A&A, 708, A201 (2026)
https://doi.org/10.1051/0004-6361/202558050

Astronomy & Astrophysics

Formaldehyde as a densitometer and thermometer in Cygnus-X, the GLOSTAR pilot region, and M8

Utilizing the H_2CO ground-state transition

I. Barlach Christensen[1,⋆,⋆⋆], C. Gieser[2,3], F. Wyrowski[1], H. Nguyen[1], T. D. Hoang[1], V. S. Veena[1],
H. Beuther[3], K. A. Kahle[3], Y. Gong[4], and K. M. Menten[1,⋆⋆⋆]

[1] Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, 53121 Bonn, Germany
[2] Max-Planck-Institut für Extraterrestrische Physik, Giessenbachstrasse 1, 85748 Garching, Germany
[3] Max Planck Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany
[4] Purple Mountain Observatory, and Key Laboratory of Radio Astronomy, Chinese Academy of Sciences, 10 Yuanhua Road, Nanjing 210023, PR China

Received 10 November 2025 / Accepted 19 February 2026

ABSTRACT

Context. Measurements of the physical conditions in molecular clumps are key to our understanding of star formation. Formaldehyde (H_2CO) is a prevalent molecule in these regions, and it can be used as a diagnostic of the physical conditions.

Aims. Here we explore a technique for determining the volume density and gas kinetic temperature in molecular clumps across various evolutionary phases and environments. The ground-state transition of H_2CO has a critical density of $n_{\rm crit} \sim 10^4$ cm^{-3}, allowing us to use this molecule as a densitometer at $n \leq 10^5$ cm^{-3}and to lessen the discrepancy between the measurements between gas densities derived from molecular tracers and those derived from dust observations.

Methods. The clumps in our study were observed with the IRAM 30-m telescope, marking the first extensive survey of the H_2CO ($1_{0,1} - 0_{0,0}$) line across a large sample of sources. These observations were complemented by the H_2CO $J = 3{-}2$ lines, obtained using the APEX telescope. These clumps have been surveyed in three regions, the Cygnus-X giant molecular cloud complex, the GLOSTAR pilot region covering the Galactic plane at longitudes $28° \leq l \leq 36°$, and the molecular cloud associated with the HII regions in the Lagoon nebula (M8).

Results. We analyzed a total of 127 clumps, including 78 from Cygnus-X, 12 from the GLOSTAR pilot region, and 37 from M8. We derived the gas kinetic temperature, volume densities and H_2CO column densities using radiative transfer modeling with `pyradex+emcee` in 102 clumps. We reproduced the observed line intensities in the sources with volume densities $n(H_2) = 5.4 \times 10^4{-}3.8 \times 10^5$ cm^{-3}, gas kinetic temperatures $T_{\rm gas} = 16{-}219$ K, and H_2CO column densities $N(H_2CO) = 6.0 \times 10^{12}{-}1.6 \times 10^{15}$ cm^{-2}.

Conclusions. The gas kinetic temperatures obtained from the non-local thermodynamic equilibrium (LTE) modeling with `pyradex+emcee` agree well with the LTE gas kinetic temperature obtained from the ratio of H_2CO ($3_{0,3} - 2_{0,2}$) and H_2CO ($3_{2,1} - 2_{2,0}$) lines at densities $n(H_2) \leq 10^{5.5}$ cm^{-3}. However, we find that, at higher densities, LTE temperatures derived from this ratio are overestimated by up to 0.5 dex. The volume densities we measured are consistent with the volume densities obtained from dust continuum measurements, thereby probing the bulk of the gas. Furthermore, we find that the volume densities and dust temperatures increase with increasing evolutionary phase. The newly available ground-state transition of H_2CO allows the physical conditions in various phases of star formation to be constrained more effectively.

Key words. astrochemistry – stars: formation – ISM: abundances – evolution – ISM: molecules – submillimeter: ISM

1. Introduction

The physical and chemical conditions during high-mass star formation are poorly understood (see e.g. McKee & Tan 2002; McKee & Ostriker 2007; Krumholz & Bonnell 2009; Tielens 2021). The formation of massive star clusters in particular is believed to take place in dense ($n(H_2) > 10^5$ cm^{-3}) and cold clumps ($T_{\rm kin} < 20$ K) within molecular clouds (e.g. Pillai et al. 2011). Knowing the initial conditions of star formation prevailing in these regions can shed light on the formation mechanisms of their stellar content (Girichidis et al. 2020). Molecular clouds are the largest structures (≥10 pc) from which gas fragments into clumps (size ~1 pc), cores (size ~0.05–0.1 pc), envelopes (size ~300–3000 AU) and disks (size ~10–200 AU) (Pokhrel et al. 2018), eventually accreting onto protostars.

Previous studies have attempted to probe the volume density and kinetic gas temperatures using various molecular tracers, but these methods have limitations. Ammonia (NH_3) inversion lines have been used to determine the gas kinetic temperature (e.g. Walmsley & Ungerechts 1983; Tafalla et al. 2004; Wienen et al. 2012), but the NH_3 abundance is strongly affected by dust evaporation (Yamato et al. 2022; Redaelli et al. 2023). To probe temperature, CH_3CCH, CH_3CN, and CH_3OH have been used,

---

[⋆] Corresponding author: `ibarlach@mpifr-bonn.mpg.de`
[⋆⋆] Member of the International Max Planck Research School (IMPRS) for Astronomy and Astrophysics at the Universities of Bonn and Cologne.
[⋆⋆⋆] Karl M. Menten was unable to witness the completion of this article, having passed away. Nonetheless, his vital contributions to our project remain deeply valued and will always be remembered.

with CH_3OH also being sensitive to density (Giannetti et al. 2017; Lin et al. 2022). The K-level population of CH_3OH is sensitive to the gas temperature (e.g. Cummins et al. 1983; Remijan et al. 2004). Higher J-lines of CH_3OH trace kinetic temperatures, and lower J-lines trace density in low-, intermediate-, and high-mass star-forming regions (Leurini et al. 2004, 2007). Notably, CH_3CN traces hot core temperatures, thus limiting its ability to measure temperatures in cold regions.

Formaldehyde (H_2CO) is a ubiquitous molecule believed to form through successive hydrogenation of CO on icy grain mantles (e.g. Mangum et al. 1990; Watanabe & Kouchi 2002; Woon 2002; Yan et al. 2019) or through the gas-phase reaction of CH_3 with atomic oxygen (e.g. Potapov & Garrod 2024; Punanova et al. 2025). Indeed, Downes et al. (1980) revealed the presence of this molecule via absorption measurements toward numerous HII regions. The H_2CO molecule is a slightly asymmetric rotor molecule that has shown its usefulness in determining the volume density and gas kinetic temperature in a variety of sources due to the collisionally governed K-levels (e.g. Mangum & Wootten 1993; Mangum et al. 2013; Ao et al. 2013; Okoh et al. 2014; Tang et al. 2018, 2021a; Gieser et al. 2021).

Due to the line's frequency (72.838 GHz) at the lower end of the receiver coverage of the 3/4 mm atmospheric window, only a limited amount of observations have been made of the $1_{0,1} - 0_{0,0}$ ground-state transition of para-H_2CO (hereafter H_2CO). Figure 1 shows how the H_2CO $(3_{0,3} - 2_{0,2})/(1_{0,1} - 0_{0,0})$ and H_2CO $(3_{2,1} - 2_{2,0})/(3_{0,3} - 2_{0,2})$ ratios depend sensitively on temperature and density. This line, first detected in the interstellar medium by Akabane et al. (1974; see also Kaifu et al. 1975), has a relatively low critical density of $n_{\rm crit} \sim 10^4$ cm^{-3}. This property allows the molecule to be used as a densitometer at $n \leq 10^5$ cm^{-3}, as it can also trace the envelopes of clumps, which exhibit a distinct physical behavior from the cores, thereby offering a more comprehensive view of how density, its spatial distribution, and the temperature vary with evolutionary stage across different spatial scales. Studies have shown that the ground-state transitions of ortho-H_2CO $(1_{1,0} - 1_{1,1})$ and $(2_{1,1} - 2_{1,2})$ at 6 and 2 cm, respectively, are density tracers (e.g. Mangum et al. 2008; Ginsburg et al. 2011, 2015; Gong et al. 2023). Tang et al. (2017) observed that the fractional abundance of H_2CO remains relatively stable across different phases of star formation. The fractional abundance with respect to H_2 of H_2CO varies by only one order of magnitude during these different phases, with $X(H_2CO) \simeq 10^{-10}$ (Tang et al. 2018).

Previous studies have revealed a significant discrepancy between gas densities derived from molecular tracers and those derived from dust observations. For example, studies using CS (Beuther et al. 2002) or higher J transitions of H_2CO (Tang et al. 2018) find higher gas densities compared to dust-based estimates. This discrepancy can be attributed to the high critical densities of the lines used in these studies. The H_2CO ground-state transition, H_2CO $(1_{0,1} - 0_{0,0})$, now offers a lower critical density line to probe the bulk material of the clumps. This ground-state transition increases the diagnostic utility of H_2CO emission lines as tracers of both temperature and density, enabling the probing of lower-density regions (see Fig. 1).

In this paper, we measure the volume density and gas kinetic temperature using H_2CO toward a sample of sources in the Cygnus-X star-forming region, the GLOSTAR pilot region, and the Messier 8 (M8) region. The combination of rotational transitions with different critical densities allowed us to probe a large range of gas physical conditions. The paper is organized as follows: in Sect. 2, we present the details of the target sources. In Sect. 3, we present the observations and data reduction. In Sect. 4, we analyze the averaged spectra to determine whether different H_2CO transitions are probing the same regions within the clumps. In Sect. 5, we describe the `pyradex+emcee` methodology and analysis. In Sect. 6, we discuss the obtained gas kinetic temperatures, volume densities, and H_2CO column densities, and we explore the physical conditions across the evolutionary phases. Finally, in Sect. 7, we present our conclusions.

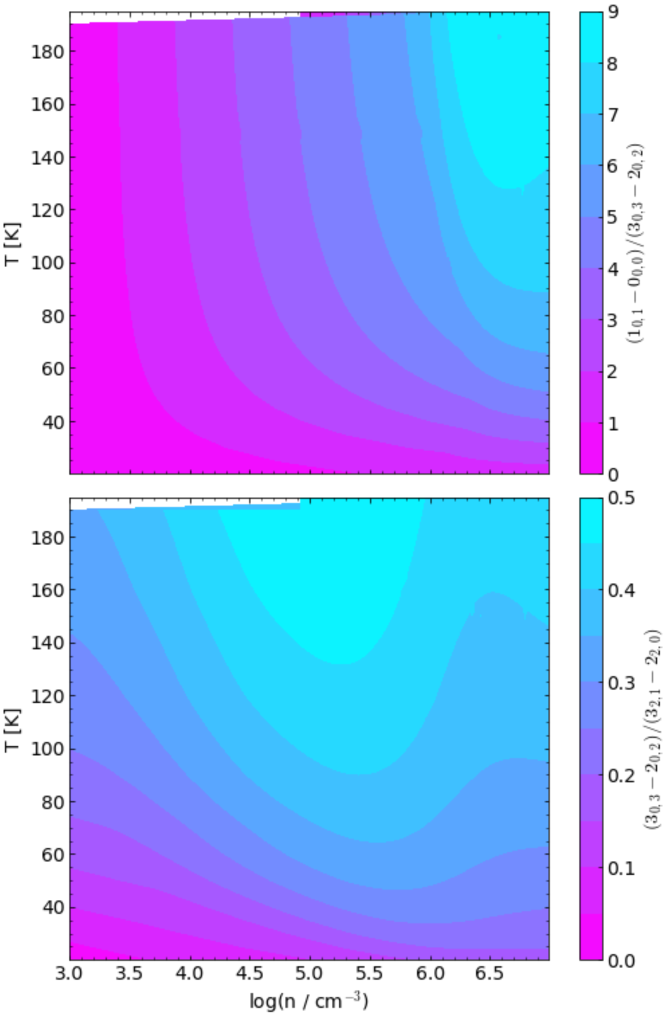


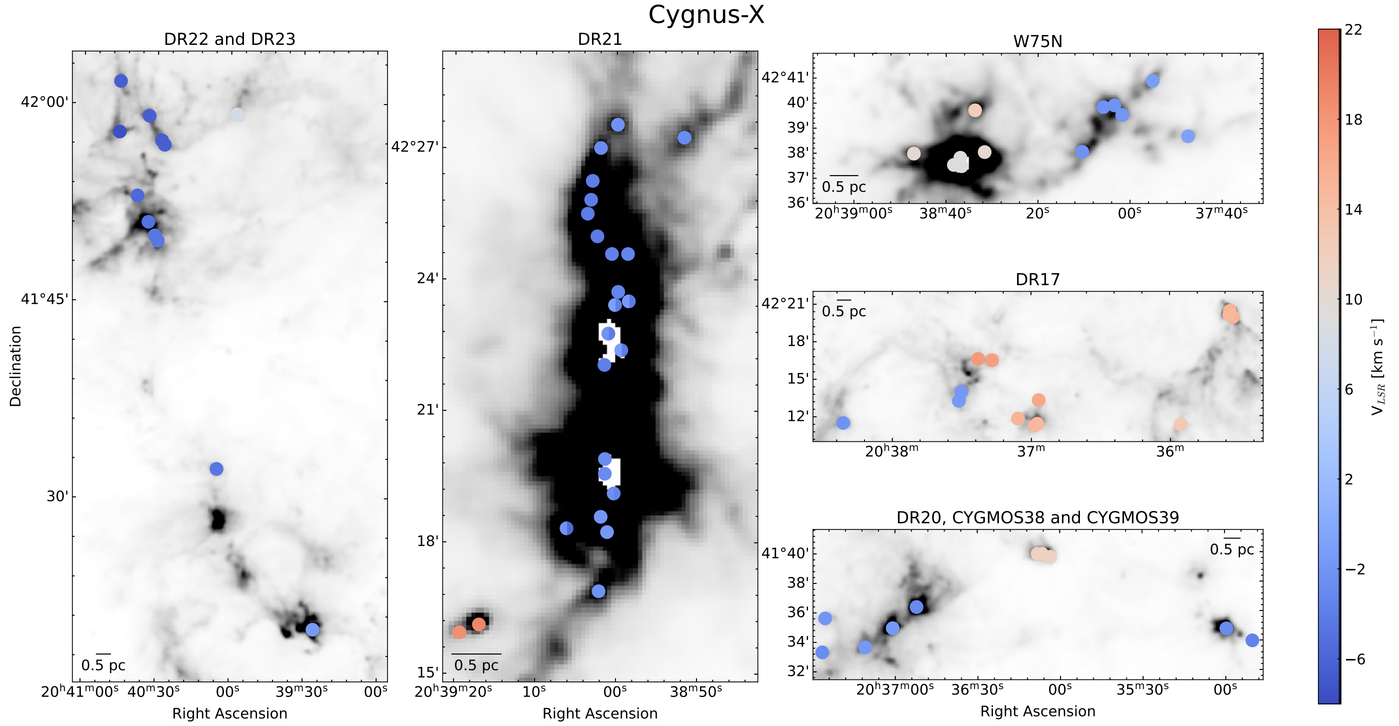
Fig. 1. Line-ratios of H_2CO $(3_{0,3} - 2_{0,2})/(1_{0,1} - 0_{0,0})$ (top) and H_2CO $(3_{2,1} - 2_{2,0})/(3_{0,3} - 2_{0,2})$ as a function of volume densities and gas kinetic temperature computed from pyradex. Bottom plot shows the "bump" at $n(H_2) \geq 10^{5.5}$ cm^{-3}, discussed in Sect 6.1.

