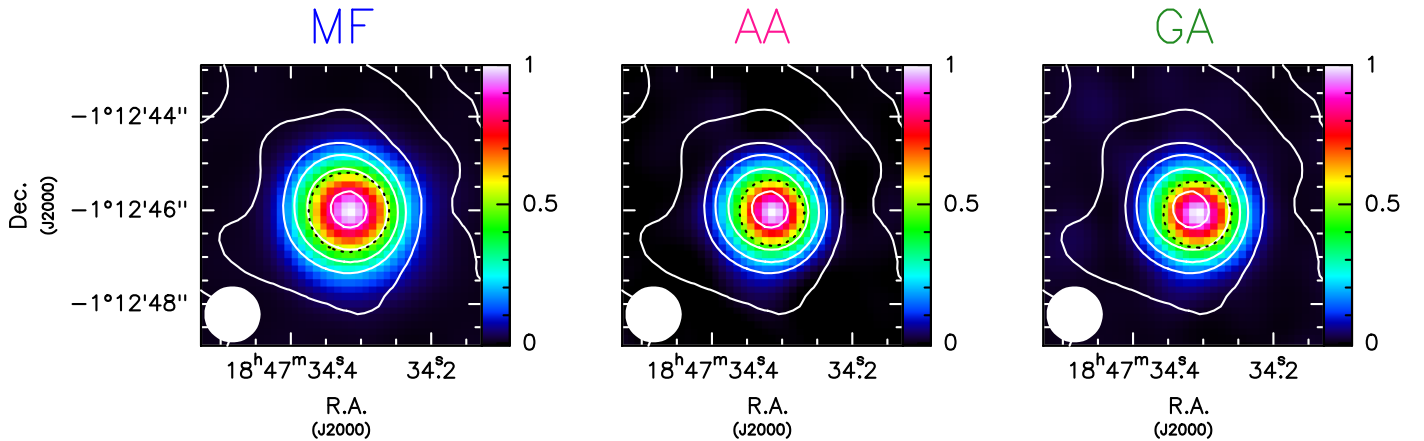


Fig. 9. Mean map for each molecule, obtained from the average of the 8 maps at different E_u in Fig. 8, rescaled to a peak of intensity 1 before averaging. The black dashed line is the contour where the emission reaches the half the peak intensity.

in Fig. 11. From the plots it is possible to exclude $T_{\text{ex}} \leq 150$ K, but it is also clear that for seven transitions out of 14, there is a complete degeneracy between T_{ex} and N_{tot} , for $T_{\text{ex}} \geq 200$ K. From this plot, we conclude that the excitation temperature of AA can be in the range 200–299 K (i.e., the value of best fit), leading to a range in column density of $4.3\text{--}8.4 \times 10^{17} \text{ cm}^{-2}$, with an uncertainty on N_{tot} of a factor ~ 2 . This implies a range in abundance X between 4.3×10^{-8} and 8×10^{-8} . The distance between the peak of the continuum emission and the AA emission is of $0''.12$, slightly larger with respect to MF but still smaller than the pixel size ($0''.15$).

4.3. Glycolaldehyde

For GA, we found an excitation temperature of $T_{\text{ex}} = 128 \pm 17$ K, the lowest value for the three isomers, a column density of $N_{\text{tot}} = (5.0 \pm 0.7) \times 10^{16} \text{ cm}^{-2}$, and an abundance $X = (5.0 \pm 1.4) \times 10^{-9}$. The low T_{ex} can be due to the small number of detected transitions with $E_u > 150$ K (see Table B.3), and consequently the presence of only two high-energy transitions used for the fit, with $E_u = 192$ and 178 K, respectively. To understand if a higher T_{ex} is compatible with the observed spectra, we performed the fit of GA for fixed values of T_{ex} in

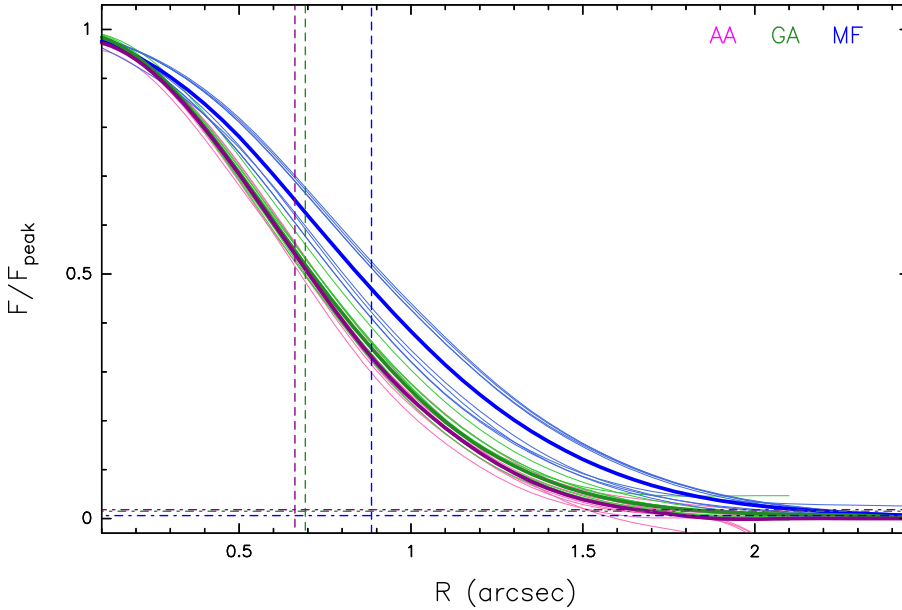


Fig. 10. Relative flux of emission maps shown in Figs. 8 (thin lines) and 9 (bold lines) as a function of R . The dashed vertical lines indicate the mean value between the two semi-axes of the elliptic 2D fit of the mean maps, calculated from the values given in Table 3. At these radii the emission of the mean maps differs slightly from 0.5 due to the fact that the fluxes plotted are calculated in circular rings, for simplicity. The horizontal dashed lines show the $3 \times \text{rms}$ level for the mean emission maps.

Table 6. N_{tot} of the best fit for AA given a grid of fixed temperatures, with $FWHM = 7.8 \text{ km s}^{-1}$ and $V_{\text{LRS}} = 0.0 \text{ km s}^{-1}$.

	T_{ex} [K]	N_{tot} [10^{17} cm^{-2}]
CH ₃ COOH	100	1.3
	150	2.6
	200	4.3
	250	6.3
	300	8.5

Table 7. N_{tot} of the best fit for GA given a grid of fixed temperatures, with $FWHM = 8.8 \text{ km s}^{-1}$ and $V_{\text{LRS}} = 0.0 \text{ km s}^{-1}$.

	T_{ex} [K]	N_{tot} [10^{17} cm^{-2}]
CH ₂ OHCHO	150	0.6
	200	0.9
	250	1.3
	300	1.7

the range 150 to 300 K, varying only the column density. The results of the fits are given in Table 7 and shown in Fig. 12. This figure shows that the transitions that are most sensitive to the value of T_{ex} are those at frequency 93.048, 93.053 and 95.756 GHz. From the comparison of the spectra simulated and observed at these frequencies (Figs. 7 and 12), the best fit seems to be compatible with $T_{\text{ex}} \leq 150 \text{ K}$, confirming the results of the XCLASS fit. The values of the column density found in G31 by Rivilla et al. (2017) and Calcutt et al. (2014) differ by a factor ~ 3.5 and ~ 0.8 , respectively. The results for GA show the largest shift in position with respect to the peak of the continuum emission, with a distance of $0''.22$, which is still small compared to the angular resolution of the data.

4.4. Methyl formate and acetic acid fit including $E_u > 250 \text{ K}$

In Sects. 4.1 and 4.2, we reported the fit of MF and AA while taking into account only the most unblended transitions with $E_u < 250 \text{ K}$. We imposed this restriction to trace the same gas, with the same physical conditions, for all the three isomers. Thus, we did not include high energy transitions (higher than the highest energy transition of GA available) that are, in principle, possibly emitted by a warmer gas. For completeness, we present the fit of AA and MF, including also the most unblended transitions of energy $E_u > 250 \text{ K}$. We selected four transitions for AA ranging from 260 to 510 K, and 7 for MF ranging from 250 to 420 K, to add to the previously selected transitions for

the fit. These new transitions are indicated by an asterisk in Tables C.1 and C.2.

The results of the new fit for AA are an excitation temperature $T_{\text{ex}} = 290 \pm 40 \text{ K}$ and a column density of $N_{\text{tot}} = (8.0 \pm 1.2) \times 10^{17} \text{ cm}^{-2}$, whereas for MF, we find $T_{\text{ex}} = 229 \pm 28 \text{ K}$ and a column density of $N_{\text{tot}} = (2.1 \pm 0.3) \times 10^{18} \text{ cm}^{-2}$. Based on a comparison with the values in Table 5, the results are very consistent within the errors. This indicates that for MF and AA in G31, even the highest energy transitions have the same excitation conditions as lower energy transitions and, thus, they are emitted by gas with the same physical conditions.

5. Discussion

5.1. Comparison with other sources

Figure 13 shows the abundances with respect to H_2 (upper panel) and the column density ratios (lower panel) of the three isomers of $\text{C}_2\text{H}_4\text{O}_2$ in G31 and in the sources in literature for which all three isomers have been observed and at least two of them detected. The upper panel shows a very restricted sample, with respect to the lower panel, because it was for only four sources that we found an estimate of $N(\text{H}_2)$ on an angular scale comparable to the size of the region over which the column densities of the three isomers have been derived. The column density ratios, the abundances and the $N(\text{H}_2)$ adopted, and the corresponding references are given in Table 8. Here, G31 has the highest values of abundance of MF and AA among the four sources shown, with

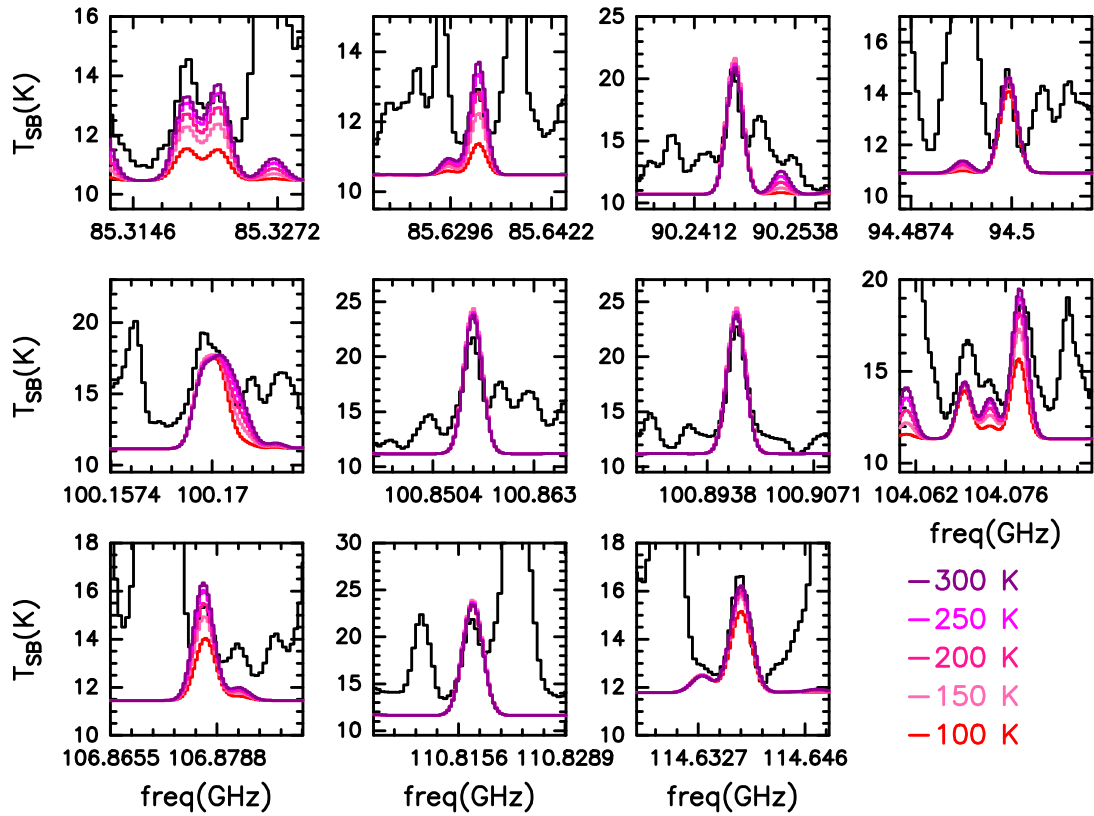


Fig. 11. Transitions used to fit acetic acid, with the LTE synthetic spectra simulated assuming different excitation temperatures, and leaving as free parameters only the column density. The value used are given in Table 6.

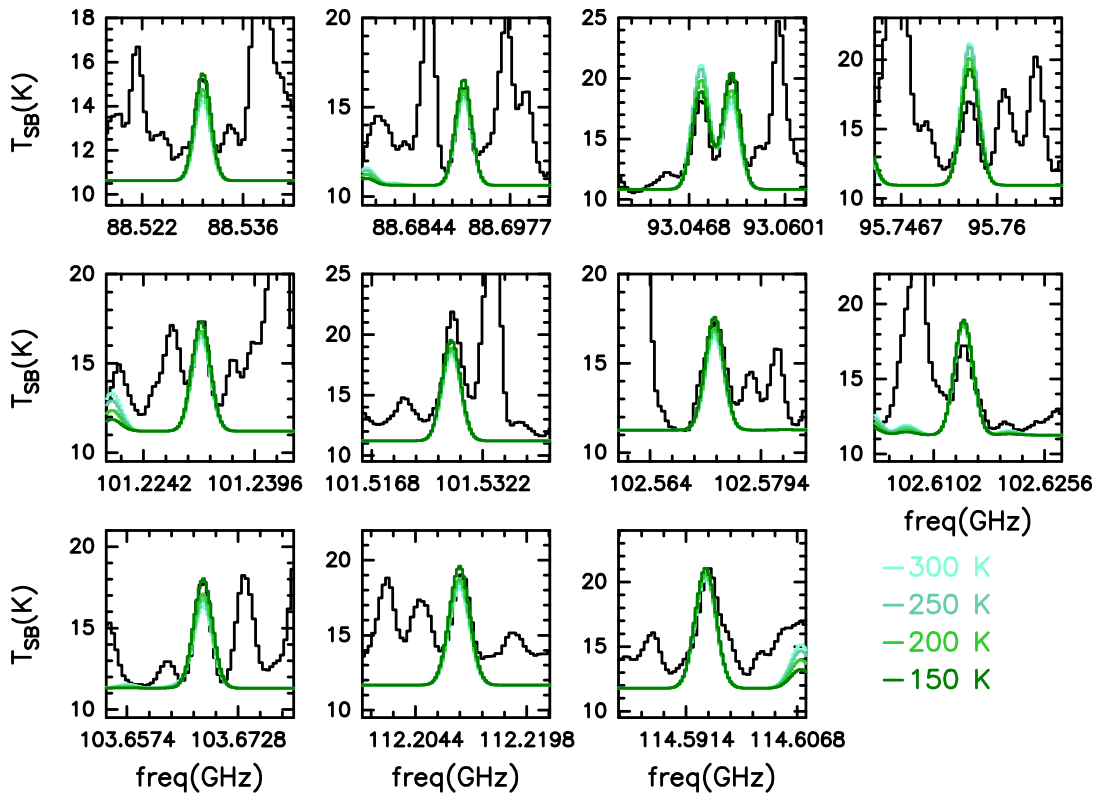


Fig. 12. Transitions used to fit glycolaldehyde, with the LTE synthetic spectra simulated assuming different excitation temperatures, and leaving as free parameters only the column density. The values we used are given in Table 7.

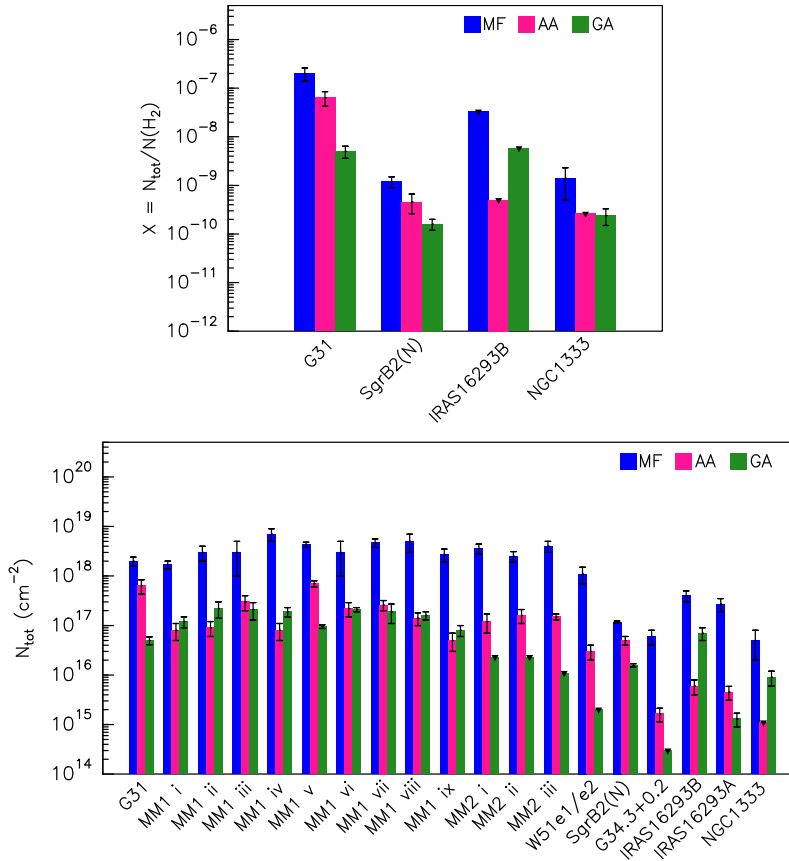


Fig. 13. *Upper panel:* abundances wrt H_2 for the sources with an estimate in Table 8. For G31, the values of $X(AA)$ given in the figure have been calculated from the mean value of the observed range in column density $4.3 \times 10^{17} - 8.4 \times 10^{17} \text{ cm}^{-2}$, while the error bars cover the entire range. For IRAS 16293 B, the value of $N(H_2)$ is a lower limit, leading to upper limits on the abundances, but not affecting the ratios between abundances. *Lower panel:* ratio of the column densities of the three isomers of $C_2H_4O_2$ for the sources listed in Table 8. For G31, the values of MF/AA and AA/GA have been calculated using the mean value of the observed range $4.3 \times 10^{17} - 8.4 \times 10^{17} \text{ cm}^{-2}$ for AA, while the error bars cover the ratios given by entire range. The sources labelled as MM1 i-viii and MM2 i-iii refer to NGC 6334I MM1, and NGC 6334I MM2, while the source labelled as NGC 1333 is NGC 1333 IRAS 4A.

a difference of one and two orders of magnitude, respectively, from the second-highest values in the sample. The abundances in G31 are calculated using $N(H_2) = (1.0 \pm 0.2) \times 10^{25} \text{ cm}^{-2}$ (see Sect. 3.1). Even considering this value as a lower limit – Rivilla et al. (2017) found a value four times larger – and adopting a value of one order of magnitude higher, the abundance of AA would still be the largest in the sample, at least one order of magnitude above the others. From the upper panel, a unique behavior for the three isomers in both high-mass and low-mass sources does not seem to exist. This is confirmed by the column density ratios in the extended sample shown in the lower panel. For IRAS 16293-2422A, Rivilla et al. (2019) found a column density of GA towards the peak one order of magnitude larger than the one in Table 8 and in Fig. 13 by Manigand et al. (2020). Manigand et al. (2020) selected an off-peak position to extract the spectrum, shifted by an amount comparable (by eye) with the emission size of AA and GA, thus possibly leading to underestimated values of these column densities.

The lower panel in Fig. 13 confirms the large fraction of acetic acid in G31, with respect to the extended sample. In fact, in only three sources (four including the lower limit in NGC 1333 IRAS 4A) $MF/AA < 10$: G31, SgrB2(N), and the position “v” in NGC 6334I MM1. The column density ratios of the isomers are plotted in Fig. 14, which are useful for the search for possible correlations. The first plot (upper-left) does not show a clear correlation between MF/AA and MF/GA, however, despite the large scatter, a hint of the dual distribution that was first seen by El-Abd et al. (2019) in the space $N_{MF} - N_{GA}$ is still visible. On the other hand, there seems to be an anti-correlation (Pearson’s $\chi^2 = 0.73$) between MF/AA and AA/GA (upper-right), where we can also possibly see the same dual distribution – although not as clearly – and a faint correlation (Pearson’s $\chi^2 = 0.51$)

between MF/GA and AA/GA (lower panel), even though with some spread. The bimodal distribution is not driven by a difference between high-mass sources and low-mass sources, but the origin must be related to particular physical or chemical conditions inside some sources (both high-mass and low-mass), as already stated by El-Abd et al. (2019).

Lastly, from our work we derived the size of the emission of the three isomers, confirming a difference in the extension and showing as AA and GA are more compact than MF. This result is consistent with what shown in the maps towards other sources by Remijan et al. (2002), Jørgensen et al. (2016), Xue et al. (2019), El-Abd et al. (2019) and Manigand et al. (2020). Towards G31, AA shows the most compact emission among the three isomers.

5.2. Comparison with chemical models and laboratory studies

For several years, the only efficient chemical pathways leading to the formation of the three isomers of $C_2H_4O_2$ were found to be on the surface of dust grains (Garrod & Herbst 2006; Garrod et al. 2008; Beltrán et al. 2009; Laas et al. 2011; Woods et al. 2012, 2013; Garrod 2013; Coutens et al. 2018; Rivilla et al. 2019). When temperatures are sufficiently high, radicals have enough energy to diffuse across the surface and react to form COMs. The main routes on warm dust grains for the formation of the three isomers from the aforementioned studies are:

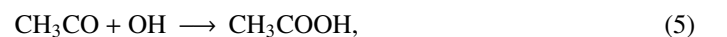


Table 8. Column density ratios between the 3 isomers of C₂H₄O₂ and abundances from star-forming regions in this work and in literature.

Source		AA/GA	MF/GA	Diameter [au]	X(MF)	X(AA)	X(GA)	$N(\text{H}_2)$ [cm ⁻²]	References
<i>High-mass sources</i>									
G31		9–17	40 ± 16	4400	(2.0 ± 0.6) × 10 ⁻⁷	(4.3–8) × 10 ⁻⁸	(5.0 ± 1.4) × 10 ⁻⁹	1.0 × 10 ²⁵	This work
SgrB2(N)-LMH		2.9 ± 0.7	7.5 ± 0.7	17 000	(1.2 ± 0.3) × 10 ⁻⁹	(4.6 ± 1.4) × 10 ⁻¹⁰	(1.6 ± 0.4) × 10 ⁻¹⁰	1.0 × 10 ^{26(a)}	(1)
NGC 6334I MM1	I	0.7 ± 0.3	15 ± 5	350	–	–	–	–	(2)
	II	0.4 ± 0.2	16 ± 7		–	–	–	–	
	III	1.4 ± 0.8	17 ± 13		–	–	–	–	
	IV	0.42 ± 0.19	38 ± 13		–	–	–	–	
	V	7.4 ± 1.2	45 ± 6		–	–	–	–	
	VI	1.0 ± 0.4	13 ± 8		–	–	–	–	
	VII	1.4 ± 0.7	25 ± 11		–	–	–	–	
	VIII	0.9 ± 0.3	30 ± 12		–	–	–	–	
	IX	0.6 ± 0.3	35 ± 15		–	–	–	–	
NGC 6334I MM2	I	>5 ^(*)	>158 ^(*)	350	–	–	–	–	(2)
	II	>7 ^(*)	>110 ^(*)		–	–	–	–	
	III	>14 ^(*)	>373 ^(*)		–	–	–	–	
W51e1/e2		>15 ^(*)	>550 ^(*)	13 000	–	–	–	–	(3),(4)
G34.3+0.2		>5 ^(*)	>200 ^(*)	29 000	–	–	–	–	(3),(5)
<i>Low-mass sources</i>									
IRAS 16293-2422B		0.09 ± 0.04	6 ± 3	60	<3.3 × 10 ⁻⁸	<5.0 × 10 ⁻¹⁰	<5.8 × 10 ⁻⁹	>1.2 × 10 ^{25(b)}	(6),(7)
IRAS 16293-2422A		3.5 ± 1.5	210 ± 90	60	–	–	–	–	(8)
NGC 1333 IRAS 4A		<1.1 ^(**)	6 ± 4 ^(**)	120	(1.4 ± 1.0) × 10 ⁻⁹	<2.6 × 10 ⁻¹⁰	(2.4 ± 1.1) × 10 ⁻¹⁰	3.7 × 10 ^{25(c)}	(5),(9)

Notes. Sources in literature where all three isomers have been detected, or only one isomer not detected with an estimate on the upper limit of the column density; column density ratio AA/GA; column density ratio MF/GA; linear scale to which the ratio refers; abundance of MF; abundance of AA; abundance of GA; $N(\text{H}_2)$ used to derive the abundances; references for the column density values of the 3 isomers of C₂H₄O₂. The high-mass sources are listed in the upper part of the table, while the low-mass sources are listed in the lower part. ^(*)The column density of GA is an upper limit, so the ratio with MF and AA are lower limits; ^(**) estimate of MF and GA inside 0.5 from Taquet et al. (2015) and upper limit of AA rescaled to the same source size (0.5) from the value in Remijan et al. (2003). Since the column density of AA is an upper limit, the ratio AA/GA is an upper limit too. Abundances have been calculated only for sources for which an estimate of $N(\text{H}_2)$ is available within an angular diameter close to the one of the derived column densities of the 3 isomers.

References. References to the works on the 3 isomers: (1) Xue et al. (2019); (2) El-Abd et al. (2019); (3) Lykke et al. (2015); (4) Remijan et al. (2002); (5) Remijan et al. (2003); (6) Jørgensen et al. (2012); (7) Jørgensen et al. (2016); (8) Manigand et al. (2020); (9) Taquet et al. (2015). Reference to the $N(\text{H}_2)$ values: ^(a)mean value for position AN01 from Sánchez-Monge et al. (2017); ^(b)lower limit from (7); ^(c)derived from emission at 145 GHz from (9).

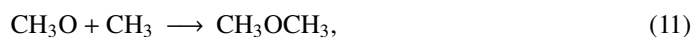
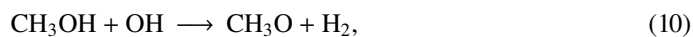


Laboratory studies have also investigated possible routes for the formation of MF and GA on the surface of dust grains at low temperatures (Fedoseev et al. 2015; Chuang et al. 2016), showing that the reactions can take place even at temperatures as low as ~10 K. This has also been confirmed by Simons et al. (2020) using microscopic kinetic Monte Carlo simulations to simulate ice chemistry. Nevertheless, the increase of temperature in more evolved sources, like G31, increases the mobility of radicals on the surface leading to higher efficiency of the routes (Eqs. (4)–(8)) than those proposed in the experiment at low temperatures.

Gas-phase chemical routes might be also efficient at low temperatures and, thus, capable of explaining the detection of COMs in a cold environment (Öberg et al. 2010; Bacmann et al. 2012; Jiménez-Serra et al. 2016). In the model by Vasyunin & Herbst (2013), the most efficient gas-phase pathway to the formation of MF in a cold environment ($T \sim 10$ K) is given by the radiative association reaction (Horn et al. 2004):



followed by a dissociative recombination and an isomerization, that leads to the formation of MF. Balucani et al. (2015) proposed the following chain of reactions for the production of MF:



starting from the non-thermal desorption of methanol at low temperature. Vasyunin et al. (2017) included these new gas-phase reactions in their network to model abundances of complex organic species in the prestellar core L1544 including MF. Reaction (13) of the chain proposed by Balucani et al. (2015) is indeed the most efficient pathway for the formation of MF in this cold environment, accounting for two-thirds of the total production. Nevertheless, the production of the reactant CH₃OCH₂ is mainly dominated by the reaction studied by Shannon et al. (2014):



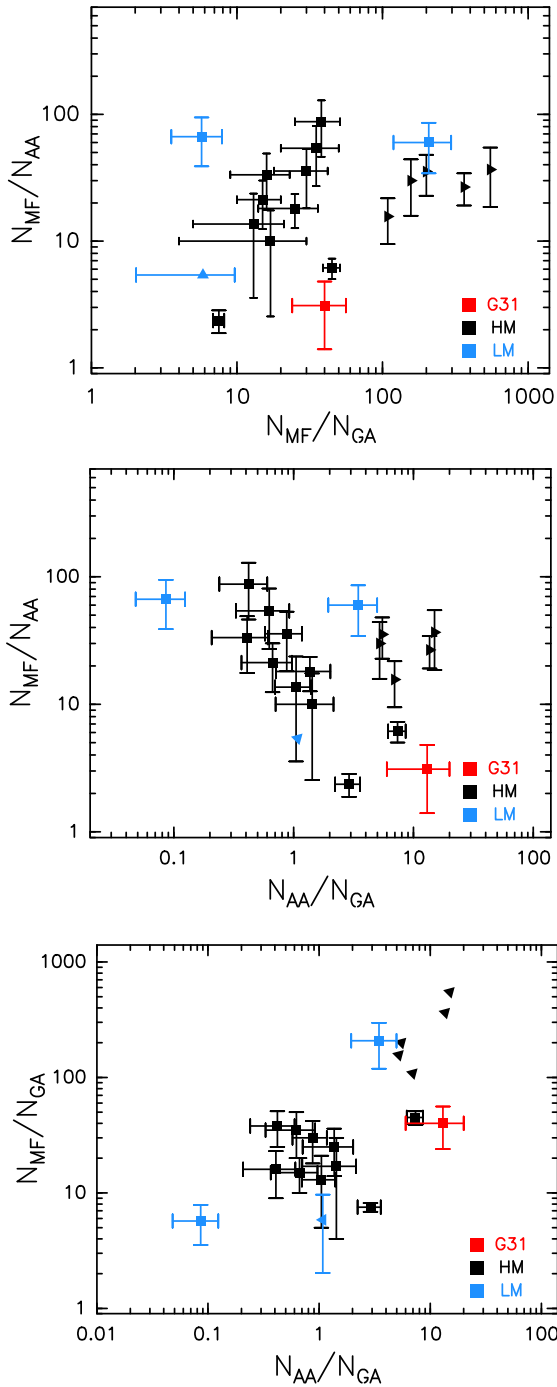


Fig. 14. Comparison of the ratios between the three isomers of $C_2H_4O_2$ for the sources listed in Table 8. For G31, the values of MF/AA and AA/GA given in the figure are calculated using the mean value of the observed range $4.3 \times 10^{17} - 8.4 \times 10^{17} \text{ cm}^{-2}$ for AA, while the error bars have been estimated from the entire range.