2. Source selection

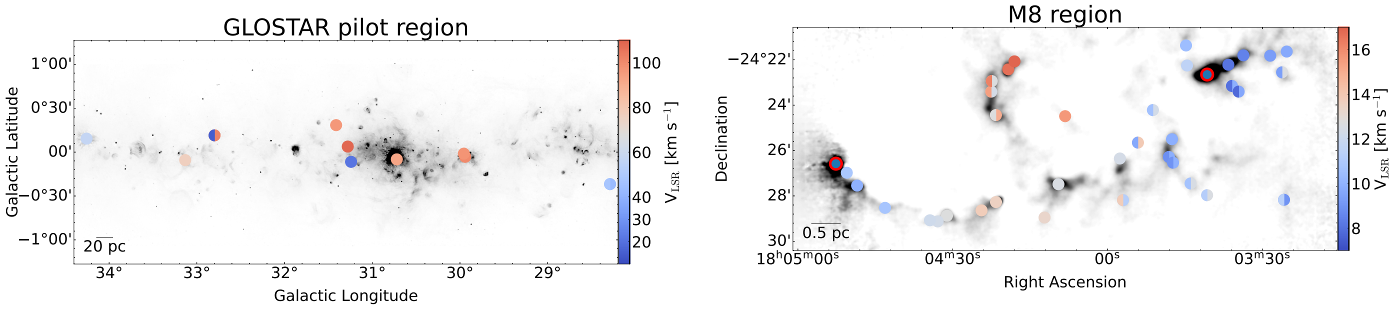
This study focuses on three specific regions: Cygnus-X (as part of the CASCADE project), the GLOSTAR pilot region, and M8. The following section gives an overview of the three target regions.

2.1. Cygnus allscale survey of chemistry and dynamical environments

Cygnus-X is a nearby ($d \sim 1.5$ kpc, Rygl et al. 2012) high-mass star-forming region with OB associations illuminating the cloud complex from the south-west (Knödlseder 2000; Beerer et al. 2010; Wright et al. 2015). Star formation is triggered by strong



(a) The Cygnus-X clumps on the *Herschel* 250 μm map. The clumps are colored based on the source velocity components, with multiple colors indicating more velocity components. The scale bar shows 0.5 pc at a distance of 1.5 kpc.



(b) Radio continuum image of the GLOSTAR pilot region from the VLA + Effelsberg combined 5 GHz data. The colors indicate the line-of-sight velocities. The scale bar shows 20 pc at a distance of 6 kpc, the median distance of the clumps.

(c) The clumps in the 870 μm filamentary structures of M8, adopted from Kahle et al. (2024). The colors indicate the line-of-sight velocities, where multi-colored circles show double velocity components. The scale bar shows 0.5 pc at a distance of 1.3 kpc. The clumps M8HG and M8E are marked with a red circle.

Fig. 2. Locations of clumps in Cygnus-X (top), the GLOSTAR pilot region (bottom left), and M8 (bottom right).

stellar winds, and several clouds are interacting with their ambient HI envelope specifically in the DR21 filament, with clumps at different evolutionary phases along the structure (Motte et al. 2007; Schneider et al. 2010; Hennemann et al. 2012; Schneider et al. 2023). Recently, a survey using the NOrthern Millimeter Array (NOEMA) and the IRAM 30-m telescope has been conducted with the goal to unveil and connect the chemical and dynamical structures of star formation across a wide range of spatial scales, namely Cygnus Allscale Survey of Chemistry and Dynamical Environments (CASCADE) (Beuther et al. 2022). Motte et al. (2007) and Cao et al. (2019) studied several regions in Cygnus-X, and, in this work, 72 clumps have been targeted. These clumps have masses in the range of $\sim 1 \times 10^4 - 2 \times 10^6\ M_\odot$, and represent sources in a variety of evolutionary phases from starless clumps to HII regions (Motte et al. 2007; Cao et al. 2019). The observed positions are shown in Fig. 2a and listed in Table A.1.

2.2. *Global view on star formation in the Milky Way - Pilot region*

The GLObal view on Star Formation in the Milky Way (GLOSTAR) project aims to characterize the unbiased statistical properties of star formation in the Galactic plane (for an overview, see Brunthaler et al. 2021). The GLOSTAR survey utilizes the Jansky Very Large Array (VLA; Perley et al. 2011) and the Effelsberg 100-m telescope[1] to conduct continuum observations covering the frequency range of 4–8 GHz. In addition, the survey includes simultaneous observations of the 6.7 GHz CH_3OH maser line, various radio recombination lines (RRL), and notably the 4.8 GHz H_2CO absorption line. The "GLOSTAR pilot region" of this survey, discussed by Brunthaler et al. (2021),

---

[1] The 100-m telescope at Effelsberg is operated by the Max-Planck Institut für Radioastronomie (MPIfR) on behalf of the Max-Planck Gesellschaft (MPG).

Table 1. Transitions of H_2CO observed with the IRAM 30-m and the APEX 12-m telescopes.

Telescope	H_2CO $J_{Ka,Kc}$	Frequency [GHz]	E_{up}/k [K]	A_{ij} [s^{-1}]	g_{up}	n_{crit} [cm^{-3}]	Detection rate [%] and noise [mK]					
							Cygnus-X 78 sources		GLOSTAR 12 sources		M8 37 sources	
IRAM 30-m	$1_{0,1} \rightarrow 0_{0,0}$	72.838	3.5	8.2(−6)	3	$2.8 \cdot 10^4$ [a]	100%	30	100%	23	95%	30
APEX	$3_{0,3} \rightarrow 2_{0,2}$	218.222	21.0	2.8(−4)	7	$7.8 \cdot 10^5$ [a]	100%	23	100%	24	95%	25
APEX	$3_{2,2} \rightarrow 2_{2,1}$	218.475	68.1	1.6(−4)	7	$3.2 \cdot 10^5$ [b]	83%	96	92%	72	49%	53
APEX	$3_{2,1} \rightarrow 2_{2,0}$	218.760	68.1	1.6(−4)	7	$3.1 \cdot 10^5$ [b]	83%	90	92%	61	65%	46

Notes. The rest frequencies, upper energies, and critical densities of the four observed transitions of H_2CO. The detection rates for each line in each region are also presented. The Einstein coefficients are written in the form of $A(-x) = A \times 10^{-x}$. References. [(a)] Shirley (2015); critical density at $T = 20$ K. [(b)] Order of critical density estimated based on collisional rates of Wiesenfeld & Faure (2013), assuming optically thin conditions and considering p-H_2 as the main collider. Molecular properties were taken from the Cologne Database for Molecular Spectroscopy (CDMS; Müller et al. 2001).

covers 28° < l < 36°, $|b| < 1°$) and its radio continuum emission is shown in Fig. 2b. This 16 square-degree pilot region shows a diversity of extended and compact sources, including HII regions, supernova remnants, and several CH_3OH masers (Medina et al. 2019; Dokara et al. 2021; Nguyen et al. 2022; Dokara et al. 2023). We observed positions in the GLOSTAR pilot region toward which H_2CO absorption was observed in the survey. These are shown in Fig. 2b and listed in Table A.2.

We determined the distance using the radial velocity of the GLOSTAR sources, although this is challenging due to the kinematic velocity distance ambiguity discussed thoroughly in Roman-Duval et al. (2009). We utilize the Parallax-Based Distance Calculator[2] for the GLOSTAR sources, as described in Reid et al. (2019). In the case of G29.95−0.01 and G31.27+0.06, we adopt the trigonometric parallax distances measured by Zhang et al. (2014). The distances are listed in Table A.2.

2.3. Messier 8

The Lagoon Nebula, also known as Messier 8 (M8 hereafter), is a bright HII region illuminated by O- and B-stars. Located at a distance of 1.3 kpc, it is one of the nearest HII regions, characterized by a high UV flux (Damiani et al. 2017; Tiwari et al. 2019). Tothill et al. (2002) identified clumps associated with the molecular cloud. A plethora of molecules are identified in the M8 photodissociation region and 38 dense dusty clumps that surround it, some of which also show photodissociation region signatures, while 38% are likely hosting protostellar objects (Kahle et al. 2024). The clumps in the M8 region were found to have dust temperatures, T_{dust}, ranging from 15.7 to 43.2 K, and H_2 column densities, $N(H_2)$, ranging from 6.7×10^{21} to 6.1×10^{22} cm^{-2}. Kahle et al. (2024) suggest that the remnant gas in the M8 region may be fragmenting due to the radiation pressure of the surrounding O- and B-stars. The data for four H_2CO transitions from 37 clumps within the M8 region in this study were reduced and analyzed by Kahle et al. (2024). The locations of the clumps can be seen in Fig. 2c with positions listed in Table A.3.

In order to determine volume densities, gas kinetic temperatures and H_2CO column densities using a consistent method, we combined data from all three regions. The considered clumps span a wide range of distances, molecular cloud conditions, kinematics, and evolutionary phases. This diversity allowed us to apply our method across a variety of environments, thereby enhancing the robustness of our analysis.

2.4. Comparison sample

The APEX Telescope Large Area Survey of the GALaxy (ATLASGAL) provided a large inventory of ~10 000 dense clumps within the inner Galactic region (Schuller et al. 2009). We compare our results, presented in Appendix C with measurements from the ATLASGAL clumps; sources representative of high-mass clumps in our inner Galaxy at different evolutionary phases (Urquhart et al. 2018).

A sub-sample of 110 clumps, namely the "ATLASGAL Top 100", is designed to contain sources that represent different evolutionary phases of high-mass star formation: starless/prestellar, mid-infrared weak cores, mid-infrared bright cores, and HII regions (König et al. 2017). The ATLASGAL selected TOP 100 sources probe the clumps on the scales of 0.12–1.95 pc (Csengeri et al. 2014; König et al. 2017; Tang et al. 2018), similar to that probed in this study (0.18–1 pc for all three regions). Tang et al. (2018) determined the volume density and gas kinetic temperatures using H_2CO ($3_{0,3} - 2_{0,2}$), ($3_{2,2} - 2_{2,1}$), ($3_{2,1} - 2_{2,0}$), ($4_{0,4} - 3_{0,3}$), ($4_{2,3} - 3_{2,2}$), and ($4_{2,2} - 3_{2,1}$). This study and Tang et al. (2018) both utilize the three $J = 3 - 2$ H_2CO transition, listed above. However, this study additionally used the $J = 1 - 0$ transition, allowing us to study the lower density material, while Tang et al. (2018) used $J = 4 - 3$.

3. Observations

We observed several H_2CO transitions toward molecular clumps in the Cygnus-X, GLOSTAR, and M8 regions described in Sect. 2. An overview of the targeted H_2CO transitions and their detection rates is presented in Table 1.

3.1. IRAM 30-m telescope observations with EMIR 090

We conducted observations with the Institut de Radioastronomie Millimétrique (IRAM; Baars et al. 1987) 30-m telescope, located on the Pico Veleta in the Spanish Sierra Nevada. We used this telescope to observe the ground-state H_2CO ($J = 1$–0) transition. For these observations, we employed the Eight MIxer Receiver (EMIR[3]; Carter et al. 2012) and the Fourier Transform

---

[2] https://www3.mpifr-bonn.mpg.de/staff/abrunthaler/bessel_calc2.0/

[3] http://www.iram.es/IRAMES/mainWiki/EmirforAstronomers

Table 2. Parameters of the IRAM and APEX observations.

Sources	Telescope (Receiver)	θ_B [″]	Tuning [GHz]	SB	η_{MB}	# pointings
Cygnus-X	IRAM (EMIR 090) APEX (nFLASH230)	33 29	72.8 218.5	LSB inner USB	0.83 0.68 – 0.86	72
GLOSTAR	IRAM (EMIR 090) APEX (nFLASH230)	33 29	72.6 218.5	LSB inner USB	0.83 0.77	10
M8	IRAM (EMIR 090) APEX (nFLASH230)	33 29	Spectral line survey Spectral line survey		0.8 0.797	37

Spectrometer FTS200 as backend, which provides a channel width of 195 kHz (corresponding to a velocity interval, $\Delta \upsilon$, of 0.80 km s^{-1}). At the line's rest frequency, 72.8 GHz, the telescope's FWHM beam width is ≈33″.

Cygnus-X was mapped as part of the CASCADE program (Project ID: 145-19 and 031-20; PIs: F. Wyrowski and H. Beuther, Beuther et al. 2022). Observations of the GLOSTAR sources were carried out with the IRAM 30-m telescope on 2016 May 28th and 29th (Project ID: 110-15; PI: H. Nguyen). A bandwidth of 8 GHz per sideband was employed, providing a total frequency coverage of 16 GHz across the Cygnus-X and GLOSTAR sources, specifically 70.2–78.2 GHz in the lower sideband and 85.9–93.9 GHz in the upper sideband. The IRAM 30-m telescope observations of the 72.8 GHz H_2CO ground-state line toward the Cygnus-X sources used on-the-fly mapping. The spectrum of each clump is extracted from a region matching the angular resolution of ∼33″. The telescope pointing and focus were checked with Mars and 1741−038 for the Cygnus-X and GLOSTAR sources, respectively. The clumps in M8 were observed during several runs between 2022 March and June (Project ID: 141-21; PI: F. Wyrowski). The M8 sources were covered as part of a spectral line survey by Kahle et al. (2024), covering a frequency range of 40.3 GHz between 70 and 117 GHz, for which the FWHM beam size varies between 34″ and 21″. The FWHM beam width of the IRAM 30-m telescope[4] θ_B, in arcseconds, is given by $2460/\nu$, where ν is the observed frequency in GHz.

We converted the observed corrected antenna temperatures, T_A^*, which assumed a forward efficiency, $\eta_{\rm eff}$, of 0.95, to main beam temperatures, $T_{\rm MB}$, using a main-beam efficiency, $\eta_{\rm MB}$, that depends on the observing frequency[5]. At 72.8 GHz, the frequency of ground-state H_2CO line, $\eta_{\rm eff} = 0.79$.

3.2. APEX observations with nFLASH230

The H_2CO ($J = 3{-}2$) observations were performed using the 12-m Atacama Pathfinder EXperiment (APEX) telescope (Güsten et al. 2006), located on the Chajnantor plateau in the Atacama desert (Güsten et al. 2006). We carried out single-pointing observations toward the selected 72 clumps in Cygnus-X during 2022 July and August (project ID: M-0109.F-9508C-2022; PI: Ivalu Barlach Christensen) and 10 clumps in the GLOSTAR pilot region during 2024 April and May (project ID: M-099.F-9519A-2017; PI: H. Nguyen and I. B. Christensen). We made use of the nFLASH 230 receiver, tuning the lower sideband (LSB) to 217.538 GHz to cover three H_2CO lines ($3_{0,3} - 2_{0,2}$, $3_{2,2} - 2_{2,1}$, and $3_{2,1} - 2_{2,0}$, see Table 1 for the summary of spectral line properties) for clumps in Cygnus-X and the GLOSTAR pilot region. The observations were performed in position switching mode, with selected clean reference positions for each subregions within Cygnus-X and a fixed offset of (+600″, 0) for each source in the GLOSTAR pilot region. The telescope focus and pointing were checked, depending on the day, with Mars, Jupiter, or Saturn, achieving an estimated pointing accuracy of ∼5″. The frequency range coverage was 200.4–208.4 GHz in the lower sideband and 216.4–224.4 GHz in the upper sideband, with a typical integration time of 5 minutes. However, the five clumps in Cygnus-X DR20 have additional observations of 30 minutes each. Additional integrations were performed for clumps in which the strongest H_2CO ($3_{0,3} - 2_{0,2}$) line was observed but the fainter H_2CO ($3_{2,2} - 2_{2,1}$) and H_2CO ($3_{2,1} - 2_{2,0}$) lines were not detected above 4σ. An overview of the reference positions is presented in Appendix A. An in-depth description of observations and data processing in the spectral survey of M8 is presented in Kahle et al. (2024) (project-ID: M-0107.F-9530C-2021; PI: Karl M. Menten). The frequency coverage of the spectral line survey in M8 clumps spans 58.3 GHz between 210 and 280 GHz.

We converted the APEX T_A^* scale to a $T_{\rm MB}$ scale using $\eta_{\rm eff} = 0.95$ and a main-beam efficiency depending on the date of observations. In the case of Cygnus-X sources, observations of Mars, Jupiter, and Saturn allowed us to estimate the beam efficiencies from observations, and we found $0.68 < \eta_{\rm MB} < 0.86$. Based on Mars observations, we found $\eta_{\rm MB} = 0.77$ for the GLOSTAR sources. For the M8 sources, $\eta_{\rm MB} = 0.80$ was adopted based on Jupiter[6]. An overview of the receiver tuning parameters of the observations of the three sub-samples is presented in Table 2.

We applied a beam correction factor of $\eta_{\rm bf} = \left(\frac{\theta_{\rm APEX}}{\theta_{\rm IRAM}}\right)^2 = 0.73$ to account for the difference in beam sizes. The observations were conducted with the IRAM and APEX telescopes, which were chosen because their beam sizes are comparable, thereby minimizing systematic uncertainties introduced by beam mismatch.

3.3. Data processing of spectroscopic data

All data of observed transitions of H_2CO were reduced using the Grenoble Image and Line Data Analysis Software (GILDAS) software developed by IRAM[7] (Pety 2018). The spectrum of each H_2CO transition for each clump in CASCADE was produced using the Continuum and Line Analysis Single-dish Software

---

[4] https://publicwiki.iram.es/Iram30mEfficiencies

[5] https://web-archives.iram.fr/ARN/feb01/node5.html

[6] The beam efficiency for the M8 sources is adopted from the publicly available line monitoring of APEX telescope receivers https://www.apex-telescope.org/telescope/efficiency/

[7] http://www.iram.fr/IRAMFR/GILDAS

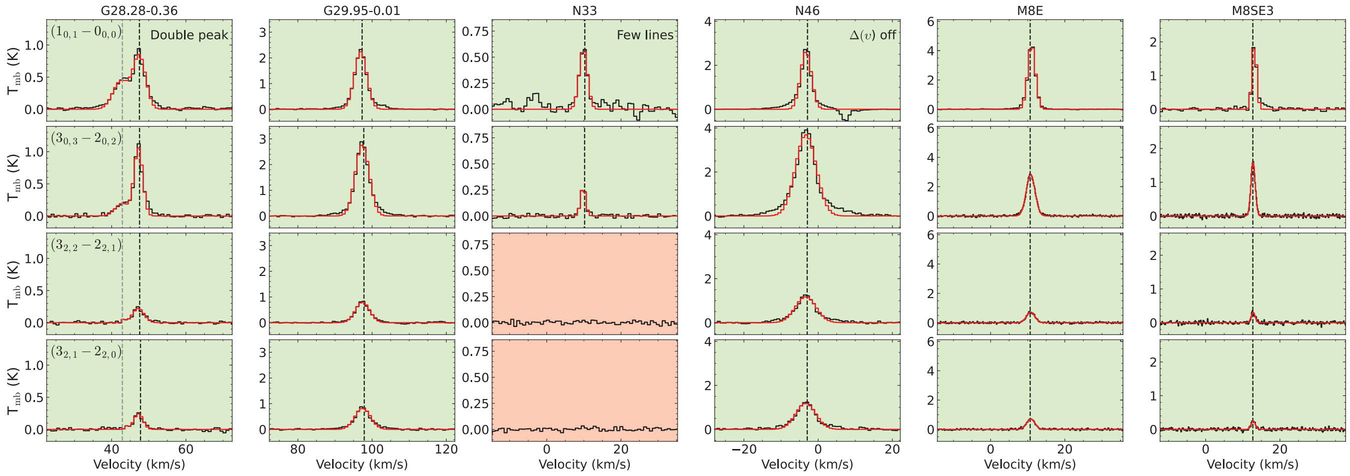


Fig. 3. Spectra of the four H_2CO transitions toward six clumps, two from each region. The background color indicates whether a line is at $\geq 4\sigma$ (green) or $<4\sigma$ (red). The dashed black line indicates the peak velocity of the fit Gaussian for each transition. Examples of clumps with too few lines (N33), differences in $\Delta\upsilon$ (N46), or a double peak (G28.28–0.36) are shown.

(CLASS). All spectra were smoothed to a spectral resolution of the H_2CO $(1_{0,1}-0_{0,0})$ line, that is, 0.8 km s^{-1}. For each H_2CO line, the spectrum of every source is centered on the source's systemic LSR velocity, $\upsilon_{\rm LSR}$, and covers a span of 160 km s^{-1}, $\upsilon_{\rm LSR} \pm 80$ km s^{-1}. This method of data processing was performed on Cygnus-X and GLOSTAR sources. The data processing of M8 is described in Kahle et al. (2024).

4. Results

Using data from the IRAM 30-m and APEX 12-m telescopes, we investigated in detail the H_2CO emission lines observed toward the clumps in Cygnus-X (78 sources), GLOSTAR (12 sources), and M8 (37 sources). We first fit Gaussian line profiles to the spectra of these clumps as input for the subsequent analysis of the physical conditions in the gas.

4.1. Spectral analysis

We analyzed the spectra of each clump focusing on the H_2CO $(1_{0,1}-0_{0,0})$, $(3_{0,3}-2_{0,2})$, $(3_{2,2}-2_{2,1})$, and $(3_{2,1}-2_{2,0})$ lines. We utilized the Gaussian fitting procedure of CLASS to obtain the integrated intensity, $\int T_{\rm MB}$, peak main-beam brightness temperature $T_{\rm MB}$, velocity of peak emission, $\upsilon_{\rm peak}$, and linewidth at the full width at half maximum (FWHM), $\Delta\upsilon$. In the case of clumps with several velocity components, as indicated in Fig. 2, a single Gaussian was fit to each velocity component. These clumps are indicated with "A" and "B" in Appendix B. The maximum number of velocity components is two for a single clump. Six sources in Cygnus-X and two sources in the GLOSTAR pilot region were found to have two velocity components.

Figure 3 shows example spectra of all the H_2CO lines and the Gaussian fit results for two sources each in Cygnus-X, the GLOSTAR pilot region, and the M8 region. The line parameters for the Cygnus-X, the GLOSTAR pilot region, and the M8 region sources are presented in Appendix B. A detailed description of the analysis can be found in Sect. 5.1. The parameters obtained for our Cygnus-X, GLOSTAR pilot region, and M8 sources are presented in Appendix B.

4.2. Detection rates

To estimate the H_2 volume density, gas kinetic temperature, and H_2CO column density, we required at least three emission lines to be detected for each clump. With a detection criterion of $\geq 4\sigma$, the highest detection rates of the formaldehyde lines are the ground-state transition H_2CO $(1_{0,1}-0_{0,0})$ and H_2CO $(3_{0,3}-2_{0,2})$ ($E_{\rm up}/k$ = 3.50 K and $E_{\rm up}/k$ = 20.96 K, respectively). These two emission lines are observed in all 72 clumps in Cygnus-X and ten clumps in the GLOSTAR pilot region. Of these 72 clumps in Cygnus-X, six clumps exhibit two velocity components, resulting in 78 H_2CO measurements. In the GLOSTAR pilot region, eight of the ten clumps have a single velocity component, while 2 have an additional velocity component. Despite possibly originating from the same source, each velocity component is treated as if originating in a separate clump.

The detection rates for Cygnus-X, the GLOSTAR pilot region, and M8 are summarized in Table 1. High detection rates (>80%) of the weaker H_2CO transitions (i.e., $3_{2,2}-2_{2,1}$ and $3_{2,1}-2_{2,0}$) are observed in both Cygnus-X and GLOSTAR clumps. In Cygnus-X, all four H_2CO lines are detected above 4σ in 60 sources. Of the remaining ten clumps, these two strong lines are detected along with either $3_{2,2}-2_{2,1}$ in five sources or $3_{2,1}-2_{2,0}$ in the other five. In the GLOSTAR pilot region, all ten clumps show all four lines above 4σ.

In contrast, M8 shows lower detection rates across all four transitions. The stronger lines are detected in 95% of sources, while the weaker H_2CO $(3_{2,2}-2_{2,1})$ and $(3_{2,1}-2_{2,0})$ transitions are detected in only 49% and 65% of clumps, respectively. Of the 37 M8 clumps observed, 13 were excluded from further analysis due to having fewer than three detected H_2CO transitions.

The high detection rates of H_2CO $(1_{0,1}-0_{0,0})$ and $(3_{0,3}-2_{0,2})$ across all regions suggest the abundant nature of H_2CO in the clumps. The different, and lower, detection rates of the higher excitation lines reveal the underlying excitation conditions (e.g. density and temperature) of these regions, which we investigate further.

4.3. Line parameter comparisons

We examined the line profile of each transition for every clump in Cygnus-X, the GLOSTAR pilot region, and M8 to ensure consistency. Based on comparison of the peak velocity, $\upsilon_{\rm LSR}$, and linewidth, $\Delta\upsilon$, of each measurement, we confirm that all lines are probing similar gas components within each clump.