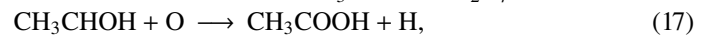
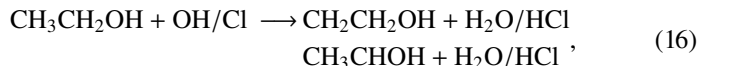
rather than reaction (12), due to the low abundances of F and Cl. Approximately the remaining one third of total MF is generated by the ion-neutral reaction



and the subsequent dissociative recombination of $HC(OH)OCH_3^+$. The abundances of COMs, such as MF, dimethyl ether, formamide, and acetaldehyde are of the order of $10^{-12} - 10^{-10}$, thus, gas-phase reactions can reproduce the

observed abundances in cold low-mass objects. However, the abundance of MF derived in, for instance, L1544 or L1689B (Bacmann et al. 2012; Jiménez-Serra et al. 2016) is two to three orders of magnitude below that obtained for G31. Coutens et al. (2018) report that in their models for star-forming regions hosting a proto-star with mass from 1 to $60 M_\odot$ MF reaches abundances in the range $7 \times 10^{-10} - 2.5 \times 10^{-8}$, when the temperatures are high (100–300 K). These results are obtained with a network considering only gas-phase reactions for MF, including the ones in Balucani et al. (2015). They found that the peak of abundance of MF from gas-phase reactions only is a factor of two smaller than the value estimated in G31 by Rivilla et al. (2017), and the inclusion of route (4) on grain-surface is able to bridge this discrepancy. However, the new estimate presented in this work is a factor of five greater, implying that grain surface reactions are likely responsible for the majority of the abundance of MF detected in G31. Garrod (2013) found that reaction (4) can efficiently lead to abundances of $\sim 1 \times 10^{-7}$. From this, MF would behave in a similar way as other COMs such as formamide (see Quénard et al. 2018), whose production at low T is dominated by gas-phase chemistry, while its formation is governed by grain surface reactions in warm or hot sources such as hot corinos or hot cores.

For GA and AA, Skouteris et al. (2018) introduced a new chain of reactions that leads to the formation of these $C_2H_4O_2$ isomers:



starting from the sublimation of ethanol in gas-phase. The GA abundance prediction of the model that simulates the chemistry in a hot corino, varies over a broad range $\sim 10^{-10} - 10^{-8}$ and shows a strong dependence on the hydroxyl radical OH, on the rate coefficient of the reaction (16), and on the initial abundance of ethanol (CH_3CH_2OH). Assuming the most favorable rate coefficient and an abundance of ethanol of $\sim 6 \times 10^{-8}$ (wrt H_2), the model of Skouteris et al. (2018) predicts an abundance of GA close to the one derived in this work, but the abundance of AA is about one order of magnitude lower than the new estimate towards G31. These values are derived before sublimated ethanol is fully consumed for a gas with temperature equal to 100 K and H nuclei density $2 \times 10^8 \text{ cm}^{-3}$. Nevertheless, Coutens et al. (2018) have shown that GA can efficiently form from reactions (6) and (8) on dust grains and reproduce the abundances in G31 observed by Rivilla et al. (2017) and the new estimate presented in this work. It must be noted that these calculations have been made assuming a mass of $25 M_\odot$ (from Osorio et al. 2009) for the proto-star embedded in G31, derived using the old estimates of distance of 7.9 kpc. The new estimate of the distance to G31, which affects the estimate of the mass and luminosity of the proto-star(s) embedded in G31, might lead to a change in the modeled abundances. Similarly to Coutens et al. (2018), Rivilla et al. (2019) have found that the main formation route for GA in IRAS 16293-2422 proto-stars is reaction (6), with possible contributions also from reaction (8). The formation of GA through grain-surface reactions in Rivilla et al. (2019) is strengthened by the fact that their grain-surface network reproduces both GA and that of its precursor HCO abundances.

Unlike GA, the abundance of AA from gas-phase reactions as a function of ethanol abundance and the different rate coefficient possibilities is not given in Skouteris et al. (2018). If we

Table 9. Summary of the chemical routes to the formation of the 3 isomers of C₂H₄O₂, and related reactions.

# in text	Reaction	Type	Comments
			MF
(4)	$\text{HCO} + \text{CH}_3\text{O} \rightarrow \text{CH}_3\text{OCHO}$	gs	Efficient pathway for the formation of MF from (a) and (b).
(9)	$\text{H}_2\text{CO}^+ + \text{H}_2\text{CO} \rightarrow \text{H}_2\text{COHOCH}_2^+ + \text{h}\nu$ $\text{H}_2\text{COHOCH}_2^+ + \text{e}^- \rightarrow \text{CH}_3\text{OCHO} + \text{H}$	gp	Efficient for the formation of MF in cold environment (c). Not efficient to reproduce the observed abundance in G31 (b).
(10)	$\text{CH}_3\text{OH} + \text{OH} \rightarrow \text{CH}_3\text{O} + \text{H}_2$		Efficient for the formation of MF in cold environment (d,e). (e) Found react.
(11)	$\text{CH}_3\text{O} + \text{CH}_3 \rightarrow \text{CH}_3\text{OCH}_3$		(14) more efficient than (12). Not efficient to reproduce the observed abundance
(12)	$\text{CH}_3\text{OCH}_3 + \text{F}/\text{Cl} \rightarrow \text{CH}_3\text{OCH}_2 + \text{HF}/\text{HCl}$	gp	in G31 (b).
(13)	$\text{CH}_3\text{OCH}_2 + \text{O} \rightarrow \text{CH}_3\text{OCHO} + \text{H}$		
(15)	$\text{CH}_3\text{OH}_2^+ + \text{HCOOH} \rightarrow \text{HC(OH)OCH}_3^+ + \text{H}_2\text{O}$ $\text{HC(OH)OCH}_3^+ + \text{e}^- \rightarrow \text{CH}_3\text{OCHO} + \text{H}$	gp	Second most efficient reaction in cold environment (e), after (13). Not efficient in reproduce the observed abundance in G31 (b).
(19)	$\text{H}_2\text{CO} + \text{H}_2\text{CO} \rightarrow \text{CH}_3\text{OCHO}$	*	Proposed by (f), presents high energy barrier, and is more efficient on grain-surface but no calculation made to predict the abundances in star-forming regions
			AA
(5)	$\text{CH}_3\text{CO} + \text{OH} \rightarrow \text{CH}_3\text{COOH}$	gs	Main route on grain-surface, not efficient enough to reproduce the observed abundance in this work (a).
(16)	$\text{CH}_3\text{CH}_2\text{OH} + \text{OH}/\text{Cl} \rightarrow \text{CH}_2\text{CH}_2\text{OH} + \text{H}_2\text{O}/\text{HCl}$ $\text{CH}_3\text{CHOH} + \text{H}_2\text{O}/\text{HCl}$	gp	Predictions by (g) span in a broad range for GA, and may cover the abundance observed in this work for AA. Strong dependence on the rate coefficient of reaction (16) and on abundances of reactant OH and ethanol, which should be constrained to reduce uncertainties.
(17)	$\text{CH}_3\text{CHOH} + \text{O} \rightarrow \text{CH}_3\text{COOH} + \text{H}$		
(19)	$\text{H}_2\text{CO} + \text{H}_2\text{CO} \rightarrow \text{CH}_3\text{OOCH}$	*	Proposed by (f), presents high energy barrier, and is more efficient on grain-surface but no calculation made to predict the abundances in star-forming regions
			GA
(6)	$\text{HCO} + \text{CH}_2\text{OH} \rightarrow \text{CH}_2\text{OHCHO}$	gs	Efficient to reproduce the abundance in this work together with (8) by (b) and in IRAS 16293-2422 A & B by (h). The estimate by (b) might change due to the new estimate of the distance to G31.
(7)	$\text{CH}_3\text{OH} + \text{HCO} \rightarrow \text{CH}_2\text{OHCHO} + \text{H}$	gs	Efficient route for the formation of GA (i), not tested in (b) and (h).
(8)	$2\text{HCO} + 2\text{H} \rightarrow \text{CH}_2\text{OHCHO}$	gs	Efficient to reproduce the abundance in this work together with (6) by (b) and possibly contributes in IRAS 16293-2422 A & B by (h). The estimate by (b) might change due to the new estimate of distance to G31.
(16)	$\text{CH}_3\text{CH}_2\text{OH} + \text{OH}/\text{Cl} \rightarrow \text{CH}_2\text{CH}_2\text{OH} + \text{H}_2\text{O}/\text{HCl}$ $\text{CH}_3\text{CHOH} + \text{H}_2\text{O}/\text{HCl}$	gp	Predictions by (g) span in a broad range, that cover the abundances observed in this work. Strong dependence on the rate coefficient of reaction (16) and on abundances of reactant OH and ethanol, which should be constrained to reduce uncertainties.
(18)	$\text{CH}_2\text{CH}_2\text{OH} + \text{O} \rightarrow \text{CH}_2\text{OHCHO} + \text{H}$		
(19)	$\text{H}_2\text{CO} + \text{H}_2\text{CO} \rightarrow \text{CH}_3\text{OCHO}$	*	Proposed by (f), presents high energy barrier, and is more efficient on grain-surface but no calculation made to predict the abundances in star-forming regions on grains surface. (j) Simulated the abundance in gas-phase in high-mass protostellar objects, and it is not efficient to reproduce G31 abundance.
			Other reactions
(14)	$\text{CH}_3\text{OCH}_3 + \text{OH} \rightarrow \text{CH}_3\text{OCH}_2 + \text{H}_2\text{O}$	gp	More efficient than reaction (12) to produce CH ₃ OCH ₂ (e) needed for reaction (13).

Notes. gs stands for grain-surface reactions; gp stands for gas-phase reactions; * stands for both grain-surface and gas-phase reactions.

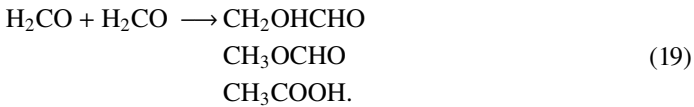
References. (a)Garrod (2013); (b)Coutens et al. (2018); (c)Vasyunin & Herbst (2013); (d)Balucani et al. (2015); (e)Vasyunin et al. (2017); (f)Ahmad et al. (2020); (g)Skouteris et al. (2018); (h)Rivilla et al. (2019); (i)Woods et al. (2012); (j)Silva et al. (2020).

assume that the ratio AA/GA~ 10 (found in the aforementioned specific case of ethanol abundance $\sim 6 \times 10^{-8}$), is conserved, in principle the abundance of AA in this work could be covered by the wide range derived from multiplying for a factor 10 the given range $\sim 10^{-10}$ – 10^{-8} of GA in Skouteris et al. (2018). However, this relies on many assumptions that need to be further tested. The observed abundance estimated this work cannot be reproduced with reaction (5) included in the chemical network

of Garrod (2013). Further investigations of reaction (16) and of the abundance of ethanol in G31 is needed to constrain the formation route of AA in G31, which might also lead to completely discard the proposed gas-phase pathways (16)–(18).

The fact that the rate coefficients of reactions (17) and (18) have a dependence on temperature implies that both gas-phase and grain-surface reactions can possibly explain the more extended emission of MF with respect to AA and GA. In

fact, the gas-phase reactions for AA and GA would be more efficient closer to the center of the core due to the increase in temperature. On the other hand, for grain-surface reactions, [Burke et al. \(2015\)](#) found that MF desorbs at lower temperature with respect to AA and GA, which follow the desorption of water ice. Recently, [Ahmad et al. \(2020\)](#) proposed a new pathway for the formation of the three isomers both in gas-phase and on grain surface, involving formaldehyde:



All the three channels show high energy barriers that can be crossed thanks to quantum tunneling or thermal hopping. The study also reports that reaction (19), in all the three cases, is more efficient on grain-surface. However, no calculations have been made to predict the abundances that can be reached thanks to this pathway on grain surface in star-forming regions. [Silva et al. \(2020\)](#) simulated the channel of reaction (19) that leads to GA in the condition of a high-mass protostellar object in gas-phase, and the maximum abundance is of $\sim 10^{-12}$, that is orders of magnitude below the value observed in G31 as well as SgrB2(N).

A summary of the reaction routes and the predictions of the different models is given in Table 9. In conclusion, for MF grain-surface reactions are likely responsible for its high abundance in G31. The results from GA and its precursor HCO in IRAS 16293-2422 A & B by [Rivilla et al. \(2019\)](#) could suggest a grain-surface formation for GA in hot sources like G31. However, the case of GA is not of straightforward interpretation as both gas-phase and grain-surface reactions could be at play considered the uncertainties in the chemical models aforementioned, and, as a results, both scenarios are able to explain the more compact morphology of GA with respect to MF. For AA, the routes included in [Garrod \(2013\)](#) are not able to reproduce the high-abundance found towards G31, while the gas-phase route proposed by [Skouteris et al. \(2018\)](#) could not be excluded and, rather, needs to be further tested in order to reduce the large uncertainties on their predictions. In case of any future discrepancy with the narrowed prediction of this route and the abundance found in G31, the main pathway(s) responsible for the high abundance of AA in G31 should be further investigated. Similarly to GA, both gas-phase and grain-surface reactions could explain the more compact morphology of AA with respect to MF.

6. Conclusions

In this paper, we present the data of the GUAPOS survey, an unbiased spectral survey performed with ALMA towards G31.41+0.31, one of the most chemically rich HMC known, located outside the GC. The survey covered a ~ 32 GHz bandwidth with a spectral resolution of 0.488 MHz and an angular resolution of $1''.2$.

We detected, for the first time, all the three isomers of $\text{C}_2\text{H}_4\text{O}_2$ towards this HMC. This increases the number of high-mass sources outside the GC, where the three isomers have been detected, to a total of three. The emission of all the isomers is compact towards the HMC and the most extended emission is from MF, while AA is the most compact. Here, MF is the most abundant of the isomers, confirming what has been seen in the other sources. Then, G31 shows the largest abundance of AA with respect to the other sources. From the comparison with the

literature, it seems that a unique behavior of the three isomers does not exist both in high- and low-mass sources.

The comparison with chemical models suggests that the MF abundance found in G31 is not reproducible using only gas-phase routes, thus revealing the need for grain-surface reactions. On the other hand, the scenario for AA and GA does not have a straightforward interpretation. Both gas-phase and grain-surface reactions are able to reproduce the observed abundance of GA, with uncertainties. However, in the case of the gas-phase reactions proposed by [Skouteris et al. \(2018\)](#), the wide range of predicted abundances for GA is strictly related to the uncertainties in the reaction of $\text{CH}_3\text{CH}_2\text{OH}$ with the hydroxyl radical, while the predictions from grain-surface reactions in [Coutens et al. \(2018\)](#) might be altered by the adoption of the new estimate of the distance to G31. For AA, the chemical model by [Garrod \(2013\)](#), including route (5) on grain-surface, is not able to reproduce the high abundance found in G31. On the other hand, the gas-phase route by [Skouteris et al. \(2018\)](#) should be further tested to reduce the uncertainties. Moreover, the more compact morphology of these two isomers with respect to MF can be explained in both scenarios (grain-surface and gas-phase) as well. To remove some of the uncertainties, laboratory and theoretical studies on reaction (16) are needed, together with an estimate of the abundance of the reactant $\text{CH}_3\text{CH}_2\text{OH}$ in G31. If as a result, this route is not efficient enough to reproduce the abundance presented here, new formation routes should be investigated. To better understand and constrain whether the main chemical routes that lead to GA are in gas-phase or in grain-surface chemistry, or whether both reactions are necessary, a systematic study is needed, which would include the new routes by [Skouteris et al. \(2018\)](#), such as the one performed in [Quénard et al. \(2018\)](#) for formamide, in which the two types of chemistry were switched off alternatively.