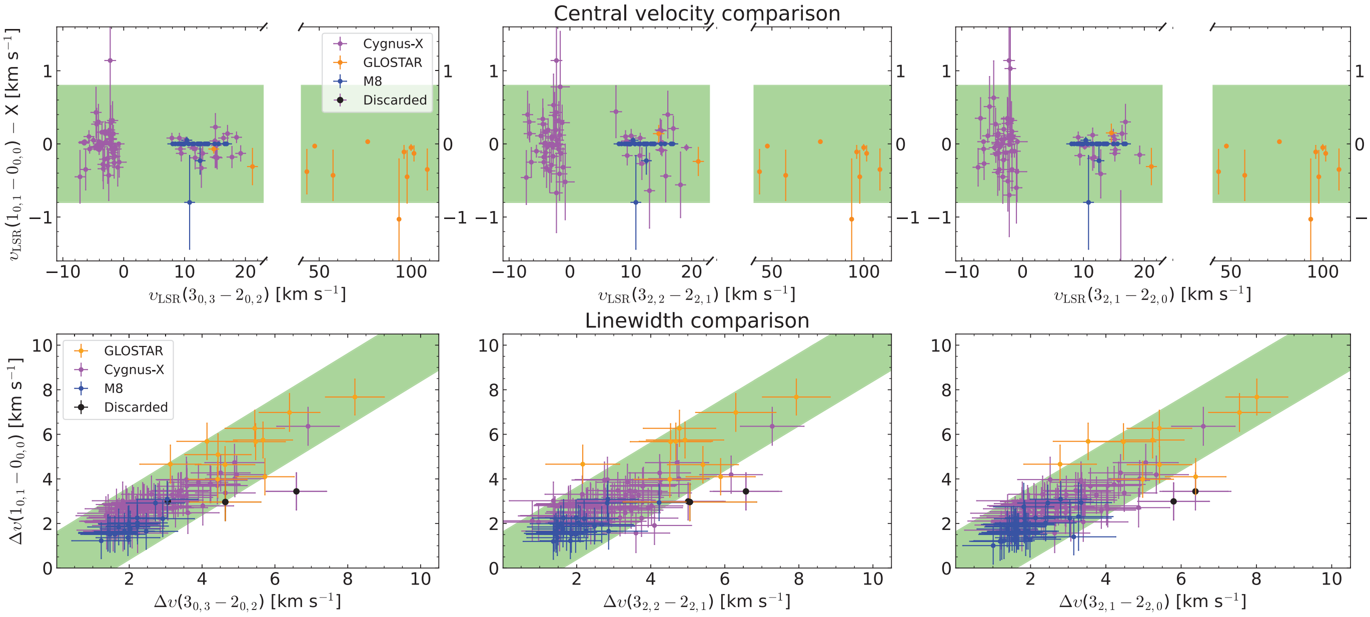


Fig. 4. *Top row*: difference in $\upsilon_{\rm peak}$ between the H_2CO ($1_{0,1}-0_{0,0}$) line and the three higher J-transitions of H_2CO. The X (in the y-axis) denotes the $\upsilon_{\rm peak}$ noted in the x-axis. *Bottom row*: linewidths of the different H_2CO transitions. In top and bottom rows, sources in Cygnus-X are shown in purple, GLOSTAR sources in orange, and M8 sources in blue. The black symbols indicate clumps that do not satisfy the velocity criterion (green shaded region) and were excluded from further analysis.

Figure 4 shows the difference in peak velocity between the four transitions (top), and the linewidth comparison of the four transitions (bottom). The green shaded areas in Fig. 4 indicate the regions of the accepted range. Both of the accepted ranges the υ_{LSR} and $\Delta\upsilon$ are based on the spectral resolution of 0.8 km s^{-1}. The accepted range for the difference between $\upsilon_{\rm peak}$ is ±0.8 km s^{-1}, and for the difference between $\Delta\upsilon$, the range is $\pm 3\times 0.8$ km s^{-1}. The peak velocities are obtained by fitting a Gaussian for each H_2CO transition (see Sect. 4.1). Each $J = 3$–2 transition's peak velocity is compared with the peak velocity of H_2CO ($1_{0,1}-0_{0,0}$). None of the sources in Cygnus-X, GLOSTAR, and M8 are discarded, as all lines are aligned. Similar linewidths further ensure that the transitions originate from the same gas component. Similar to the peak velocity comparison, we also compare the linewidths of $J = 3$–2 with the strongest transition H_2CO ($1_{0,1}-0_{0,0}$). Sources with a linewidth different than $\Delta\upsilon$(H_2CO ($1_{0,1}-0_{0,0}$)) are discarded. Two sources in Cygnus-X and none in the GLOSTAR pilot region and M8 region are discarded due to the difference of linewidths.

H_2CO ($3_{2,2}-2_{2,1}$) and H_2CO ($3_{2,1}-2_{2,0}$) are expected to exhibit similar emission line profiles, as their values for $E_{\rm up}$ and $A_{\rm ij}$ are identical (see Table 1 for reference, Müller et al. 2001). If, as one expects, these transitions originate under the same physical conditions, their line integrated intensities should be comparable. This comparison is shown in Fig. 5, which plots the intensities of the two transitions from the three different regions in this study: within the uncertainties, the integrated intensities of both lines are identical.

To summarize, we discarded sources based on the following criteria: fewer than three lines detected, misaligned velocities within a source, differing linewidths among transitions, and insufficient agreement between line intensities of the H_2CO ($3_{2,2}-2_{2,1}$) and H_2CO ($3_{2,1}-2_{2,0}$) transitions. As a result, a total of 10 sources are discarded in Cygnus-X, none are discarded in the GLOSTAR pilot region, and 13 are discarded in M8. An overview of the sources discarded is presented in Table B.4.

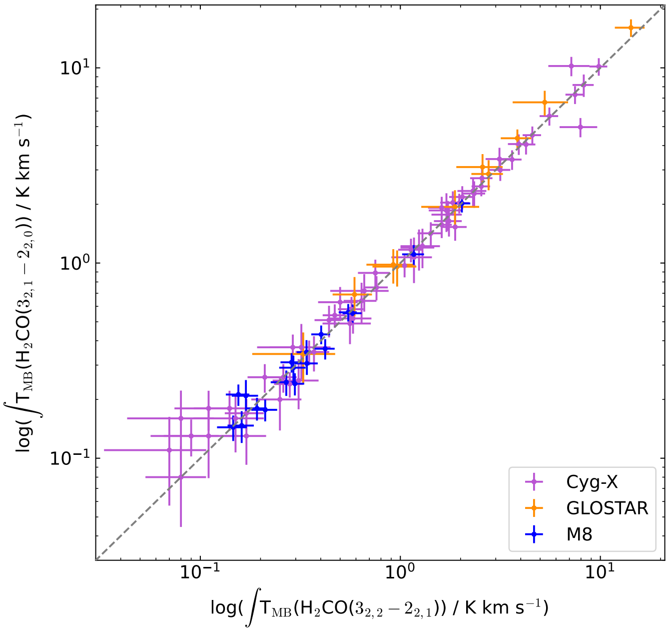


Fig. 5. Comparison of H_2CO ($3_{2,2}-2_{2,1}$) and H_2CO ($3_{2,1}-2_{2,0}$) integrated intensities. The sources in Cygnus-X, the GLOSTAR pilot region, and M8 are shown in purple, orange, and blue, respectively.

5. Analysis

In the following section, we analyze the volume density, gas kinetic temperature and H_2CO column density for clumps in Cygnus-X (68 sources), the GLOSTAR pilot region (ten sources), and the M8 region (24 sources). Furthermore, based on the modeled H_2 volume densities, the calculation of H_2 column densities are described.

5.1. Modeling the physical conditions using H_2CO

The goal was to determine the volume density, gas kinetic temperature, and column density of H_2CO by comparing the line intensities with a model. We utilize `pyradex`[8] under non-local thermodynamic equilibrium (LTE) conditions, the publicly available Python wrapper for `RADEX` (van der Tak et al. 2007), in conjunction with the Python package `emcee`[9]. The following description outlines the `pyradex+emcee` analysis approach for a single clump.

The modeled line parameters for a given set of physical conditions are obtained using `pyradex`. The background temperature, $T_{\rm bg}$, taken to be 2.725 K (the temperature of the cosmic microwave background; Fixsen 2009), and the linewidth are assumed to be fixed, and it is assumed that the clump behaves as an expanding sphere so that the radiative transfer can be treated under the local velocity gradient (LVG) assumption. This assumption implies that there are no non-local radiative effects, and that the line-of-sight molecular column density and H_2 volume density are independent. The `pyradex+emcee` modeling assumes an LVG geometry and yields uniformly distributed physical conditions. We adopted a spherical geometry and set the beam filling factor to $f = 1$ for the modeling, as the beam dilution has been accounted for in the input line intensities. The linewidth was determined as the weighted average of the transitions measured toward the clump, with the condition that the transitions have similar linewidths, indicating that they probe the same gas. The collision rate coefficients for H_2CO were adopted from Wiesenfeld & Faure (2013), who calculated them for 21 temperatures in the range from 10 to 300 K. The collisional partners of H_2CO were assumed to be p- and o-H_2, with an assumed H_2 ortho-to-para ratio of 0.01 (Pagani et al. 2009; Dislaire et al. 2012).

The model parameters are $p = (n(H_2), T_{\rm kin}, N(H_2CO))$. We define a Bayesian approach, with the posterior probability distribution ($\Pr\left(p|I^{\rm data}\right)$) of the model parameters p (notation following Yang et al. 2017):

$$\Pr\left(p|I^{\rm data}\right) = \frac{\Pr(p)\Pr\left(I^{\rm data}|p\right)}{\Pr\left(I^{\rm data}\right)}, \tag{1}$$

defined in terms of the prior probability $\Pr(p)$, the likelihood of the observed intensity of H_2CO given the parameter values, $\Pr\left(I^{\rm data}|p\right)$, and the probability of the data (commonly referred to as "evidence"), $\Pr\left(I^{\rm data}\right)$, serving as a normalizing factor. The likelihood, assuming the noise is independent, is the product of Gaussian probability density functions:

$$\Pr\left(I^{\rm data}|p\right) = \prod_i \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{\left(I_i^{\rm data} - I_i^{\rm model}(p)\right)^2}{2\sigma_i^2}\right). \tag{2}$$

Here, σ_i is the respective error of the observed intensity of H_2CO in $I_i^{\rm data}$, and $I_i^{\rm model}$ is the respective modeled intensities of H_2CO with `pyradex`. We used the logarithmic form of Eq. (2). To ensure that the parameters remain within reasonable astrophysical conditions while still allowing freedom to explore the parameter space, boundaries are set for the ranges of $\log(n(H_2)/{\rm cm}^3)$, $\log(T_{\rm kin}/{\rm K})$, and $\log(N(H_2CO)/{\rm cm}^2)$. These boundaries are defined as $\log(n(H_2)) = 0 - 7$, $\log(T_{\rm kin}) = 1 - 2.47$ (constrained by the availability of collisional rate coefficients; see above) and $\log(N(H_2CO)/{\rm cm}^2) = 10–17$.

We utilized 400 "walkers" to explore the parameter space. We designated a "burn-in" stage consisting of 100 iterations. We used `scipy.curve_fit` together with `pyradex` to obtain a preliminary fit of the physical model parameter, generating an initial position in the parameter space. After the first 100 iterations, the initial position of the parameter space was "forgotten", and the following stage ("walking" stage) consisted of 1000 iterations sampling the posterior distribution.

The H_2 volume density, gas kinetic temperature, and H_2CO column densities are obtained by the median of the posterior probability distribution for each parameter, and the maximum posterior probability within the 16 and 84% quantiles ($= \pm 1\sigma_{\rm model}$). These modeled parameters are presented in Table C.1.

5.2. H_2 volume density, gas kinetic temperature, and H_2CO abundance

Figure 6 shows the distribution of the H_2 volume density, gas kinetic temperature, H_2CO column density, and H_2CO abundance of sources in Cygnus-X, the GLOSTAR pilot region, and M8. We compared the results of this study with the ATLASGAL Top 100 sources obtained by Tang et al. (2018). We accept the modeled physical parameters with a relative error tolerance of $\sigma_{\rm model,rel} \leq 50\%$, where $\sigma_{\rm model}$ is obtained from the posterior distribution. The H_2 volume densities comprising sources in Cygnus-X, the GLOSTAR pilot region, and the M8 region, span a range of $n(H_2) = 5.4 \times 10^4–3.8 \times 10^5$ cm^{-3}. Figure 1 illustrates how the H_2CO $(3_{0,3} - 2_{0,2})/(1_{0,1} - 0_{0,0})$ and H_2CO $(3_{2,1} - 2_{2,0})/(3_{0,3} - 2_{0,2})$ ratios, computed with `pyradex`, vary as a function of H_2 volume density and temperature. The gas kinetic temperatures vary within the range of $T_{\rm gas} = 16–219$ K, while the H_2CO column densities span the range of $N(H_2CO) = 6.0 \times 10^{12}–1.6 \times 10^{15}$ cm^{-2}.

We estimated the H_2 column density, with the line-of-sight length, L, assuming a spherical structure of the clumps. The ratio of the column density and volume density yields the line-of-sight scale length

$$L = \frac{N(H_2)}{n(H_2)}. \tag{3}$$

The size of the clumps was taken as the angular resolution, $\theta_{\rm APEX} = 29''$. To determine the scale length, L, of a source, it is essential to have well-determined distances. The clumps in Cygnus-X have sizes of 0.21 pc, and those of the clumps in the M8 region are 0.18 pc, as the distances to the entire cloud of Cygnus-X and M8 are constant (Rygl et al. 2012; Damiani et al. 2017). The clumps of the GLOSTAR pilot region have varying distances (as described in Sect. 2.2; Reid et al. 2019), resulting in clump sizes spanning 0.26–1.72 pc. Utilizing this H_2 column density, we obtained the fractional abundance of H_2CO. The mean of $X(H_2CO)$ in all three catalogs is $\geq 10^{-10}$ (see the rightmost part of Fig. 6 for the distribution). We obtain H_2CO abundances comparable to those in dense cores, photodissociation regions, and cloud scales, ranging from $\sim 10^{-11}$ to $\sim 10^{-9}$ (Guzmán et al. 2013; Gerin et al. 2024). We also find H_2CO abundances higher than 10^{-9}, in the Cygnus-X region.

---

[8] https://github.com/keflavich/pyradex

[9] https://emcee.readthedocs.io/en/stable/

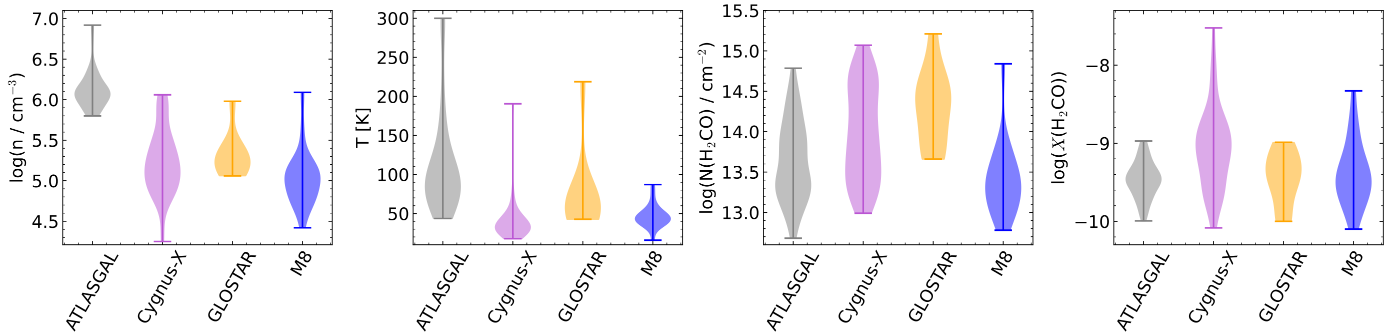


Fig. 6. Distribution of (left to right) the H_2 volume density, gas kinetic temperature, the H_2CO column density, and the fractional abundance. The sources in Cygnus-X, the GLOSTAR pilot region, and M8 are shown in purple, orange, and blue across all distributions, respectively. Similar high-mass sources from the ATLASGAL Top 100, measured using higher J-transitions of H_2CO by Tang et al. (2018) are depicted in gray.

6. Discussion

Here, we discuss the results of all the clumps (102 in total) in the three regions and their obtained physical conditions. We discuss the obtained gas kinetic temperatures and volume densities further in Sects. 6.1 and 6.2.

The higher J-transitions analyzed by Tang et al. (2018) probe higher critical densities, with the critical density of the additional transition ($n_{\rm crit}$(H_2CO ($4_{0,4} - 3_{0,3}$))$\sim 10^6$ cm^{-3}) being an order of magnitude higher than that of the highest J-transition used in this study ($n_{\rm crit}$(H_2CO ($1_{0,1} - 0_{0,0}$)) $\sim 10^5$ cm^{-3}). We find the volume densities of sources in Cygnus-X, the GLOSTAR pilot region, and the M8 region to be an order of magnitude lower than the ATLASGAL sources. Furthermore, the upper levels of the higher J-transitions are higher than that of the J-transitions used in this study. The gas kinetic temperatures are found to be slightly higher, albeit overlapping with some higher temperature clumps in Cygnus-X and GLOSTAR pilot region.

6.1. LTE and non-LTE gas kinetic temperatures

In this section, we present a comparison of the gas kinetic temperature obtained using a non-LTE analysis using `pyradex+emcee` (described in Sect. 5.1) and $T_{\rm kin}$ values derived from the H_2CO ($3_{0,3} - 2_{0,2}$) to H_2CO ($3_{2,1} - 2_{2,0}$) ratio under the LTE assumption. Mangum & Wootten (1993) and Tang et al. (2018) found a dependence of the LTE gas kinetic temperature and the H_2CO ($3_{0,3} - 2_{0,2}$) and H_2CO ($3_{2,1} - 2_{2,0}$) integrated intensity ratio:

$$T_{\rm LTE} = \frac{47.1}{\ln\left(0.556 \times \frac{I(3_{0,3}-2_{0,2})}{I(3_{2,1}-2_{2,0})}\right)} \quad {\rm K}. \tag{4}$$

Figure 7 plots the non-LTE gas kinetic temperature versus the H_2CO ($3_{0,3} - 2_{0,2}$) and H_2CO ($3_{2,1} - 2_{2,0}$) integrated intensity ratio for the sources in our target samples and the ATLASGAL sources discussed by Tang et al. (2018).

The ATLASGAL sources exhibit higher kinetic temperatures compared to the Cygnus-X and M8 sources, but their kinetic temperatures overlap with those sources in the GLOSTAR pilot region (see Fig. 6). The clumps in the GLOSTAR pilot region trace the envelopes of HII regions, and therefore likely trace heated gas. Furthermore, we expanded the relation toward lower temperatures. Tang et al. (2018) discussed that at higher densities, the gas kinetic temperatures exceeding 60 K obtained with this ratio might be overestimated at volume densities of 10^{5-7} cm^{-3}.

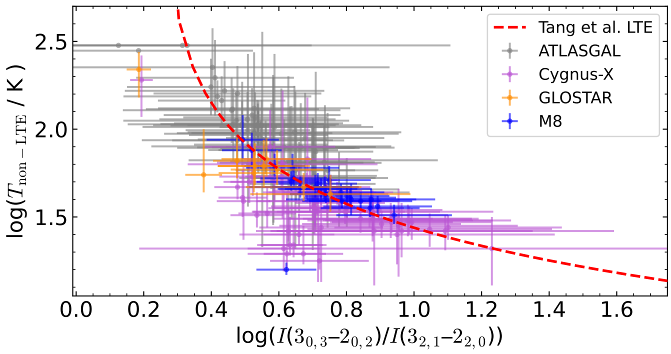


Fig. 7. Gas kinetic temperature of sources in Cygnus-X (purple), the GLOSTAR pilot region (orange), and M8 (blue) using `pyradex+emcee`. The gray points show the temperatures obtained for the ATLASGAL sources by Tang et al. (2018). The dashed red line shows the LTE gas kinetic temperature function presented in Eq. (4).

The LTE gas kinetic temperatures calculated using Eq. (4) are higher than non-LTE measurements as the volume density increases (see Fig. 8). Tang et al. (2018) discussed a large “bump” in the H_2CO ($3_{0,3} - 2_{0,2}$) and H_2CO ($3_{2,1} - 2_{2,0}$) line intensity ratios in their temperature-volume density plot (see their Fig. 4). The bump begins at $n(H_2) \simeq 10^{5.5}$ cm^{-3}. This bump suggests that for a constant gas kinetic temperature and constant H_2CO ($3_{0,3} - 2_{0,2}$) and H_2CO ($3_{2,1} - 2_{2,0}$) line intensity ratio, no unique volume densities can be determined. The bump occurs because the H_2CO ($3_{2,1} - 2_{2,0}$) line intensity increases more rapidly than the line intensity of H_2CO ($3_{0,3} - 2_{0,2}$) for increasing densities $n(H_2) > 10^5$ cm^{-3}. Tang et al. (2018) noted that the bump becomes more pronounced as the line ratios increase. Thus, the increasing difference between LTE and non-LTE gas kinetic temperature found in Fig. 8 is due to the LTE gas kinetic temperature calculation not accounting for the rapid increase of H_2CO ($3_{2,1} - 2_{2,0}$) line compared with H_2CO ($3_{0,3} - 2_{0,2}$).

6.2. Volume densities

We investigated the obtained volume densities from our set of H_2CO lines, and we compared them with the gas densities obtained from dust continuum measurements. Figure 9 shows the comparison between the volume densities obtained from the H_2CO emission lines and the 1.2-mm continuum emission, as reported by Motte et al. (2007) for Cygnus-X and by Tothill et al. (2002) for M8. Figure 9 reveals that the regions where

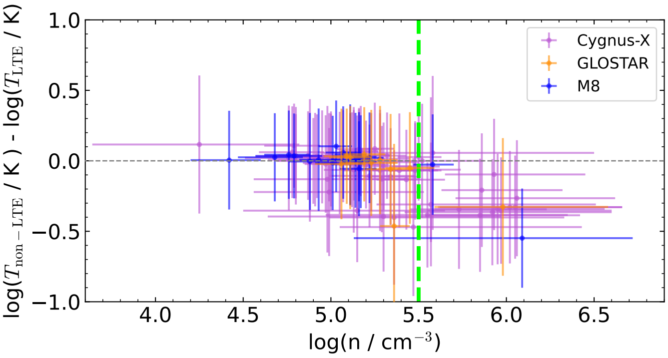


Fig. 8. Difference between the gas kinetic temperature of sources in Cygnus-X (purple), the GLOSTAR pilot region (orange), and M8 (blue) using `pyradex+emcee` compared with the LTE gas kinetic temperature function presented in Tang et al. (2018). The dashed green line marks $n(H_2) = 10^{5.5}$ cm^{-3}.

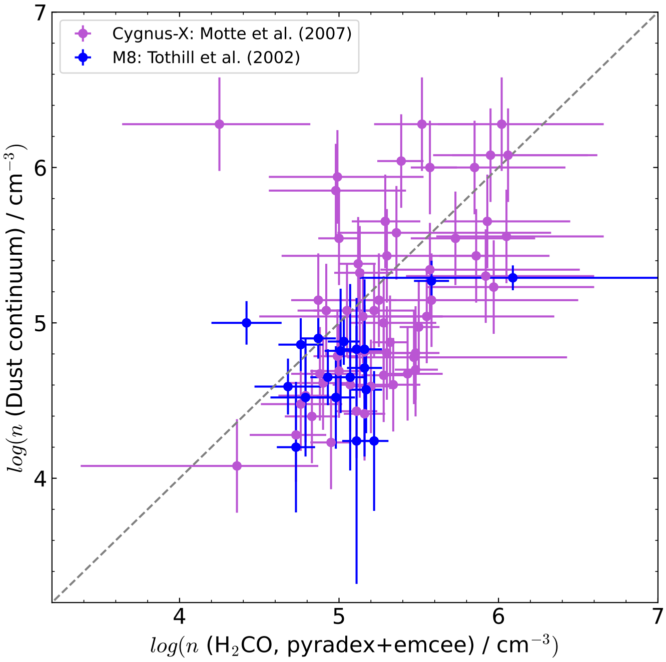


Fig. 9. Comparison of volume densities in Cygnus-X (purple) and M8 (blue) obtained with `pyradex+emcee` and the 1.2-mm continuum emission by Motte et al. (2007) for Cygnus-X and Tothill et al. (2002) for M8. The dashed gray line shows $x = y$.

the gas and dust are probed overlap with the detection of H_2CO. Line emission measurements typically yield higher volume densities compared to dust continuum measurements. This bias occurs because emission lines require specific densities for efficient excitation, while continuum measurements probe all line emissions combined. Beuther et al. (2002) compared volume densities derived from dust continuum data with values determined through the LVG modeling of CS emission. They found that molecular line emission measurements yielded volume densities an order of magnitude larger than those obtained from dust continuum measurements. Given that $n_{\mathrm{crit}}(H_2CO\ (1_{0,1} - 0_{0,0}))$ is 2.8×10^4 cm^{-3}, this line appears well suited for probing the bulk of the gas in massive clumps that have an average H_2 density of $\sim 10^4$ cm^{-3}(Csengeri et al. 2014; Urquhart et al. 2018).

Weaker lines (i.e. $3_{2,2} - 2_{2,1}$ and $3_{2,1} - 2_{2,0}$) result in the higher number of clumps discarded in the M8 regions. Kahle et al. (2024) determined the H_2 column density from the spectral energy distributions (SEDs) of the clumps. Clumps with the two weaker lines not detected in the M8 region have an order of magnitude lower H_2 column densities than clumps analyzed with pyradex+emcee in the M8 region. The clumps M8SC5, M8C2, and M8C3 show particular cases of non-detection in either of the two stronger lines (i.e. $1_{0,1} - 0_{0,0}$ and $3_{0,3} - 2_{0,2}$). H_2CO ($1_{0,1} - 0_{0,0}$) was not detected at $>3\sigma$ in M8C2, although a faint line was observed. For the clump M8C3, we did not observe the H_2CO ($1_{0,1} - 0_{0,0}$). Only H_2CO ($1_{0,1} - 0_{0,0}$) was detected in M8SC5. The remaining $J = 3 - 2$ transitions could likely not be detected as the IRAM 30-m observations are offset from the APEX observations in the case of M8SC5 (and M8WC3 and M8SE8).

The physical processes contributing to fragmentation during star formation remains an open question. The clump density structure has been proposed to determine the fragmentation (e.g. Pandian et al. 2024). Observations of clumps indicate that clumps form stars at a higher rate that expected from their average spatial densities. This higher rate is proposed to be a result from the dynamic free-fall (e.g. Tan et al. 2006; Parmentier 2020), as seen from the more evolved star-forming regions exhibiting steeper density slopes (Lin et al. 2022).

Conversely, Girichidis et al. (2011) suggests that the density structure of the large-scale clump (inside-out formation process) gives insight to the initial conditions and the fragmentation of the region. Beuther et al. (2024) found that the density structure of clumps is flatter compared to that observed on core scales by Gieser et al. (2021), possibly because the star-forming regions are still embedded within larger-scale structures. Nevertheless, in an attempt to understand the density structures of star-forming regions, constraining the volume density at different scales provides important insights into understanding star formation.