Acknowledgements. The authors thank the referee Brett A. McGuire for the constructive comments which helped to improve the quality and readability of the paper. The authors wish to thank Holger S.P. Müller and Christian Endres for inserting acetic acid in the CDMS catalog. C.M. and L.C. acknowledge support from the Italian Ministero dell'Istruzione, Università e Ricerca through the grant Progetti Premiali 2012 – iALMA (CUP C52113000140001). I.J.-S. has received partial support from the Spanish FEDER (project number ESP2017-86582-C4-1-R). V.M.R. has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 664931. This paper makes use of the following ALMA data: ADS/JAO.ALMA#2017.1.00501.S. ALMA is a partnership of ESO (representing its member states), NSF (USA) and NINS (Japan), together with NRC (Canada), MOST and ASIAA (Taiwan), and KASI (Republic of Korea), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO and NAOJ.

References

- Ahmad, A., Shivani, Misra, A., & Tandon, P. 2020, *Res. A&A*, 20, 014
- Araya, E., Hofner, P., Kurtz, S., Olmi, L., & Linz, H. 2008, *ApJ*, 675, 420
- Arce, H. G., Santiago-García, J., Jørgensen, J. K., Tafalla, M., & Bachiller, R. 2008, *ApJ*, 681, L21
- Bacmann, A., Taquet, V., Faure, A., Kahane, C., & Ceccarelli, C. 2012, *A&A*, 541, L12
- Balucani, N., Ceccarelli, C., & Taquet, V. 2015, *MNRAS*, 449, L16
- Bauder, A. 1979, *J. Phys. Chem. Ref. Data*, 8, 583
- Beckwith, S. V. W., & Sargent, A. I. 1991, *ApJ*, 381, 250
- Belloche, A., Müller, H. S. P., Menten, K. M., Schilke, P., & Comito, C. 2013, *A&A*, 559, A47
- Belloche, A., Müller, H. S. P., Garrod, R. T., & Menten, K. M. 2016, *A&A*, 587, A91
- Belloche, A., Garrod, R. T., Müller, H. S. P., et al. 2019, *A&A*, 628, A10
- Beltrán, M. T., Cesaroni, R., Neri, R., et al. 2004, *ApJ*, 601, L187
- 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

- Beltrán, M. T., Cesaroni, R., Rivilla, V. M., et al. 2018, *A&A*, **615**, A141
- Beltrán, M. T., Padovani, M., Girart, J. M., et al. 2019, *A&A*, **630**, A54
- Beuther, H., Zhang, Q., Bergin, E. A., et al. 2007, *A&A*, **468**, 1045
- Bonato, M., Liuzzo, E., Giannetti, A., et al. 2018, *MNRAS*, **478**, 1512
- Bottinelli, S., Ceccarelli, C., Williams, J. P., & Lefloch, B. 2007, *A&A*, **463**, 601
- Briggs, D. S. 1995, *AAS Meeting Abstracts*, **187**, 112.02
- Brown, R. D., Crofts, J. G., Gardner, F. F., et al. 1975, *ApJ*, **197**, L29
- Burke, D. J., Puletti, F., Brown, W. A., et al. 2015, *MNRAS*, **447**, 1444
- Butler, R. A. H., De Lucia, F. C., Petkie, D. T., et al. 2001, *ApJS*, **134**, 319
- Calcutt, H., Viti, S., Codella, C., et al. 2014, *MNRAS*, **443**, 3157
- Carroll, P. B., Drouin, B. J., & Widicus Weaver, S. L. 2010, *ApJ*, **723**, 845
- Carvajal, M., Willaert, F., Demaison, J., & Kleiner, I. 2007, *J. Mol. Spectr.*, **246**, 158
- Cazaux, S., Tielens, A. G. G. M., Ceccarelli, C., et al. 2003, *ApJ*, **593**, L51
- Cernicharo, J., Marcelino, N., Roueff, E., et al. 2012, *ApJ*, **759**, L43
- Cesaroni, R. 2019, *A&A*, **631**, A65
- Cesaroni, R., Churchwell, E., Hofner, P., Walmsley, C. M., & Kurtz, S. 1994a, *A&A*, **288**, 903
- Cesaroni, R., Olmi, L., Walmsley, C. M., Churchwell, E., & Hofner, P. 1994b, *ApJ*, **435**, L137
- Cesaroni, R., Hofner, P., Walmsley, C. M., & Churchwell, E. 1998, *A&A*, **331**, 709
- Cesaroni, R., Hofner, P., Araya, E., & Kurtz, S. 2010, *A&A*, **509**, A50
- Cesaroni, R., Beltrán, M. T., Zhang, Q., Beuther, H., & Fallscheer, C. 2011, *A&A*, **533**, A73
- Cesaroni, R., Sánchez-Monge, Á., Beltrán, M. T., et al. 2017, *A&A*, **602**, A59
- Chuang, K. J., Fedoseev, G., Ioppolo, S., van Dishoeck, E. F., & Linnartz, H. 2016, *MNRAS*, **455**, 1702
- Coletta, A., Fontani, F., Rivilla, V. M., et al. 2020, *A&A*, **641**, A54
- Coutens, A., Viti, S., Rawlings, J. M. C., et al. 2018, *MNRAS*, **475**, 2016
- Csengeri, T., Belloche, A., Bontemps, S., et al. 2019, *A&A*, **632**, A57
- Demaison, J., Dubrulle, A., Boucher, D., Burie, J., & van Eijck, B. P. 1982, *J. Mol. Spectr.*, **94**, 211
- Demaison, J., Boucher, D., Dubrulle, A., & Van Eijck, B. P. 1983, *J. Mol. Spectr.*, **102**, 260
- 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, *J. Mol. Spectr.*, **327**, 95
- Favre, C., Carvajal, M., Field, D., et al. 2014, *ApJS*, **215**, 25
- 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
- Galametz, M., Maury, A. J., Valdivia, V., et al. 2019, *A&A*, **632**, A5
- Garrod, R. T. 2013, *ApJ*, **765**, 60
- Garrod, R. T., & Herbst, E. 2006, *A&A*, **457**, 927
- Garrod, R. T., Widicus Weaver, S. L., & Herbst, E. 2008, *ApJ*, **682**, 283
- Gibb, A. G., Wyrowski, F., & Mundy, L. G. 2004, *ApJ*, **616**, 301
- Girart, J. M., Beltrán, M. T., Zhang, Q., Rao, R., & Estalella, R. 2009, *Science*, **324**, 1408
- Herbst, E., & van Dishoeck, E. F. 2009, *ARA&A*, **47**, 427
- Hitoshi, O., Kazumi, O., Kojiro, T., & Shozo, T. 2003, *Molecules*, **8**
- Hofner, P., & Churchwell, E. 1993, *Astrophysical Masers* (Berlin: Springer), 412, 164
- Hollis, J. M., Lovas, F. J., & Jewell, P. R. 2000, *ApJ*, **540**, L107
- Hollis, J. M., Vogel, S. N., Snyder, L. E., Jewell, P. R., & Lovas, F. J. 2001, *ApJ*, **554**, L81
- Horn, A., Møllendal, H., Sekiguchi, O., et al. 2004, *ApJ*, **611**, 605
- Ilyushin, V. V., Alekseev, E. A., Dyubko, S. F., et al. 2001, *J. Mol. Spectr.*, **205**, 286
- Ilyushin, V. V., Alekseev, E. A., Dyubko, S. F., & Kleiner, I. 2003, *J. Mol. Spectr.*, **220**, 170
- Ilyushin, V., Kryvda, A., & Alekseev, E. 2009, *J. Mol. Spectr.*, **255**, 32
- Ilyushin, V. V., Endres, C. P., Lewen, F., Schlemmer, S., & Drouin, B. J. 2013, *J. Mol. Spectr.*, **290**, 31
- Immer, K., Li, J., Quiroga-Nuñez, L. H., et al. 2019, *A&A*, **632**, A123
- Isokoski, K., Bottinelli, S., & van Dishoeck, E. F. 2013, *A&A*, **554**, A100
- Jacobsen, S. K., Jørgensen, J. K., Di Francesco, J., et al. 2019, *A&A*, **629**, A29
- Jiménez-Serra, I., Vasyunin, A. I., Caselli, P., et al. 2016, *ApJ*, **830**, L6
- Jørgensen, J. K., Favre, C., Bisschop, S. E., et al. 2012, *ApJ*, **757**, L4
- Jørgensen, J. K., van der Wiel, M. H. D., Coutens, A., et al. 2016, *A&A*, **595**, A117
- Kalenskii, S. V., & Johansson, L. E. B. 2010, *Astron. Rep.*, **54**, 1084
- Karakawa, Y., Oka, K., Odashima, H., Takagi, K., & Tsunekawa, S. 2001, *J. Mol. Spectr.*, **210**, 196
- Krisher, L. C., & Saegebarth, E. 1971, *J. Chem. Phys.*, **54**, 4553
- Laas, J. C., Garrod, R. T., Herbst, E., & Widicus Weaver, S. L. 2011, *ApJ*, **728**, 71
- Lykke, J. M., Favre, C., Bergin, E. A., & Jørgensen, J. K. 2015, *A&A*, **582**, A64
- Maeda, A., Medvedev, I. R., De Lucia, F. C., Herbst, E., & Groner, P. 2008a, *ApJS*, **175**, 138
- Maeda, A., De Lucia, F. C., & Herbst, E. 2008b, *J. Mol. Spectr.*, **251**, 293
- Manigand, S., Jørgensen, J. K., Calcutt, H., et al. 2020, *A&A*, **635**, A48
- Marstokk, K. M., & Møllendal, H. 1970, *J. Mol. Struct.*, **5**, 205
- Marstokk, K. M., & Møllendal, H. 1973, *J. Mol. Struct.*, **16**, 259
- Martín, S., Martín-Pintado, J., Blanco-Sánchez, C., et al. 2019, *A&A*, **631**, A159
- Mayen-Gijón, J. M., Anglada, G., Osorio, M., et al. 2014, *MNRAS*, **437**, 3766
- McGuire, B. A., Burkhardt, A. M., Kalenskii, S., et al. 2018, *Science*, **359**, 202
- McMullin, J. P., Waters, B., Schiebel, D., Young, W., & Golap, K. 2007, *ASP Conf. Ser.*, **376**, 127
- Mehring, D. M., Snyder, L. E., Miao, Y., & Lovas, F. J. 1997, *ApJ*, **480**, L71
- Möller, T., Endres, C., & Schilke, P. 2017, *A&A*, **598**, A7
- Müller, H. S. P., Thorwirth, S., Roth, D. A., & Winnewisser, G. 2001, *A&A*, **370**, L49
- Müller, H. S. P., Schlöder, F., Stutzki, J., & Winnewisser, G. 2005, *J. Mol. Struct.*, **742**, 215
- Öberg, K. I., Bottinelli, S., Jørgensen, J. K., & van Dishoeck, E. F. 2010, *ApJ*, **716**, 825
- Oesterling, L. C., Albert, S., De Lucia, F. C., Sastry, K. V. L. N., & Herbst, E. 1999, *ApJ*, **521**, 255
- Ogata, K., Odashima, H., Takagi, K., & Tsunekawa, S. 2004, *J. Mol. Spectr.*, **225**, 14
- Osorio, M., Anglada, G., Lizano, S., & D'Alessio, P. 2009, *ApJ*, **694**, 29
- Ossenkopf, V., & Henning, T. 1994, *A&A*, **291**, 943
- Ott, S. 2010, *ASP Conf. Ser.*, **434**, 139
- Pickett, H. M., Poynter, R. L., Cohen, E. A., et al. 1998, *J. Quant. Spectr. Rad. Transf.*, **60**, 883
- Plummer, G. M., Herbst, E., De Lucia, F., & Blake, G. A. 1984, *ApJS*, **55**, 633
- Plummer, G. M., Herbst, E., De Lucia, F. C., & Blake, G. A. 1986, *ApJS*, **60**, 949
- Quénard, D., Jiménez-Serra, I., Viti, S., Holdship, J., & Coutens, A. 2018, *MNRAS*, **474**, 2796
- Remijan, A., Snyder, L. E., Liu, S. Y., Mehringer, D., & Kuan, Y. J. 2002, *ApJ*, **576**, 264
- Remijan, A., Snyder, L. E., Friedel, D. N., Liu, S. Y., & Shah, R. Y. 2003, *ApJ*, **590**, 314
- Rivilla, V. M., Beltrán, M. T., Cesaroni, R., et al. 2017, *A&A*, **598**, A59
- Rivilla, V. M., Beltrán, M. T., Vasyunin, A., et al. 2019, *MNRAS*, **483**, 806
- Sakai, Y., Kobayashi, K., & Hirota, T. 2015, *ApJ*, **803**, 97
- Sánchez-Monge, Á., Schilke, P., Schmiedeke, A., et al. 2017, *A&A*, **604**, A6
- Sánchez-Monge, Á., Schilke, P., Ginsburg, A., Cesaroni, R., & Schmiedeke, A. 2018, *A&A*, **609**, A101
- Shannon, R. J., Caravan, R. L., Blitz, M. A., & Heard, D. E. 2014, *Phys. Chem. Chem. Phys.*, **16**, 3466
- Silva, S. G. S., Vichiatti, R. M., Haiduke, R. L. A., Machado, F. B. C., & Spada, R. F. K. 2020, *MNRAS*, **497**, 4486
- Simons, M. A. J., Lamberts, T., & Cuppen, H. M. 2020, *A&A*, **634**, A52
- Skouteris, D., Balucani, N., Ceccarelli, C., et al. 2018, *ApJ*, **854**, 135
- Taquet, V., López-Sepulcre, A., Ceccarelli, C., et al. 2015, *ApJ*, **804**, 81
- van Eijck, B. P., & van Duijneveldt, F. B. 1983, *J. Mol. Spectr.*, **102**, 273
- van Eijck, B. P., van Opheusden, J., van Schaik, M. M. M., & van Zoeren, E. 1981, *J. Mol. Spectr.*, **86**, 465
- Vastel, C., Ceccarelli, C., Lefloch, B., & Bachiller, R. 2014, *ApJ*, **795**, L2
- Vasyunin, A. I., & Herbst, E. 2013, *ApJ*, **769**, 34
- Vasyunin, A. I., Caselli, P., Dulieu, F., & Jiménez-Serra, I. 2017, *ApJ*, **842**, 33
- Widicus Weaver, S. L., Butler, R. A. H., Drouin, B. J., et al. 2005, *ApJS*, **158**, 188
- Włodarczak, G., & Demaison, J. 1988, *A&A*, **192**, 313
- 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
- Xue, C., Remijan, A. J., Burkhardt, A. M., & Herbst, E. 2019, *ApJ*, **871**, 112
- Ysard, N., Koehler, M., Jiménez-Serra, I., Jones, A. P., & Verstraete, L. 2019, *A&A*, **631**, A88

Appendix A: Catalog entries documentation for the three isomers of C₂H₄O₂

A.1. Methyl Formate

The data set used in JPL catalog is based on Ilyushin et al. (2009) and includes data from Brown et al. (1975), Bauder (1979), Demaison et al. (1983), Plummer et al. (1984, 1986), Oesterling et al. (1999), Karakawa et al. (2001), Hitoshi et al. (2003), Ogata et al. (2004), Carvajal et al. (2007), Maeda et al. (2008b,a)⁸.

A.2. Acetic Acid

The data set used in the CDMS catalog is based on Ilyushin et al. (2013) and includes data from Krisher & Saegebarth (1971), van Eijck et al. (1981), Demaison et al. (1982, 1983), van Eijck & van Duijneveldt (1983), Wlodarczak & Demaison (1988), Ilyushin et al. (2001, 2003)⁹.

A.3. Glycolaldehyde

The data set used in JPL catalog is based on Marstokk & Møllendal (1970, 1973), Butler et al. (2001), Widicus Weaver et al. (2005), Carroll et al. (2010)¹⁰.

Appendix B: Optical depth of the lines

Equation (B.28) of Möller et al. (2017) gives the optical depth of a line as a function of frequency:

$$\tau_\nu = \frac{c^2}{8\pi\nu_{u,l}^2} A_{u,l} N_u (e^{(E_u - E_l)/k_B T_{\text{ex}}} - 1) \phi_{u,l}(\nu), \quad (\text{B.1})$$

where c is the speed of light, $\nu_{u,l}$ is the frequency of the transition, $A_{u,l}$ is the Einstein's coefficient of the transition, N_u is the

column density of the molecule in the upper state of the transition, and E_u and E_l are the energy level of the upper and lower states of the transition and $\phi_{u,l}(\nu)$ is the normalized line profile (i.e., $\int \phi_{u,l}(\nu) d\nu = 1$). N_u is related to the total column density of the molecule, N_{tot} , through the Boltzmann distribution,

$$N_u = \frac{g_u}{Q(T_{\text{ex}})} N_{\text{tot}} e^{-E_u/k_B T_{\text{ex}}}, \quad (\text{B.2})$$

where g_u is the degeneracy of the level u , E_u the corresponding and $Q(T_{\text{ex}})$ the partition function of the molecular species, calculated at the temperature of excitation T_{ex} . If we assume that τ_ν has a Gaussian profile of the type

$$\tau_\nu = \tau_0 e^{-\frac{4 \ln 2 (\nu - \nu_{u,l})^2}{\Delta\nu^2}} \quad (\text{B.3})$$

where τ_0 is the opacity at the line center and $\Delta\nu$ is the FWHM, then the integral of the optical depth over frequency will be related to the central opacity of the line by

$$\int \tau_\nu d\nu = \frac{\sqrt{\pi}}{2\sqrt{\ln 2}} \tau_0 \Delta\nu. \quad (\text{B.4})$$

Integrating Eq. (B.1) on ν and using Eqs. (B.2) and (B.4) we obtain the expression for the optical depth at the center of the line

$$\tau_0 = \frac{\sqrt{\ln 2}}{4\pi\sqrt{\pi}} \frac{c^3}{\nu_{u,l}^3 \Delta\nu} \frac{g_u}{Q(T_{\text{ex}})} A_{u,l} N_{\text{tot}} e^{-E_l/k_B T_{\text{ex}}} (1 - e^{h\nu_{u,l}/k_B T_{\text{ex}}}) \quad (\text{B.5})$$

where $\Delta\nu$ is the FWHM in velocity units, namely $\Delta\nu = \Delta\nu c/\nu_{u,l}$.

⁸ More information is available at <https://spec.jpl.nasa.gov/ftp/pub/catalog/doc/d060003.pdf>

⁹ More information is available at <https://cdms.astro.uni-koeln.de/cdms/portal/catalog/1605/#>

¹⁰ More information is available at <https://spec.jpl.nasa.gov/ftp/pub/catalog/doc/d060006.pdf>

Appendix C: Tables of identified transitions

Tables C.1-C.3 are only available at the CDS.

Appendix D: Full spectrum

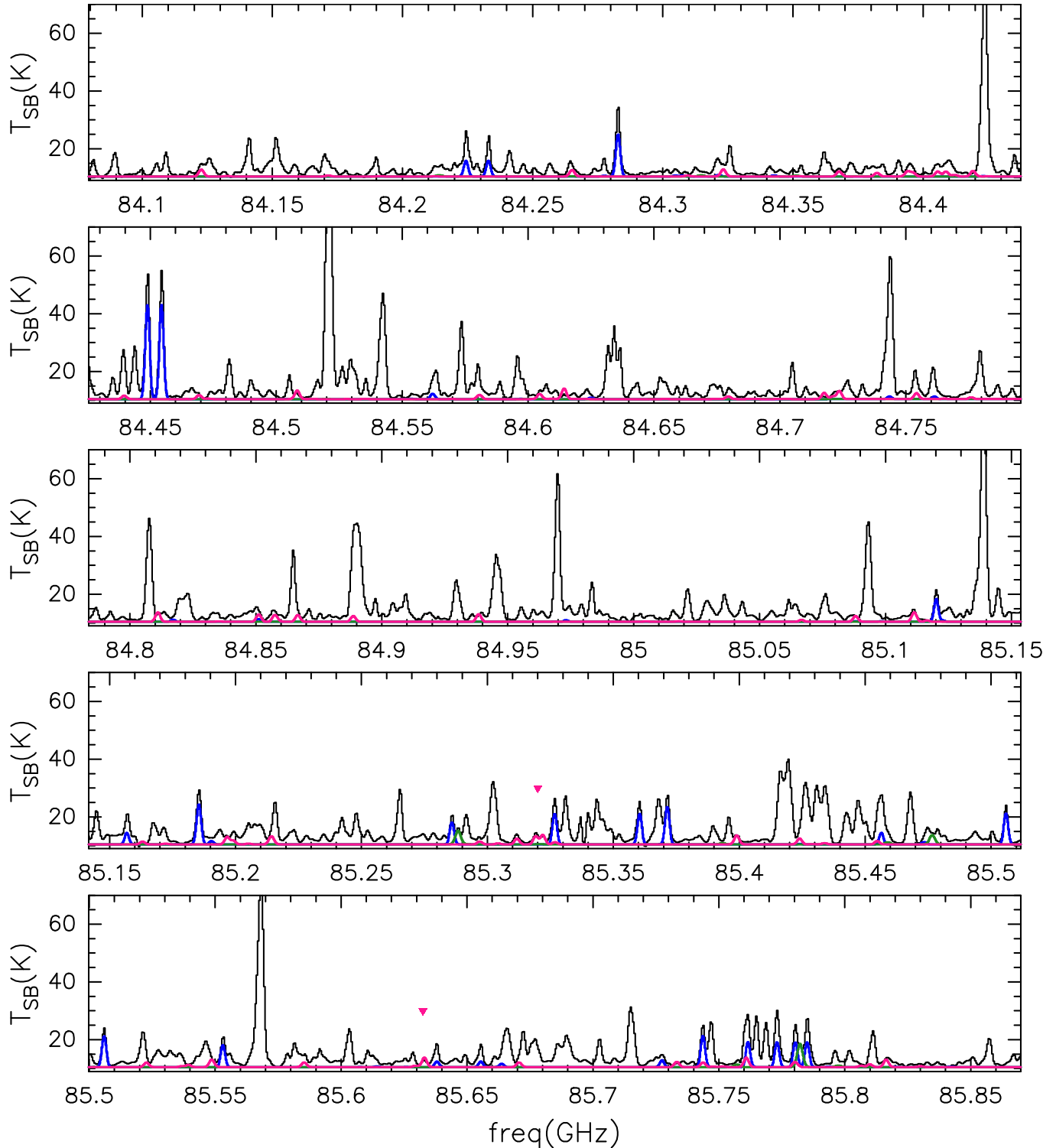


Fig. D.1. Total spectra of the GUAPOS project in black. In blue, we show the synthetic spectrum of the best fit of methyl formate (MF); in green the synthetic spectrum of the best fit of glycolaldehyde (GA); in pink the synthetic spectrum of the best fit of acetic acid (AA). The colored triangles indicate the transitions used to constrain the fitting procedure of the 3 different molecules. Closer views of those transitions are given in Figs. 5–7.

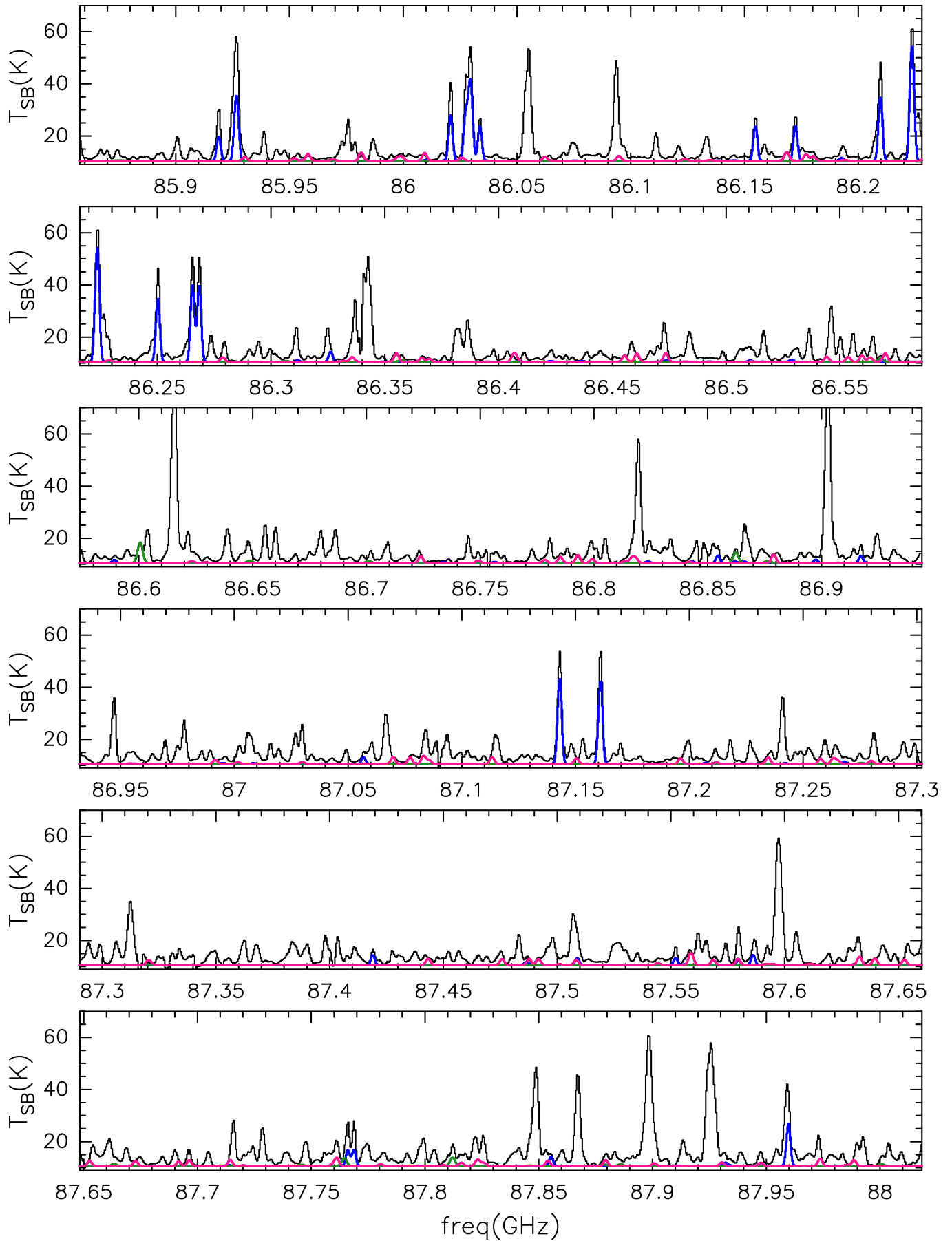


Fig. D.1. continued.