6.3. Physical conditions in different evolutionary stages of clumps

We utilize the method described by Urquhart et al. (2022) to determine the evolutionary phase of each clump in Cygnus-X and the M8 region. The SEDs of the quiescent phase and star-forming clumps can be described as gray body radiation, a modified version of blackbody radiation.

Quiescent clumps are dark at wavelengths of 70 µm or less. As protostellar objects form within clumps, their accretion luminosity warms the envelope, shifting the SED peak to shorter wavelengths. This warming process makes the protostar detectable first at 70 µm, then at mid-infrared (≤24 µm), and finally at near-infrared wavelengths as it evolves into a young stellar object (YSO). High-mass stars produce significant UV flux, ionizing their surroundings and creating HII regions detectable by compact radio and mid-infrared emission. This gradual evolution in the SED allows the use of carefully selected far- and mid-infrared wavelength images (3.6, 24, and 70 µm) to distinguish between quiescent and star-forming clumps and classify embedded objects.

The Spitzer Legacy Survey of the Cygnus-X Region (Kraemer et al. 2010; Beerer et al. 2010) provides observations of 3.6 and 24 µm emission in Cygnus-X. The Herschel imaging survey of OB Young Stellar objects program (Motte et al. 2010) supplies the 70 µm emission data. The continuum maps of the clumps in Cygnus-X are obtained with MUSTANG-2 camera on the Green Bank Telescope under project GBT22A-280 (Kim et al. 2025).

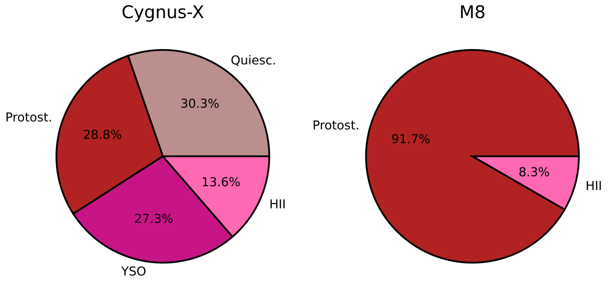


Fig. 10. Distribution of the evolutionary phases of the clumps in the Cygnus-X (left) and M8 region (right). "Protost". denotes protostellar and "Quiesc". denotes quiescent stage.

For the M8 region, we obtain the 3.6 μm and 24 μm emission from archival data of the *Spitzer* telescope (Carey et al. 2009; Churchwell et al. 2009). We use the AKARI data covering 65 μm, calibrated by Kahle et al. (2024), to estimate the 70 μm emission. The AKARI satellite has a lower spatial resolution than this work ($\theta_{\rm beam} = 63.4''$; Kahle et al. 2024). We extract the 65 μm emission centered at each clump over the entire beam size. The continuum emission of the M8 region are obtained with the Submillimetre Common-User Bolometer Array (SCUBA; Holland et al. 1999) of the James Clerk Maxwell Telescope[10] by Tothill et al. (2002) (see Fig. 2c). Following the approach by Urquhart et al. (2022), we visually examine the evolutionary phases to determine if emission is present at each of these wavelengths. For the 65 μm emission in the M8 region, we rely on the work of Kahle et al. (2024) to decide whether a clump shows detectable emission at 65μm.

We determine the evolutionary phases of clumps for which the physical conditions have been measured. These are in Cygnus-X (66) and clumps in M8 (24). We discard the clumps where the evolutionary phase was ambiguous (2 in Cygnus-X). Fig. 10 shows the distribution of the evolutionary phases of Cygnus-X and M8 clumps. The quiescent phase is the most numerous in the Cygnus-X region. In M8 region, only protostellar and HII phases are observed, with the protostellar phase being the most prevalent. The sample in the GLOSTAR pilot regions are all in the HII phase (Brunthaler et al. 2021; Dokara et al. 2021, H. Nguyen priv., comm.).

The determined evolutionary phase in the M8 region identifies two of the 24 clumps as HII regions, while the remaining clumps are classified as protostellar. The clumps, M8HG and M8E (marked in red circle in Fig. 2c), have previously been determined as HII regions (see e.g. Molinari et al. 1998; Tothill et al. 2008). The "S*" clumps in the southern part of M8 are associated with a larger structure known as "the central ridge", which has been identified as a prominent site of star formation activity (Tothill et al. 2002, 2008). All sources in the central ridge are classified as protostellar. The clumps, designated EC 1–5, lie near the open cluster NGC 6530 (Tothill et al. 2008; Chen et al. 2007), and are also classified as protostellar clumps. The open cluster has triggered star formation in M8E (Tiwari et al. 2020), and M8HG represents one of the most recent starbursts in the M8 region (Arias et al. 2006).

Figure 11 shows the distribution of the H_2 volume density, temperature and H_2CO fractional abundance with the evolutionary phases in comparison with the ATLASGAL sample (Urquhart et al. 2022) and densities derived with CH_3OH of the ATLASGAL Top 100 sample (Giannetti et al. 2025). The three physical parameters are discussed in the following sections.

6.3.1. H_2 volume density

The left-most plot in Fig. 11 demonstrates an increase in H_2 volume density across evolutionary phases, with medians increasing ~0.5 dex. This increase is particularly significant as it highlights the dynamic processes at play in star formation regions, and reinforces that the density increases with evolutionary phase. Comparatively, the mean H_2 volume densities of the ATLASGAL sources do not exhibit the increasing trend observed in our sample. Specifically, the H_2 volume densities of our sample, spanning from the quiescent to the YSO evolutionary phase, overlap with those of the ATLASGAL sources. While some overlap occurs at the HII phase, our sample generally shows higher volume densities. Similarly, the H_2 volume densities of the ATLASGAL Top 100, derived from CH_3OH observations, increase as the clumps evolve (Giannetti et al. 2025). These densities generally exceed those derived for our source sample, potentially because CH_3OH is tracing higher densities.

Theoretical models predict such an increase in H_2 volume density as gas continuously accretes onto clumps throughout the four phases (e.g. Shu et al. 1987; Offner et al. 2022). Observationally, Lin et al. (2022) observed that clumps becomes denser with increasing evolutionary phase. For instance, during the initial phases of star formation, the gravitational collapse of molecular clouds leads to the formation of denser regions where material can accumulate. The observed increasing trend across these phases indicates that gas accretion likely outpaces negative feedback mechanisms, resulting in denser clumps. This is critical because it suggests that the conditions necessary for star formation are being met more effectively than previously understood. Moreover, Coletta et al. (2025) observed this increasing density trend, interpreted as accretion over time, in massive dense cores across various Galactic locations. We contribute to the understanding of multi-scale processes and find this trend on larger-scale structures. The similar behavior at different scales points to self-similar physical processes.

The `pyradex+emcee` modeling assumes that the clumps are uniform within the modeled size. Gieser et al. (2023) and Beuther et al. (2024) observed a steepening of the power-law density profile from clumps to cores. In a simple idealized scenario with the various cloud scales interconnected, clumps harboring star-forming regions grow in density, mass and size due to accretion (Vázquez-Semadeni et al. 2019, 2024). The power-law density profile in the clumps of Cygnus-X will be analyzed in a forthcoming paper (Christensen et al., in prep.).

6.3.2. Temperatures

We observed that the gas kinetic temperature remains rather constant throughout the evolutionary phases from quiescent to clumps containing YSO (see middle plot in Fig. 11). We find that a wider range of gas kinetic temperatures are seen in the HII phase, many tracing the higher temperature regime. In contrast, the dust temperature, as presented by Bonne et al. (2023) based on data from the Herschel imaging survey of OB Young Stellar objects program (Motte et al. 2010) for Cygnus-X, and Kahle et al. (2024) for M8, increases with evolutionary phases. We find that our sample generally shows dust temperatures higher than the ATLASGAL sources in all but the HII stage. In contrast, the HII phase shows overlapping dust temperatures.

---

[10] https://www.eaobservatory.org/jcmt/

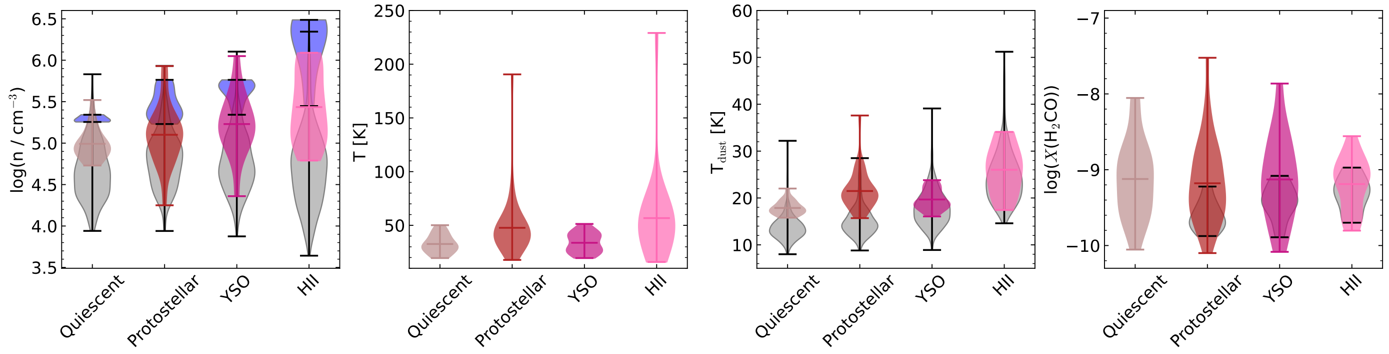


Fig. 11. Distribution of (left to right) the H_2 volume density, gas kinetic temperature, dust temperature, and H_2CO fractional abundance with evolutionary phases of our sample. The horizontal lines within the distribution marks the mean of our sub-sample, gray distributions in the H_2 volume density and dust temperature shows the mean of the ATLASGAL sample by Urquhart et al. (2022), and blue distribution in the H_2 volume density is the ATLASGAL Top 100 sample by Giannetti et al. (2025).

The differences between the gas kinetic and dust temperatures across the evolutionary stages are consistently significant within the errors. One clump in particular is an outlier with the largest difference; N44B (commonly known as DR21(OH)) with $T_{\rm gas-dust} = 190.5^{72}_{-73}$ K. We do not infer the unique velocity component nor spatially resolve unique cores from the continuum emissions; we use them to estimate the evolutionary phase. This study found that N44 has two velocity components; we determined both components' evolutionary phase as protostellar. Mookerjea et al. (2012) found that the cores in DR21(OH) (N44A coincides as MM1, and N44B as MM2, based on their velocities) are more evolved, using carbon-chain molecules as an evolutionary diagnostics.

The finding that gas kinetic temperatures remain constant is surprising. Mean values consistently range between 30 and 40 K, which appears high, in comparison with the dust temperature. For instance, in the CMZ, dust temperatures reach around 20 K, while gas kinetic temperatures can reach up to 100 K (Marsh et al. 2016; Tang et al. 2021b; Henshaw et al. 2023; Battersby et al. 2025). The temperature difference between the dust temperature and gas kinetic temperature occurs in areas where gas and dust are not coupled well ($\leq 10^{4.5}$ cm^{-3} Goldsmith 2001). The gas and dust exchange energy efficiently through collisions at higher densities, and they equalize the gas and dust temperatures. However, dust cooling is more effective via infrared cooling in diffuse regions, and it leads to lower dust temperatures. This could explain the discrepancy between the H_2CO and dust temperatures, as we find that the gas kinetic temperature is consistently higher than the dust temperature.

6.3.3. Fractional abundance of H_2CO

The right-most plot in Fig. 11 shows a constant H_2CO fractional abundance of $\sim 10^{-10}$ across evolutionary phases, with mean values differing by less than a factor of two. This apparent uniformity, also noted by Tang et al. (2017), suggests that the formation and destruction of H_2CO is not strongly influenced by the dynamical evolution of the star-forming environment. Given that H_2CO can form through both gas-phase reactions (e.g., $CH_3 + O \rightarrow H + H_2CO$) and grain-surface processes (e.g., $H + HCO \rightarrow H_2CO$ Punanova et al. 2025; Potapov & Garrod 2024), the observed stability may reflect a balance between these formation pathways across different phases. Consequently, we affirm that H_2CO is a reliable tracer of dense gas in various stages of star formation, largely due to its chemically resilient abundance profile.