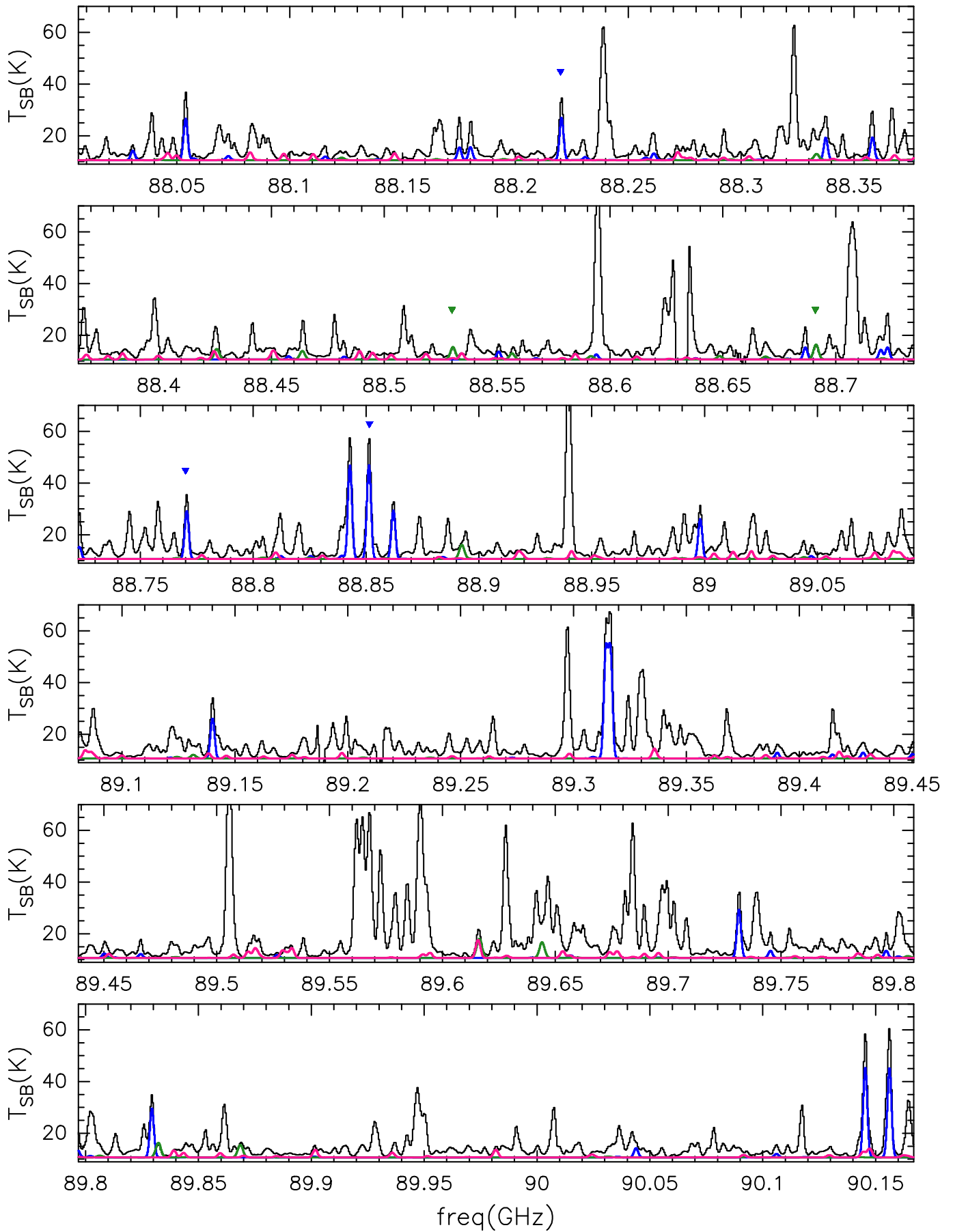
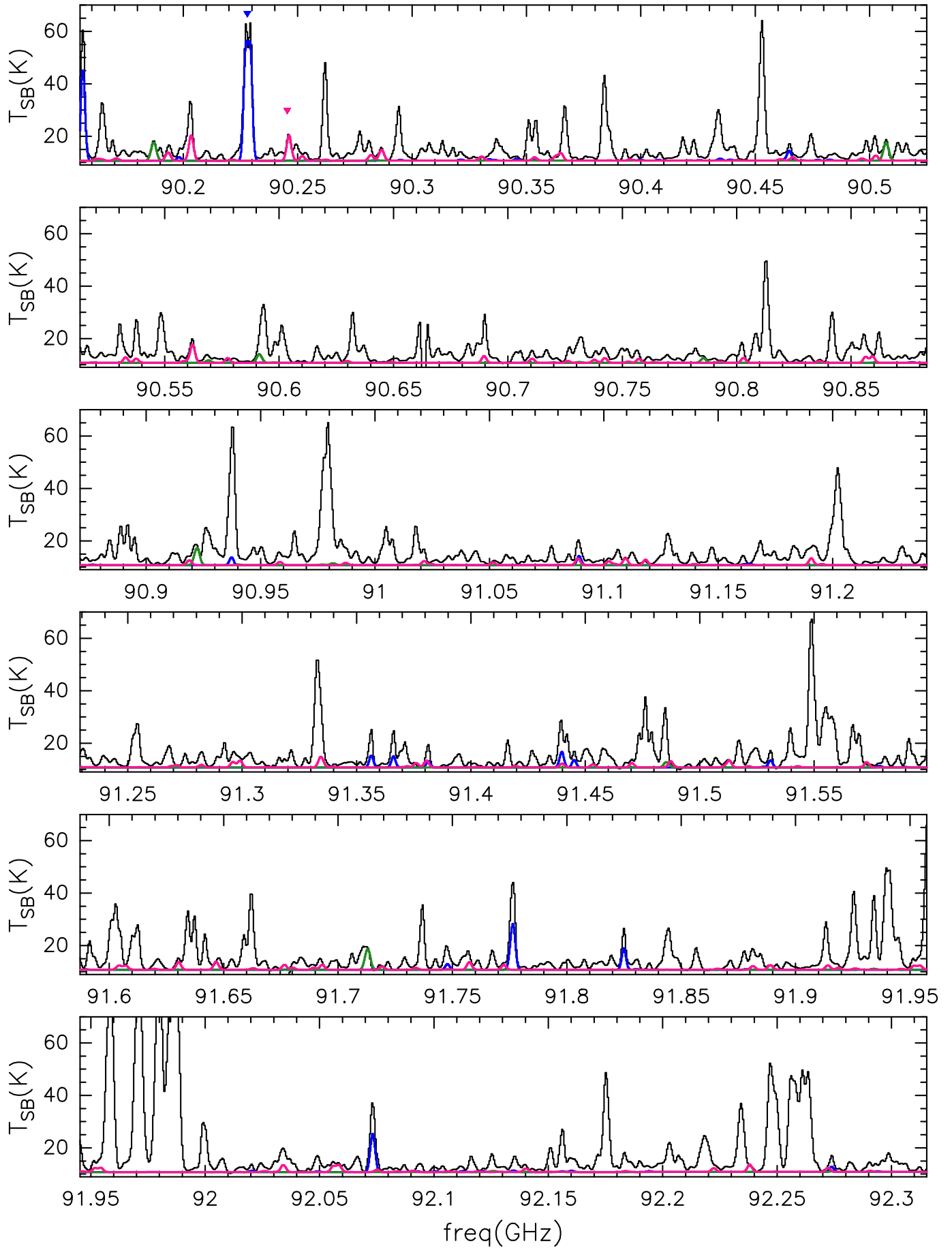


Fig. D.1. continued.

**Fig. D.1.** continued.

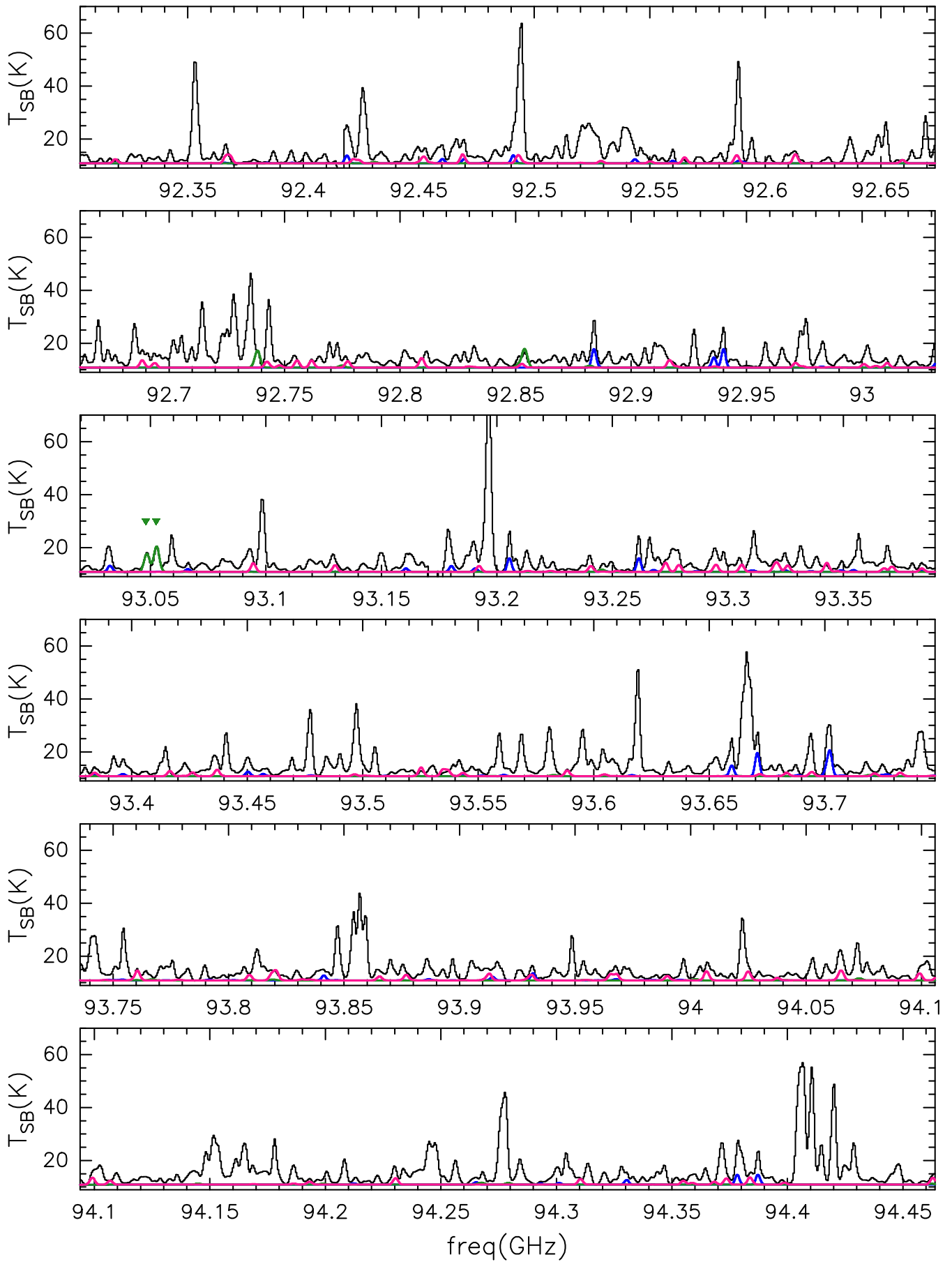


Fig. D.1. continued.

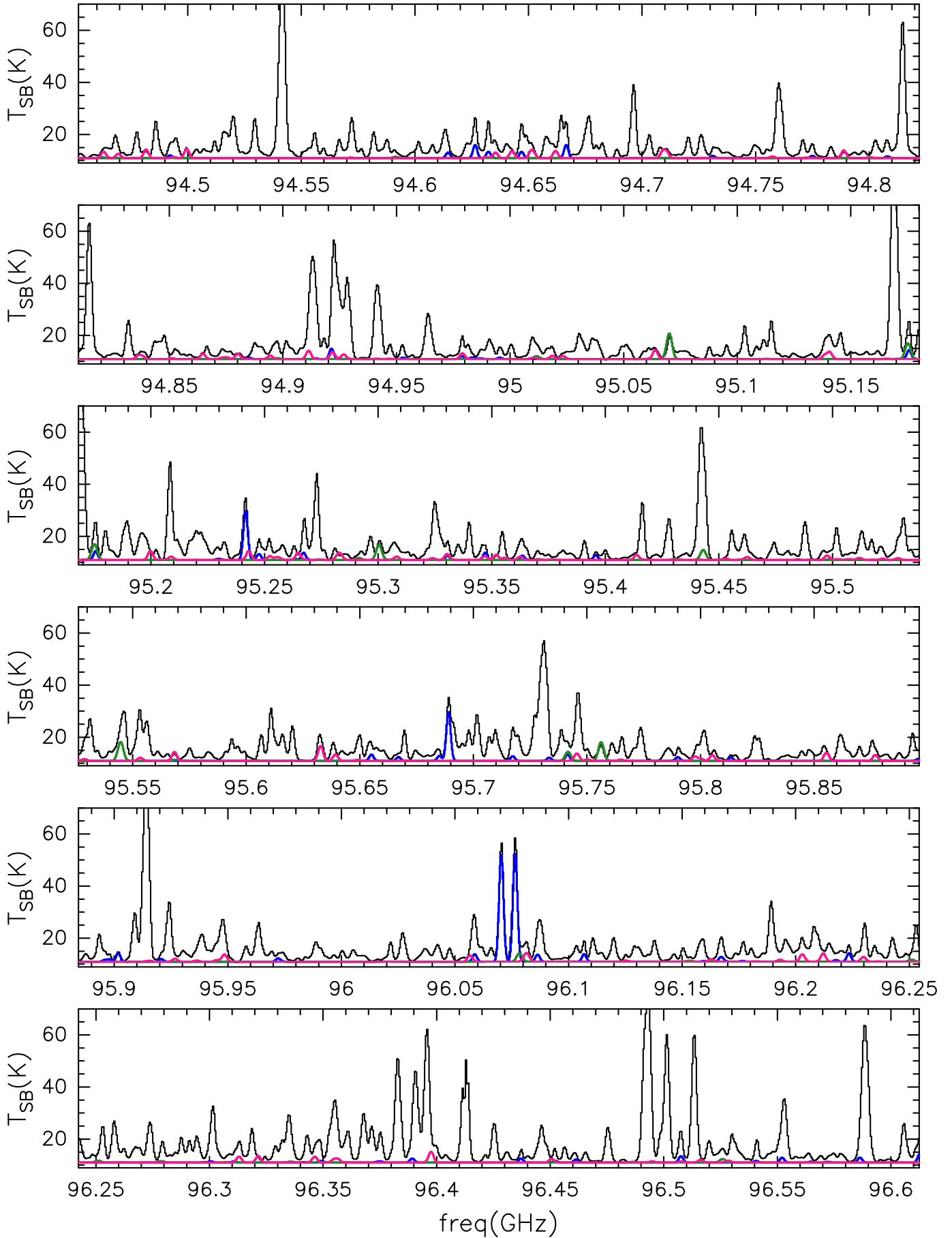


Fig. D.1. continued.

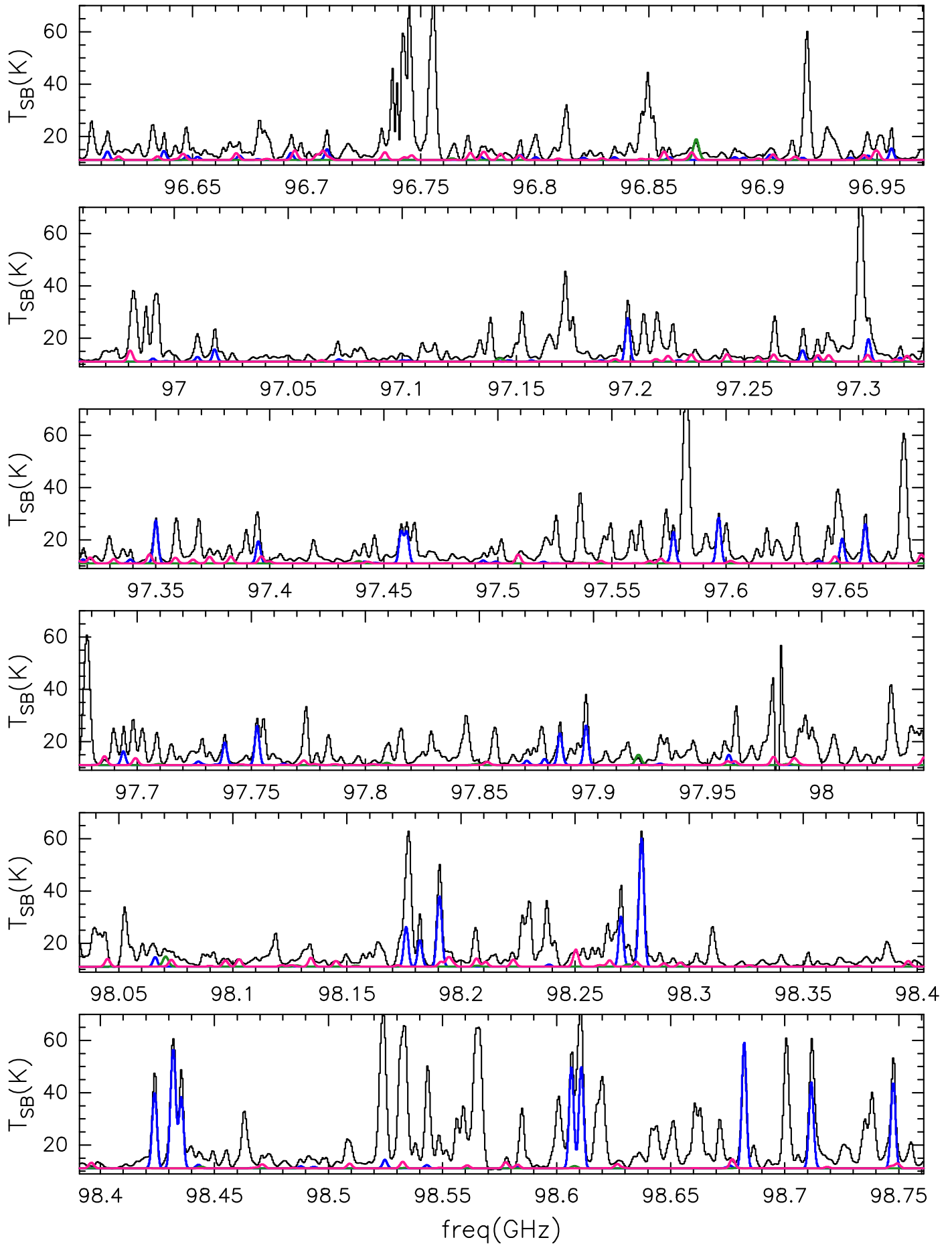


Fig. D.1. continued.