7. Summary and conclusion

In this paper, we have determined the gas kinetic temperature, volume density, and column density of clumps in Cygnus-X, the GLOSTAR pilot region, and M8 using the `pyradex+emcee` approach. Our analysis encompasses four transitions of H_2COobserved with the IRAM 30-m and APEX 12-m telescopes. The main results are summarized as follows:

1. We observed a high detection rate of the four transitions. Regarding H_2CO ($1_{0,1} - 0_{0,0}$) and H_2CO ($3_{0,3} - 2_{0,2}$), they were detected in 100% of sources in Cygnus-X and the GLOSTAR pilot region. The weaker emission lines H_2CO ($3_{2,2} - 2_{2,1}$) and H_2CO ($3_{2,1} - 2_{2,0}$) have detection rates of greater than 80% in these two regions. In M8, H_2CO ($1_{0,1} - 0_{0,0}$) and H_2CO ($3_{0,3} - 2_{0,2}$) were detected in 94% of sources, while the detection rate for H_2CO ($3_{2,2} - 2_{2,1}$) was 48%, the lowest among the transitions. The consistently high detection rates of these low-energy H_2CO lines highlight their effectiveness as widespread tracers of molecular gas in diverse environments.
2. The `pyradex+emcee` method, assuming non-LTE conditions succeeds in reproducing the intensities of H_2CO transitions that probe the same gas. We examined 102 sources across the three source catalogs and successfully obtained the physical conditions.
3. We obtained volume densities of $n(H_2) = 5.4 \times 10^4 – 3.8 \times 10^5$ cm^{-3}, gas kinetic temperatures of $T_{\rm gas} = 16–219$ K, and H_2CO column densities of $N(H_2CO) = 6.0 \times 10^{12} – 1.6 \times 10^{15}$ cm^{-2} in the Cygnus-X region, GLOSTAR pilot region, and M8 region. The non-LTE and LTE gas kinetic temperatures agree well below the critical density of H_2CO ($3_{0,3} - 2_{0,2}$) ($\sim 10^{5.5}$ cm^{-3}). Above this value, the LTE gas kinetic temperatures, obtained from the intensity ratios of $I(3_{0,3} - 2_{0,2})/I(3_{2,1} - 2_{2,0})$, overestimate the modeled non-LTE gas kinetic temperatures. The volume densities that we measured in Cygnus-X and M8 using H_2CO align well with those determined from the 1.2-mm continuum emission.
4. We determined the evolutionary phases of 90 clumps in the Cygnus-X and the M8 region. We find a significant order-of-magnitude increase in the H_2 volume density across evolutionary phases, indicating dynamic processes in star-forming

regions. This observed trend, predicted by theoretical models, suggests that gas accretion outpaces negative feedback, leading to denser clumps and more effective conditions for star formation. The dust temperature increases with evolutionary phases. We observed that the dust temperatures during the evolutionary phases consistently remain lower than the gas kinetic temperatures. Surprisingly, we find that the gas kinetic temperatures, obtained with `pyradex+emcee`, remain constant throughout the evolutionary phases. The H_2CO may trace more diffuse gas despite its critical densities, suggesting it should exist in denser regions (e.g., the case for HCN; Kauffmann et al. 2017). The temperature discrepancy found between the dust and gas indicates that the gas and dust are decoupled.

5. The H_2CO fractional abundance remains relatively constant at $\sim 10^{-10}$ across different evolutionary phases. This stable abundance throughout star formation affirms that H_2CO can be effectively used to study gas in different phases and is consistent with findings from Tang et al. (2017).

The H_2CO ($1_{0,1} - 0_{0,0}$) line, newly accessible with the IRAM 30-m telescope, along with beam-matched data for the H_2CO ($3_{0,3} - 2_{0,2}$), H_2CO ($3_{2,2} - 2_{2,1}$), and H_2CO ($3_{2,1} - 2_{2,0}$) lines from APEX 12-m allowed us to constrain a wide range of physical conditions in molecular clouds. The ground-state transition allowed us to probe the bulk of the gas volume densities. We identified an increasing trend of the H_2 volume densities with evolving clumps. The CASCADE survey has performed equivalent observations of Cygnus-X with the NOEMA telescope, making it possible to expand this analysis to smaller core scales.

Data availability

The appendix Tables A.1, A.2, A.3, B.1, B.2, B.3, C.1, and D.1 are available at the CDS via `https://cdsarc.cds.unistra.fr/viz-bin/cat/J/A+A/708/A201`.

Acknowledgements. The authors would like to thank the anonymous referee for their suggestions and constructive comments that improved the overall quality of the manuscript. The authors are grateful to the staff at the Pico Veleta observatories and APEX for their support of these observations. I.B.C. is a fellow of the International Max-Planck-Research School (IMPRS) for A&A at the Universities of Bonn and Cologne. This work is based on observations carried out under project number 145-19, 031-20, 110-15, and 141-21 with the IRAM 30-m telescope. IRAM is supported by INSU/CNRS (France), MPG (Germany) and IGN (Spain). This publication is based on data acquired with the Atacama Pathfinder Experiment (APEX) under program ID M-0109.F9508C-2022, M-099.F9519A-2017, and M-0107.F9530C-2021. APEX has been a collaboration between the Max-Planck-Institut für Radioastronomie, the European Southern Observatory, and the Onsala Space Observatory. I.B.C. thanks A. Cheema for help with the determination of evolutionary phases in the M8 clumps. Y.G. was supported by the Ministry of Science and Technology of China under the National Key R&D Program (Grant No. 2023YFA1608200), the National Natural Science Foundation of China (Grant No. 12427901), and the Strategic Priority Research Program of the Chinese Academy of Sciences (Grant No. XDB0800301).

References

Akabane, K., Morimoto, M., Nagane, K., et al. 1974, PASJ, 26, 1
Ao, Y., Henkel, C., Menten, K. M., et al. 2013, A&A, 550, A135
Arias, J. I., Barbá, R. H., Maíz Apellániz, J., Morrell, N. I., & Rubio, M. 2006, MNRAS, 366, 739
Baars, J. W. M., Hooghoudt, B. G., Mezger, P. G., & de Jonge, M. J. 1987, A&A, 175, 319
Battersby, C., Walker, D. L., Barnes, A., et al. 2025, ApJ, 984, 156
Beerer, I. M., Koenig, X. P., Hora, J. L., et al. 2010, ApJ, 720, 679
Beuther, H., Gieser, C., Soler, J. D., et al. 2024, A&A, 682, A81
Beuther, H., Schilke, P., Menten, K. M., et al. 2002, ApJ, 566, 945
Beuther, H., Wyrowski, F., Menten, K. M., et al. 2022, A&A, 665, A63
Bonne, L., Bontemps, S., Schneider, N., et al. 2023, ApJ, 951, 39
Brunthaler, A., Menten, K. M., Dzib, S. A., et al. 2021, A&A, 651, A85
Cao, Y., Qiu, K., Zhang, Q., et al. 2019, ApJS, 241, 1
Carey, S. J., Noriega-Crespo, A., Mizuno, D. R., et al. 2009, PASP, 121, 76
Carter, M., Lazareff, B., Maier, D., et al. 2012, A&A, 538, A89
Chen, L., de Grijs, R., & Zhao, J. L. 2007, AJ, 134, 1368
Churchwell, E., Babler, B. L., Meade, M. R., et al. 2009, PASP, 121, 213
Coletta, A., Molinari, S., Schisano, E., et al. 2025, A&A, 696, A151
Csengeri, T., Urquhart, J. S., Schuller, F., et al. 2014, A&A, 565, A75
Cummins, S. E., Green, S., Thaddeus, P., & Linke, R. A. 1983, ApJ, 266, 331
Damiani, F., Bonito, R., Prisinzano, L., et al. 2017, A&A, 604, A135
Dislaire, V., Hily-Blant, P., Faure, A., et al. 2012, A&A, 537, A20
Dokara, R., Brunthaler, A., Menten, K. M., et al. 2021, A&A, 651, A86
Dokara, R., Gong, Y., Reich, W., et al. 2023, A&A, 671, A145
Downes, D., Wilson, T. L., Bieging, J., & Wink, J. 1980, A&AS, 40, 379
Fixsen, D. J. 2009, ApJ, 707, 916
Gerin, M., Liszt, H., Pety, J., & Faure, A. 2024, A&A, 686, A49
Giannetti, A., Leurini, S., Wyrowski, F., et al. 2017, A&A, 603, A33
Giannetti, A., Leurini, S., Schisano, E., et al. 2025, A&A, 698, A90
Gieser, C., Beuther, H., Semenov, D., et al. 2021, A&A, 648, A66
Gieser, C., Beuther, H., Semenov, D., et al. 2023, A&A, 674, A1605
Ginsburg, A., Bally, J., Battersby, C., et al. 2015, A&A, 573, A106
Ginsburg, A., Darling, J., Battersby, C., Zeiger, B., & Bally, J. 2011, ApJ, 736, 149
Girichidis, P., Federrath, C., Banerjee, R., & Klessen, R. S. 2011, MNRAS, 413, 2741
Girichidis, P., Offner, S. S. R., Kritsuk, A. G., et al. 2020, Space Sci. Rev., 216, 68
Goldsmith, P. F. 2001, ApJ, 557, 736
Gong, Y., Ortiz-León, G. N., Rugel, M. R., et al. 2023, A&A, 678, A130
Güsten, R., Nyman, L. Å., Schilke, P., et al. 2006, A&A, 454, L13
Guzmán, V. V., Goicoechea, J. R., Pety, J., et al. 2013, A&A, 560, A73
Hennemann, M., Motte, F., Schneider, N., et al. 2012, A&A, 543, L3
Henshaw, J. D., Barnes, A. T., Battersby, C., et al. 2023, ASP Conf. Ser., 534, 83
Holland, W. S., Robson, E. I., Gear, W. K., et al. 1999, MNRAS, 303, 659
Kahle, K. A., Wyrowski, F., König, C., et al. 2024, A&A, 687, A162
Kaifu, N., Iguchi, T., & Morimoto, M. 1975, ApJ, 196, 719
Kauffmann, J., Goldsmith, P. F., Melnick, G., et al. 2017, A&A, 605, L5
Kim, W. J., Beuther, H., Wyrowski, F., et al. 2025, A&A, 694, A30
Knödlseder, J. 2000, A&A, 360, 539
König, C., Urquhart, J. S., Csengeri, T., et al. 2017, A&A, 599, A139
Kraemer, K. E., Hora, J. L., Egan, M. P., et al. 2010, AJ, 139, 2319
Krumholz, M. R., & Bonnell, I. A. 2009, in Structure Formation in Astrophysics, ed. G. Chabrier (Cambridge: Cambridge University Press), 288
Leurini, S., Schilke, P., Menten, K. M., et al. 2004, A&A, 422, 573
Leurini, S., Schilke, P., Wyrowski, F., & Menten, K. M. 2007, A&A, 466, 215
Lin, Y., Wyrowski, F., Liu, H. B., et al. 2022, A&A, 658, A128
Mangum, J. G., & Wootten, A. 1993, ApJS, 89, 123
Mangum, J. G., Wootten, A., Loren, R. B., & Wadiak, E. J. 1990, ApJ, 348, 542
Mangum, J. G., Darling, J., Menten, K. M., & Henkel, C. 2008, ApJ, 673, 832
Mangum, J. G., Darling, J., Henkel, C., & Menten, K. M. 2013, ApJ, 766, 108
Marsh, K. A., Ragan, S. E., Whitworth, A. P., & Clark, P. C. 2016, MNRAS, 461, L16
McKee, C. F., & Ostriker, E. C. 2007, ARA&A, 45, 565
McKee, C. F., & Tan, J. C. 2002, Nature, 416, 59
Medina, S. N. X., Urquhart, J. S., Dzib, S. A., et al. 2019, A&A, 627, A175
Molinari, S., Brand, J., Cesaroni, R., Palla, F., & Palumbo, G. G. C. 1998, A&A, 336, 339
Mookerjea, B., Hassel, G. E., Gerin, M., et al. 2012, A&A, 546, A75
Motte, F., Bontemps, S., Schilke, P., et al. 2007, A&A, 476, 1243
Motte, F., Zavagno, A., Bontemps, S., et al. 2010, A&A, 518, L77
Müller, H. S. P., Thorwirth, S., Roth, D. A., & Winnewisser, G. 2001, A&A, 370, L49
Nguyen, H., Rugel, M. R., Murugeshan, C., et al. 2022, A&A, 666, A59
Offner, S. S. R., Taylor, J., Markey, C., et al. 2022, MNRAS, 517, 885
Okoh, D., Esimbek, J., Zhou, J. J., et al. 2014, Ap&SS, 350, 657
Pagani, L., Vastel, C., Hugo, E., et al. 2009, A&A, 494, 623
Pandian, J. D., Chatterjee, R., Csengeri, T., et al. 2024, ApJ, 966, 54
Parmentier, G. 2020, ApJ, 893, 32
Perley, R. A., Chandler, C. J., Butler, B. J., & Wrobel, J. M. 2011, ApJ, 739, L1
Pety, J. 2018, in Submillimetre Single-dish Data Reduction and Array Combination Techniques, 11
Pillai, T., Kauffmann, J., Wyrowski, F., et al. 2011, A&A, 530, A118
Pokhrel, R., Myers, P. C., Dunham, M. M., et al. 2018, ApJ, 853, 5
Potapov, A., & Garrod, R. T. 2024, A&A, 692, A252
Punanova, A. F., Borshcheva, K., Fedoseev, G. S., et al. 2025, MNRAS, 537, 3686
Redaelli, E., Bizzocchi, L., Caselli, P., & Pineda, J. E. 2023, A&A, 674, L8