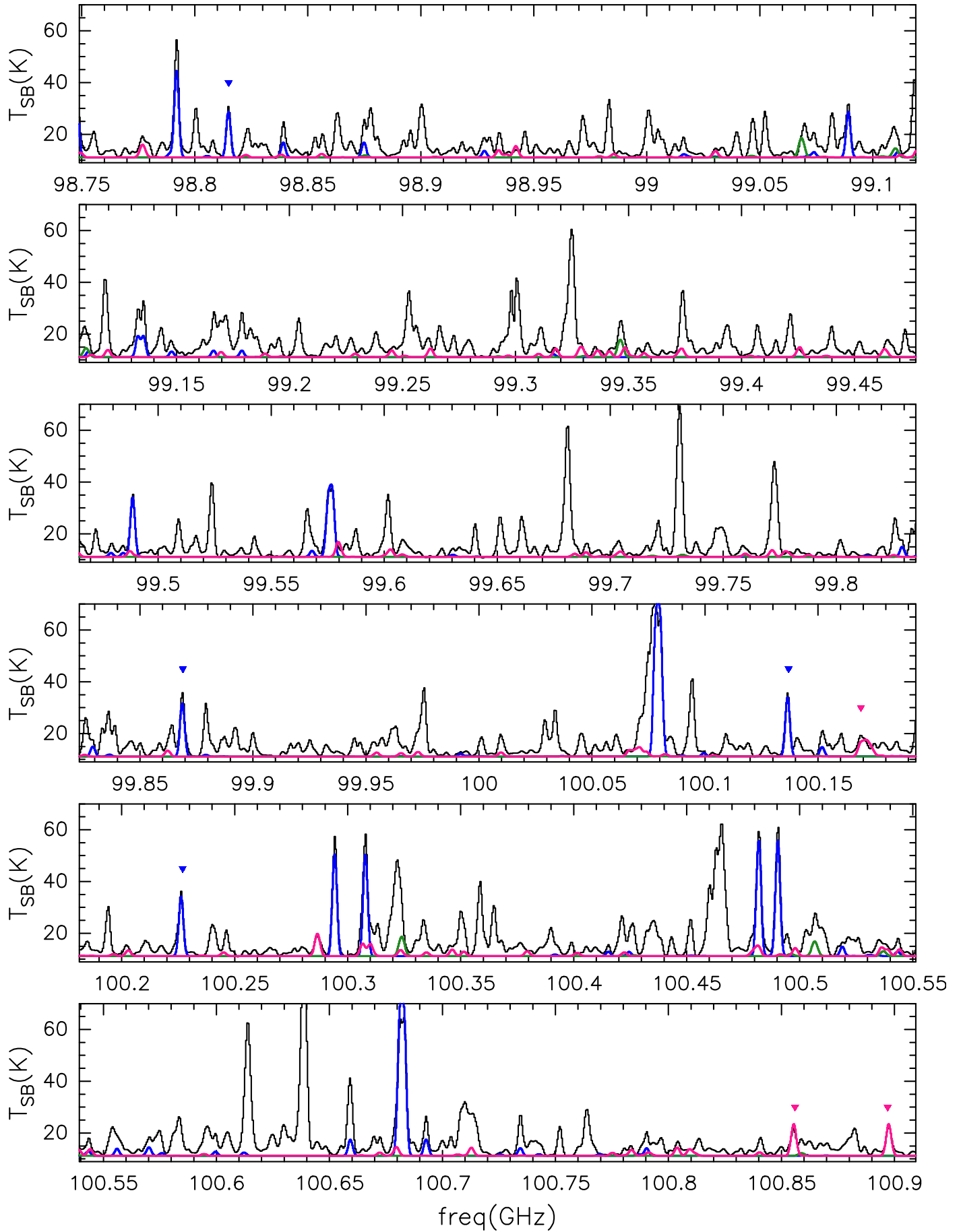


Fig. D.1. continued.

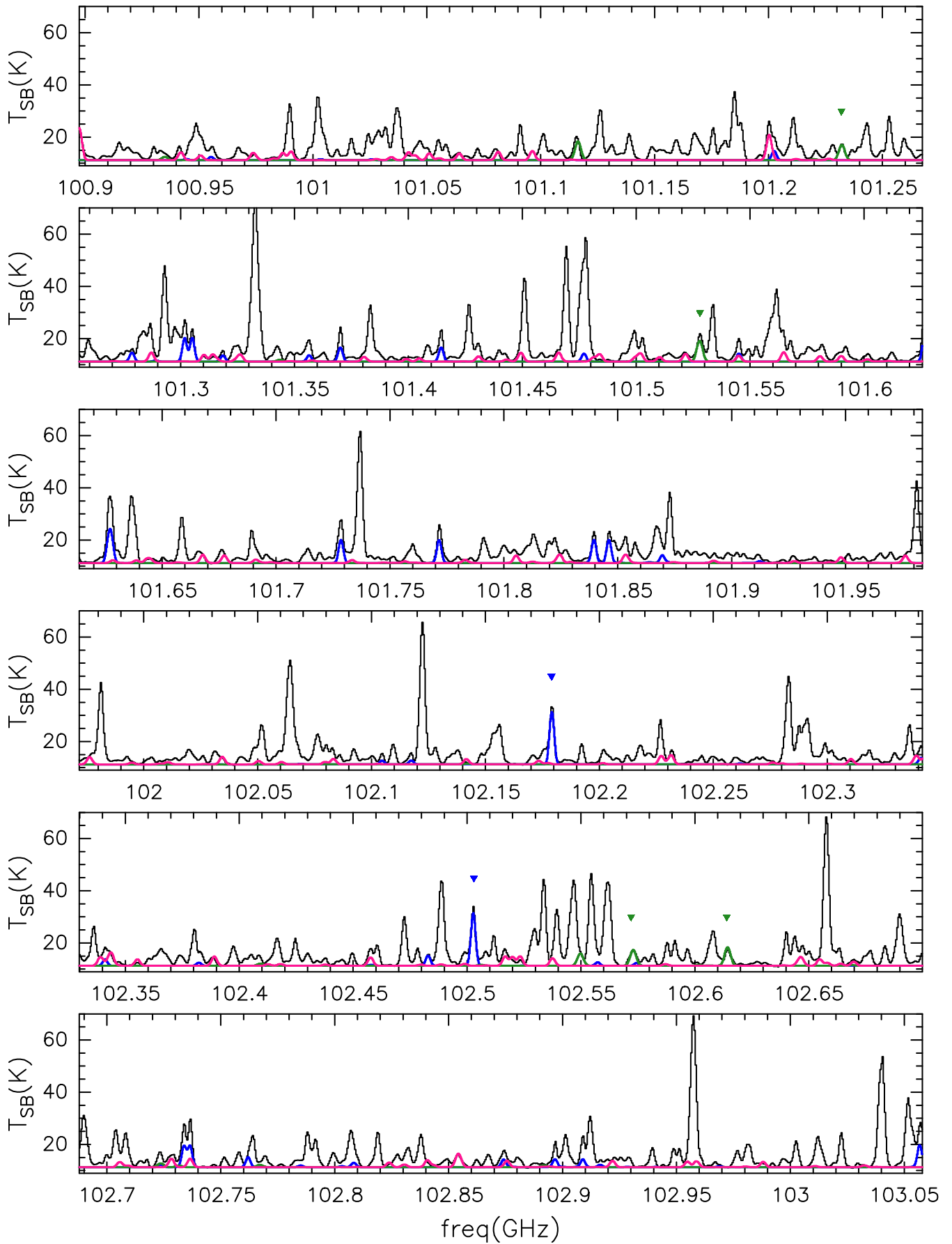


Fig. D.1. continued.

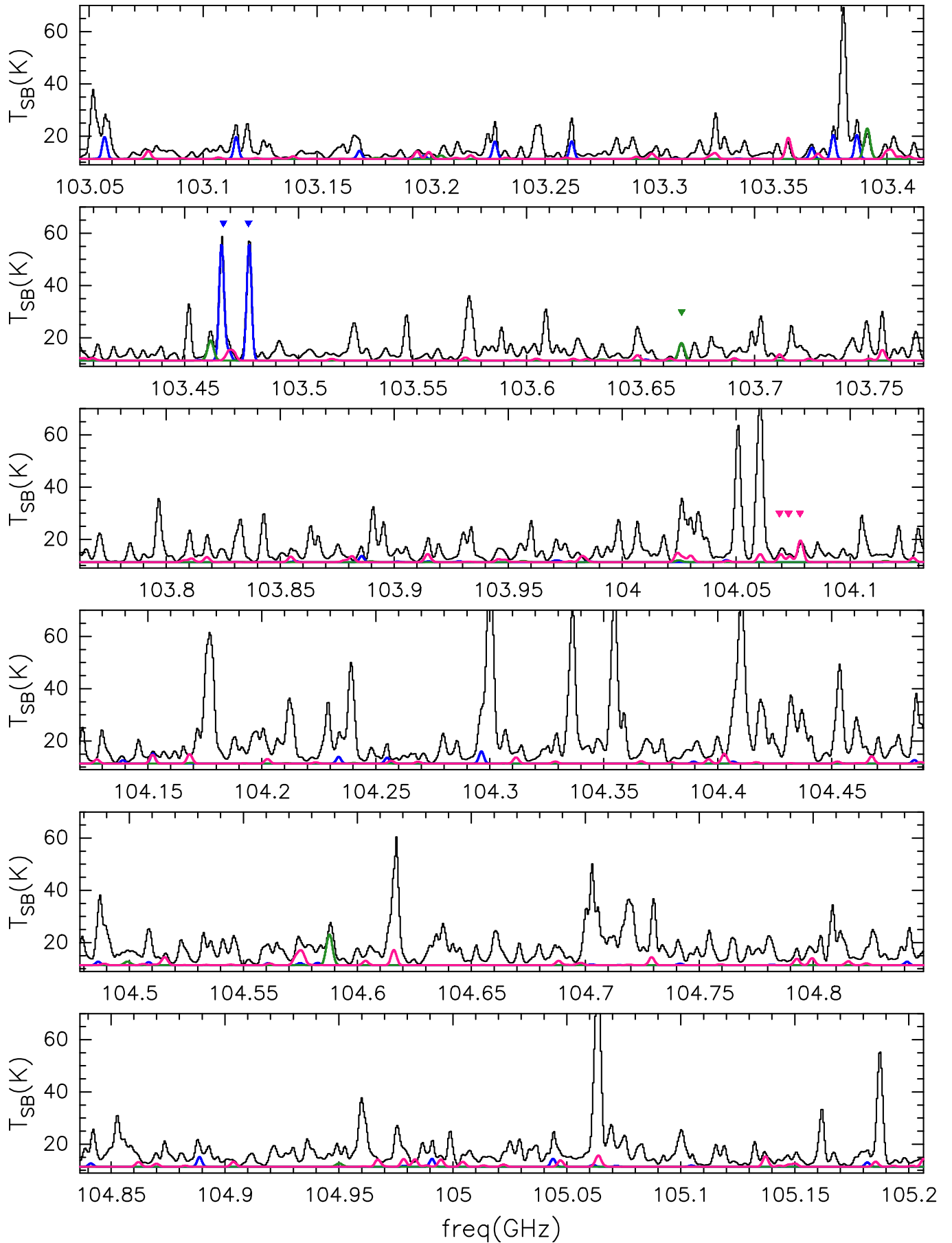


Fig. D.1. continued.

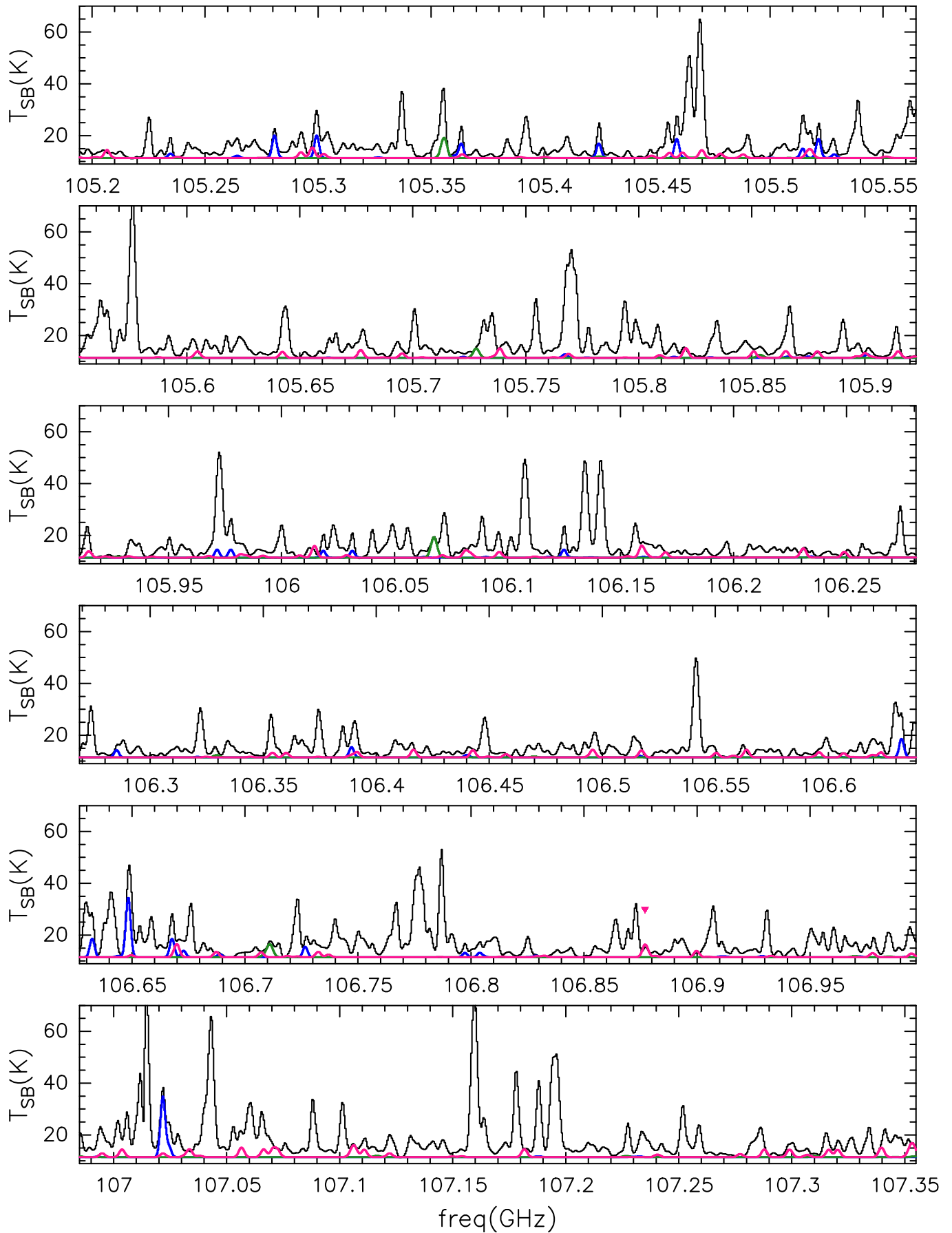


Fig. D.1. continued.

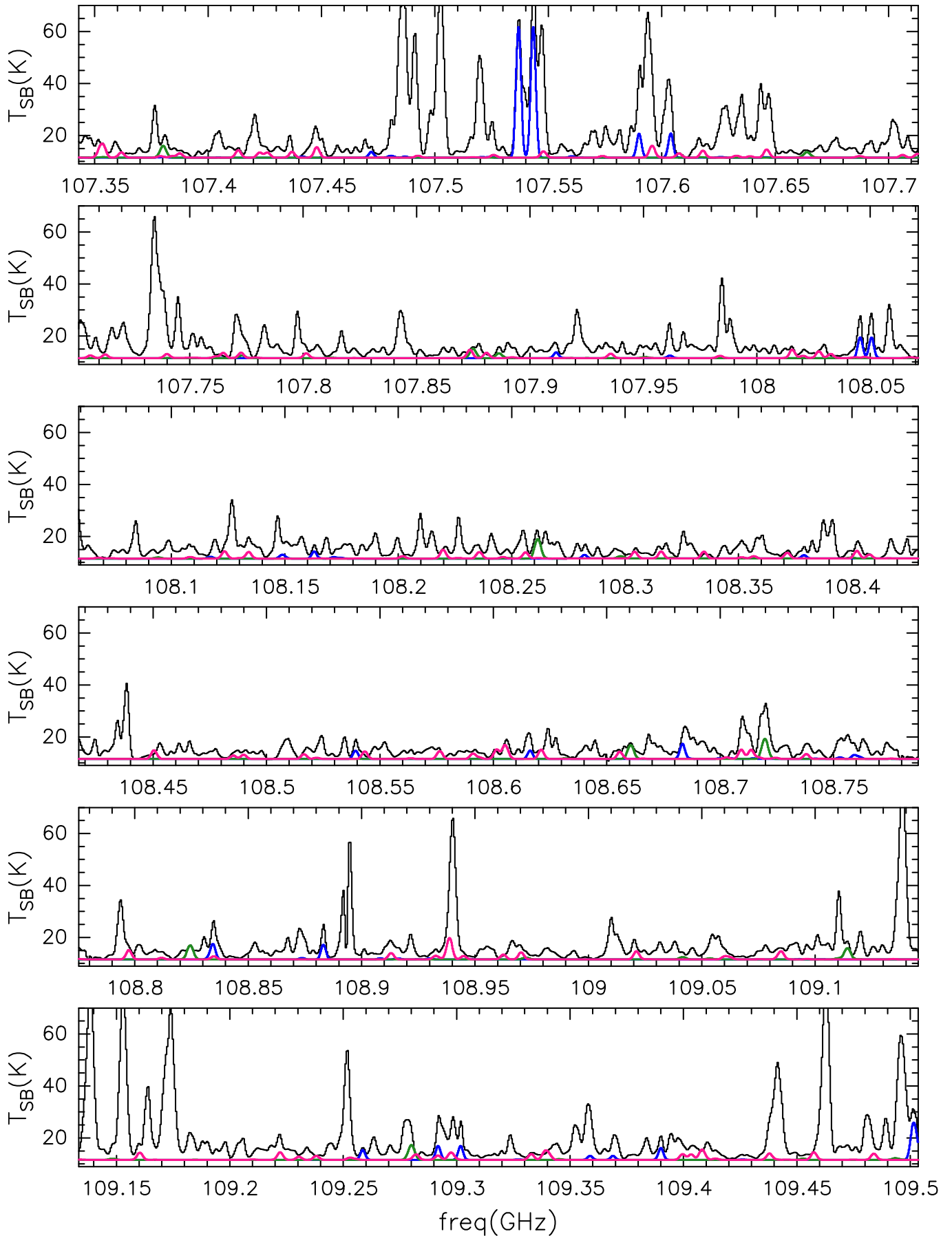


Fig. D.1. continued.

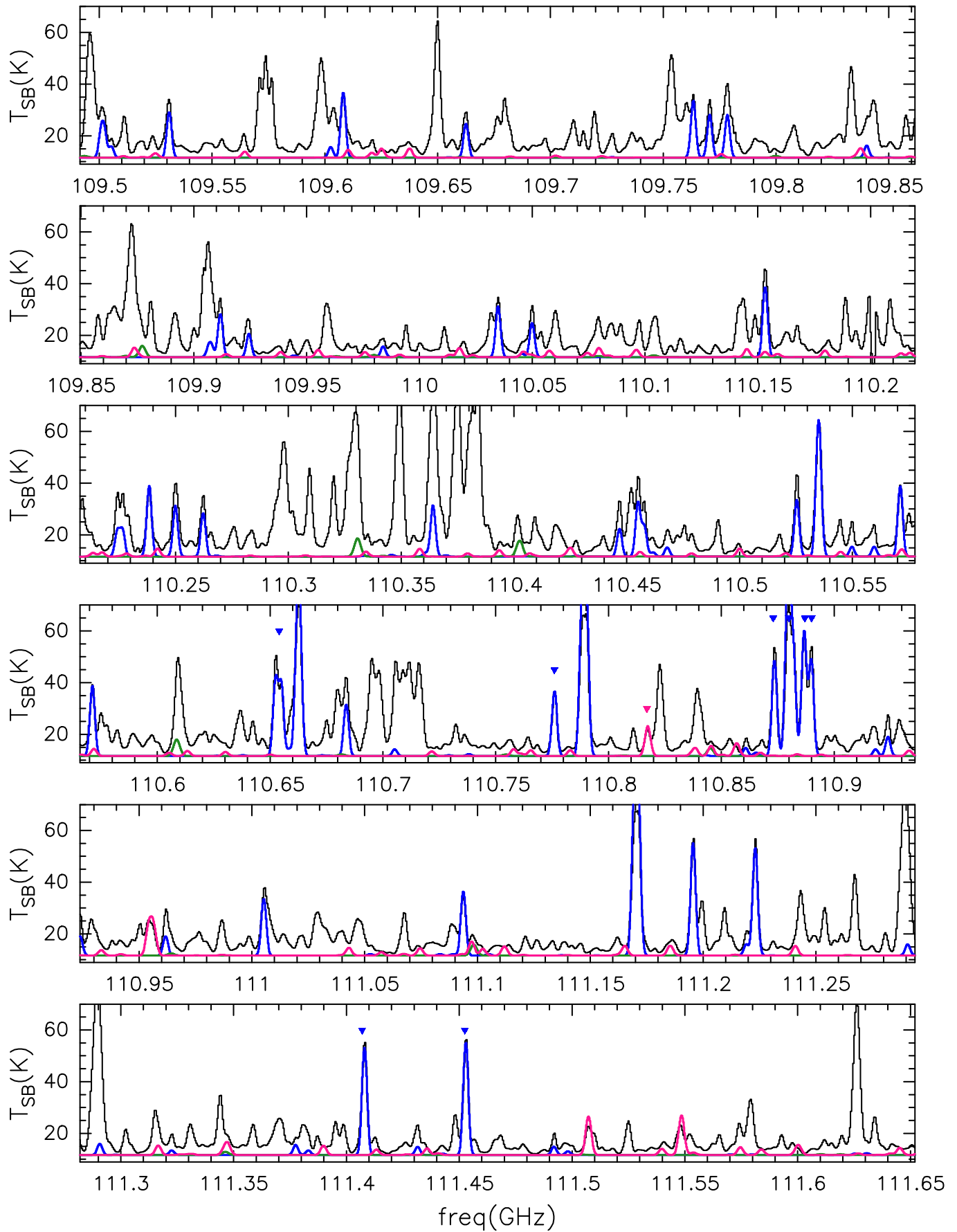


Fig. D.1. continued.

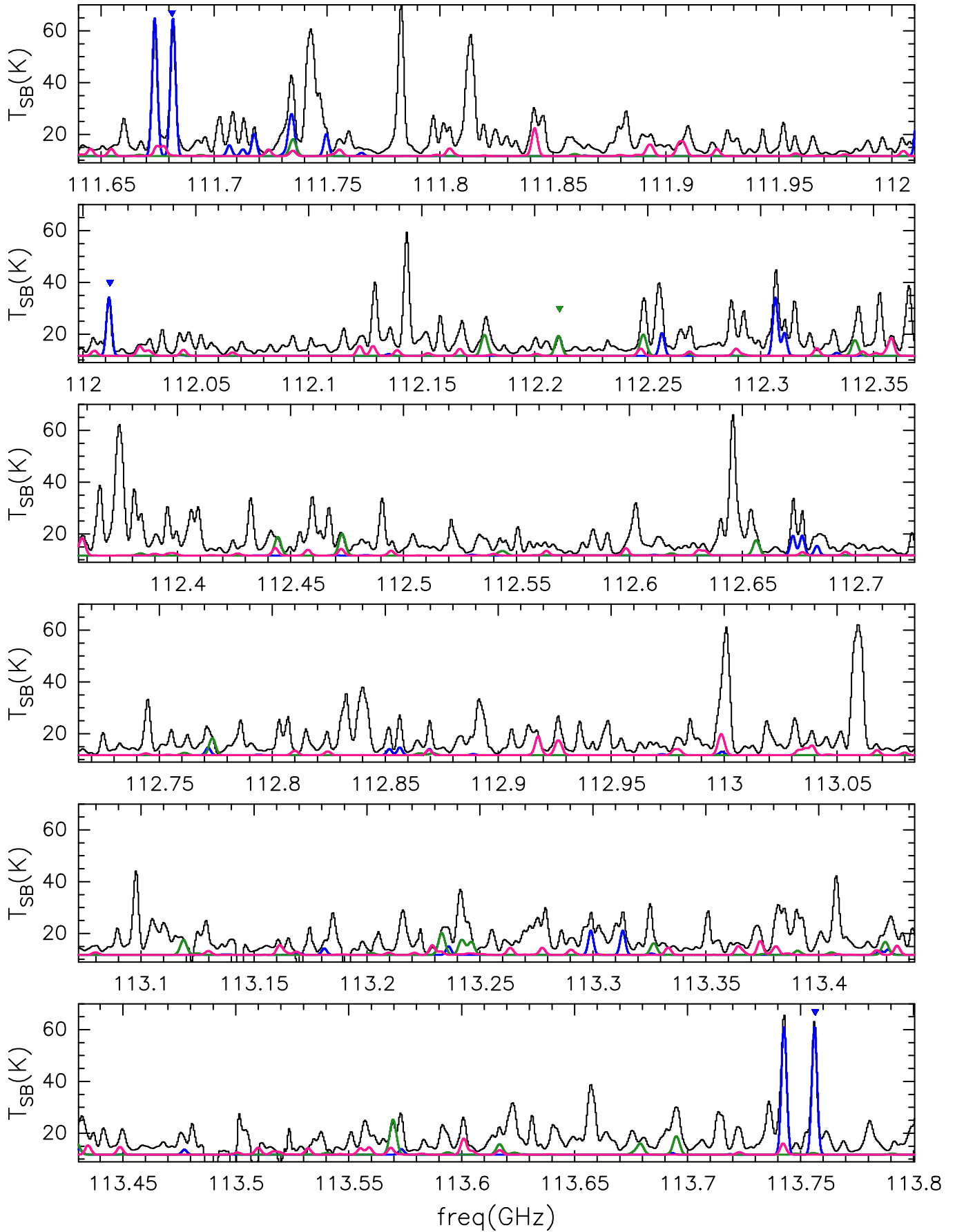


Fig. D.1. continued.

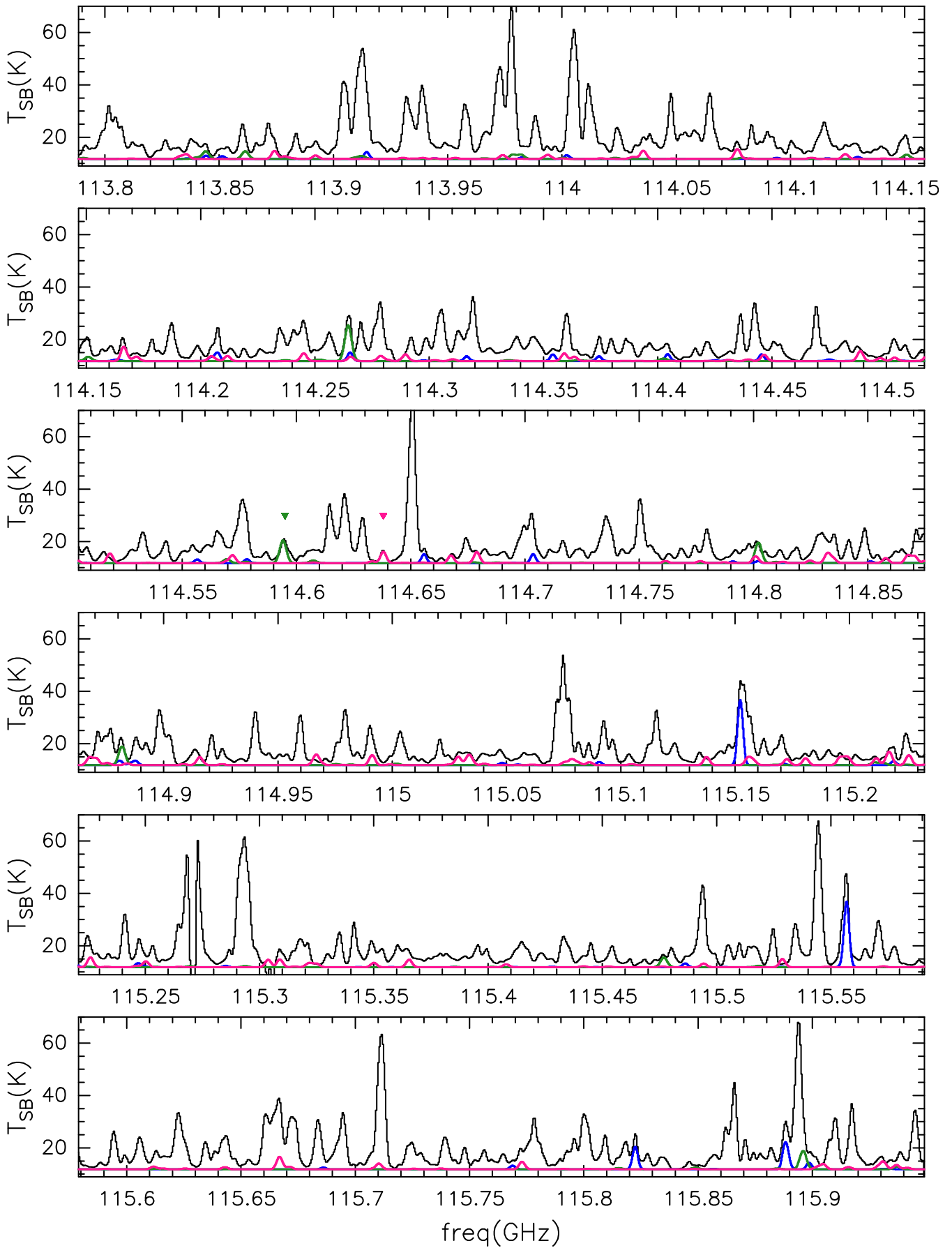


Fig. D.1. continued.