Reid, M. J., Menten, K. M., Brunthaler, A., et al. 2019, ApJ, 885, 131
Remijan, A., Sutton, E. C., Snyder, L. E., et al. 2004, ApJ, 606, 917
Roman-Duval, J., Jackson, J. M., Heyer, M., et al. 2009, ApJ, 699, 1153
Rygl, K. L. J., Brunthaler, A., Sanna, A., et al. 2012, A&A, 539, A79
Schneider, N., Csengeri, T., Bontemps, S., et al. 2010, A&A, 520, A49
Schneider, N., Bonne, L., Bontemps, S., et al. 2023, Nat. Astron., 7, 546
Schuller, F., Menten, K. M., Contreras, Y., et al. 2009, A&A, 504, 415
Shirley, Y. L. 2015, PASP, 127, 299
Shu, F. H., Adams, F. C., & Lizano, S. 1987, ARA&A, 25, 23
Tafalla, M., Myers, P. C., Caselli, P., & Walmsley, C. M. 2004, A&A, 416, 191
Tan, J. C., Krumholz, M. R., & McKee, C. F. 2006, ApJ, 641, L121
Tang, X. D., Henkel, C., Chen, C. H. R., et al. 2017, A&A, 600, A16
Tang, X. D., Henkel, C., Wyrowski, F., et al. 2018, A&A, 611, A6
Tang, X. D., Henkel, C., Menten, K. M., et al. 2021a, A&A, 655, A12
Tang, Y., Wang, Q. D., & Wilson, G. W. 2021b, MNRAS, 505, 2377
Tielens, A. 2021, Molecular Astrophysics (Cambridge: Cambridge University Press)
Tiwari, M., Menten, K. M., Wyrowski, F., et al. 2019, A&A, 626, A28
Tiwari, M., Menten, K. M., Wyrowski, F., et al. 2020, A&A, 644, A25
Tothill, N. F. H., White, G. J., Matthews, H. E., et al. 2002, ApJ, 580, 285
Tothill, N. F. H., Gagné, M., Stecklum, B., & Kenworthy, M. A. 2008, ASP Conf. Ser., 5, 533
Urquhart, J. S., König, C., Giannetti, A., et al. 2018, MNRAS, 473, 1059
Urquhart, J. S., Wells, M. R. A., Pillai, T., et al. 2022, MNRAS, 510, 3389
van der Tak, F. F. S., Black, J. H., Schöier, F. L., Jansen, D. J., & van Dishoeck, E. F. 2007, A&A, 468, 627
Vázquez-Semadeni, E., Palau, A., Ballesteros-Paredes, J., Gómez, G. C., & Zamora-Avilés, M. 2019, MNRAS, 490, 3061
Vázquez-Semadeni, E., Gómez, G. C., & González-Samaniego, A. 2024, MNRAS, 530, 3445
Walmsley, C. M., & Ungerechts, H. 1983, A&A, 122, 164
Watanabe, N., & Kouchi, A. 2002, ApJ, 571, L173
Wienen, M., Wyrowski, F., Schuller, F., et al. 2012, A&A, 544, A146
Wiesenfeld, L., & Faure, A. 2013, MNRAS, 432, 2573
Woon, D. E. 2002, ApJ, 569, 541
Wright, N. J., Drew, J. E., & Mohr-Smith, M. 2015, MNRAS, 449, 741
Yamato, Y., Furuya, K., Aikawa, Y., et al. 2022, ApJ, 941, 75
Yan, Y. T., Zhang, J. S., Henkel, C., et al. 2019, ApJ, 877, 154
Yang, C., Omont, A., Beelen, A., et al. 2017, A&A, 608, A144
Zhang, B., Moscadelli, L., Sato, M., et al. 2014, ApJ, 781, 89

Appendix A: Clumps locations in Cygnus-X, the GLOSTAR pilot region, and M8

The pointing coordinates of clumps in Cygnus-X, the GLOSTAR pilot region, and M8 are given in Tables A.1, A.2, and A.3, respectively.

Table A.1: Clump positions in Cygnus-X.

Source	RA (J2000)	Dec (J2000)
N1	$20^{\mathrm{h}}35^{\mathrm{m}}32.7^{\mathrm{s}}$	+42°19′53″
N2	$20^{\mathrm{h}}35^{\mathrm{m}}34.0^{\mathrm{s}}$	+42°20′23″
N3	$20^{\mathrm{h}}35^{\mathrm{m}}34.2^{\mathrm{s}}$	+42°20′05″
N5	$20^{\mathrm{h}}36^{\mathrm{m}}06.5^{\mathrm{s}}$	+41°39′58″
N6	$20^{\mathrm{h}}36^{\mathrm{m}}08.1^{\mathrm{s}}$	+41°39′58″
N10	$20^{\mathrm{h}}36^{\mathrm{m}}52.2^{\mathrm{s}}$	+41°36′23″
N11	$20^{\mathrm{h}}36^{\mathrm{m}}56.8^{\mathrm{s}}$	+42°13′20″
N12	$20^{\mathrm{h}}36^{\mathrm{m}}57.4^{\mathrm{s}}$	+42°11′27″
N13	$20^{\mathrm{h}}36^{\mathrm{m}}58.4^{\mathrm{s}}$	+42°11′18″
N14	$20^{\mathrm{h}}37^{\mathrm{m}}00.9^{\mathrm{s}}$	+41°34′57″

Notes. The full table is available at CDS.

Table A.2: Clump positions in the GLOSTAR pilot region.

Source	RA (J2000)	Dec (J2000)	Distance (kpc)
G28.28-0.36A	$18^{\mathrm{h}}44^{\mathrm{m}}15.2^{\mathrm{s}}$	-04°17′57″	7.13 ± 3.24
G28.28-0.36B	$18^{\mathrm{h}}44^{\mathrm{m}}15.2^{\mathrm{s}}$	-04°17′54″	7.13 ± 3.24
G29.94-0.05	$18^{\mathrm{h}}46^{\mathrm{m}}10.1^{\mathrm{s}}$	-02°41′10″	7.61 ± 0.69
G29.95-0.01	$18^{\mathrm{h}}46^{\mathrm{m}}04.2^{\mathrm{s}}$	-02°39′22″	5.26 ± 0.62
G30.72-0.08	$18^{\mathrm{h}}47^{\mathrm{m}}41.9^{\mathrm{s}}$	-02°00′20″	7.11 ± 0.81
G31.24-0.11	$18^{\mathrm{h}}48^{\mathrm{m}}45.0^{\mathrm{s}}$	-01°33′17″	1.88 ± 0.30
G31.27+0.06	$18^{\mathrm{h}}48^{\mathrm{m}}11.9^{\mathrm{s}}$	-01°26′30″	4.27 ± 0.85
G31.41+0.30	$18^{\mathrm{h}}47^{\mathrm{m}}34.3^{\mathrm{s}}$	-01°12′44″	5.09 ± 0.52
G32.79+0.19A	$18^{\mathrm{h}}50^{\mathrm{m}}30.8^{\mathrm{s}}$	-00°01′57″	12.22 ± 0.34
G32.79+0.19B	$18^{\mathrm{h}}50^{\mathrm{m}}30.8^{\mathrm{s}}$	-00°01′57″	12.22 ± 0.34
G33.13-0.09	$18^{\mathrm{h}}52^{\mathrm{m}}07.6^{\mathrm{s}}$	+00°08′12″	6.72 ± 0.27
G34.25+0.15	$18^{\mathrm{h}}53^{\mathrm{m}}18.5^{\mathrm{s}}$	+01°14′56″	–

Notes. The table is available at CDS.

Table A.3: Clump positions observed with the APEX 12-m and the IRAM 30-m telescopes. Sources marked with a * have been observed at deviating positions with the IRAM 30 meter telescope (M8WC3: $18^{\mathrm{h}}03^{\mathrm{m}}44.8^{\mathrm{s}}$, -24°21′03″, M8SE8: $18^{\mathrm{h}}04^{\mathrm{m}}50.5^{\mathrm{s}}$, -24°27′33″, M8SC5: $18^{\mathrm{h}}03^{\mathrm{m}}40.7^{\mathrm{s}}$, -24°26′59″) to properly match the dust continuum peaks.

Source	RA (J2000)	Dec (J2000)
M8HG	$18^{\mathrm{h}}03^{\mathrm{m}}40.7^{\mathrm{s}}$	-24°22′40″
M8WC1	$18^{\mathrm{h}}03^{\mathrm{m}}36.6^{\mathrm{s}}$	-24°22′14″
M8WC2	$18^{\mathrm{h}}03^{\mathrm{m}}33.7^{\mathrm{s}}$	-24°21′49″
M8WC3*	$18^{\mathrm{h}}03^{\mathrm{m}}44.8^{\mathrm{s}}$	-24°21′23″
M8WC4	$18^{\mathrm{h}}03^{\mathrm{m}}44.6^{\mathrm{s}}$	-24°22′16″
M8WC5	$18^{\mathrm{h}}03^{\mathrm{m}}35.9^{\mathrm{s}}$	-24°23′10″
M8WC6	$18^{\mathrm{h}}03^{\mathrm{m}}34.6^{\mathrm{s}}$	-24°23′25″
M8WC7	$18^{\mathrm{h}}03^{\mathrm{m}}26.2^{\mathrm{s}}$	-24°22′34″
M8WC8	$18^{\mathrm{h}}03^{\mathrm{m}}28.5^{\mathrm{s}}$	-24°21′50″
M8WC9	$18^{\mathrm{h}}03^{\mathrm{m}}25.3^{\mathrm{s}}$	-24°21′39″
M8SW1	$18^{\mathrm{h}}03^{\mathrm{m}}25.8^{\mathrm{s}}$	-24°28′11″

Notes. The full table is available at CDS.

Table A.4: Overview of the reference positions and the number of clumps in the regions of Cygnus-X, including the number of clumps.

Region	No. of Clumps	Vel [km s^{-1}]	Ref. R.A.	Ref. Dec
CYGMOS38	3	11.5	$20^{\mathrm{h}}35^{\mathrm{m}}33.0^{\mathrm{s}}$	+41°47′53″
CYGMOS39	2	11.5	$20^{\mathrm{h}}35^{\mathrm{m}}33.0^{\mathrm{s}}$	+41°47′53″
DR17	4	11.4	$20^{\mathrm{h}}34^{\mathrm{m}}29.3^{\mathrm{s}}$	+42°03′26″
DR20	5	-2.0	$20^{\mathrm{h}}35^{\mathrm{m}}33.0^{\mathrm{s}}$	+41°47′53″
DR21	24	-3.0	$20^{\mathrm{h}}39^{\mathrm{m}}30.1^{\mathrm{s}}$	+42°40′55″
DR22	2	-4.5	$20^{\mathrm{h}}41^{\mathrm{m}}38.3^{\mathrm{s}}$	+41°35′06″
DR23	9	-4.5	$20^{\mathrm{h}}41^{\mathrm{m}}38.3^{\mathrm{s}}$	+41°35′06″
DRE21	1	-1.5	$20^{\mathrm{h}}39^{\mathrm{m}}30.1^{\mathrm{s}}$	+42°40′55″
DRE23	1	-4.5	$20^{\mathrm{h}}41^{\mathrm{m}}38.3^{\mathrm{s}}$	+41°35′06″
DRW17	8	-1.5	$20^{\mathrm{h}}37^{\mathrm{m}}34.7^{\mathrm{s}}$	+42°30′35″
W75N	12	5.0	$20^{\mathrm{h}}37^{\mathrm{m}}34.7^{\mathrm{s}}$	+42°30′35″
WS75N	1	5.0	$20^{\mathrm{h}}37^{\mathrm{m}}34.7^{\mathrm{s}}$	+42°30′35″

Appendix B: Line profile measurements of H_2CO

The obtained integrated intensities and linewidths of H_2CO ($1_{0,1} - 0_{0,0}$), H_2CO ($3_{0,3} - 2_{0,2}$), H_2CO ($3_{2,2} - 2_{2,1}$), and H_2CO ($3_{2,1} - 2_{2,0}$) in Cygnus-X are presented in Table B.1. Similarly, Table B.2 presented the sources in the GLOSTAR pilot region, and Table B.3 presents the sources in the M8 region.

Table B.1: The integrated intensities and mean linewidth of the four H_2CO transitions.

Source	$\int T_{mb}(1_{0,1} - 0_{0,0})\,d\upsilon$ [K km s^{-1}]	$\int T_{mb}(3_{0,3} - 2_{0,2})\,d\upsilon$ [K km s^{-1}]	$\int T_{mb}(3_{2,2} - 2_{2,1})\,d\upsilon$ [K km s^{-1}]	$\int T_{mb}(3_{2,1} - 2_{2,0})\,d\upsilon$ [K km s^{-1}]	Linewidth $\Delta\upsilon$ [km s^{-1}]
C09-1	3.66 ± 0.43	2.78 ± 0.32	0.56 ± 0.15	0.49 ± 0.11	2.71 ± 0.88
C09-2	0.88 ± 0.15	0.35 ± 0.05	-	-	1.93 ± 0.88
DR21-12	4.08 ± 0.43	1.64 ± 0.17	0.23 ± 0.06	-	2.42 ± 0.86
DR21-19	2.87 ± 0.31	0.98 ± 0.11	0.11 ± 0.04	0.13 ± 0.05	2.60 ± 0.98
N03-2	1.71 ± 0.27	0.51 ± 0.07	-	0.03 ± 0.02	2.47 ± 0.97
N05-5	2.52 ± 0.32	0.52 ± 0.08	-	-	2.66 ± 0.84
N05-6	1.88 ± 0.22	1.38 ± 0.14	0.21 ± 0.04	0.26 ± 0.04	1.90 ± 0.87
N1	8.11 ± 0.87	8.35 ± 0.85	1.75 ± 0.28	1.64 ± 0.27	3.24 ± 0.84

Notes. The integrated intensities are obtained with the Gaussian fit. The full table is available at CDS.

Table B.2: GLOSTAR integrated intensities and mean linewidth of the four H_2CO transitions.

Source	$\int T_{mb}(1_{0,1} - 0_{0,0})\,d\upsilon$ [K km s^{-1}]	$\int T_{mb}(3_{0,3} - 2_{0,2})\,d\upsilon$ [K km s^{-1}]	$\int T_{mb}(3_{2,2} - 2_{2,1})\,d\upsilon$ [K km s^{-1}]	$\int T_{mb}(3_{2,1} - 2_{2,0})\,d\upsilon$ [K km s^{-1}]	Linewidth $\Delta\upsilon$ [km s^{-1}]
G28.28-0.36A	2.37 ± 0.61	0.92 ± 0.19	-	-	4.76 ± 0.92
G28.28-0.36B	4.22 ± 0.57	3.25 ± 0.36	0.59 ± 0.14	0.69 ± 0.16	3.18 ± 0.91
G29.94-0.05	3.22 ± 0.38	1.93 ± 0.24	0.33 ± 0.15	0.34 ± 0.11	4.51 ± 0.89
G29.95-0.01	9.48 ± 0.96	13.04 ± 1.32	3.85 ± 0.67	4.36 ± 0.47	4.48 ± 0.85
G30.72-0.08	9.46 ± 0.97	6.97 ± 0.81	1.88 ± 0.62	1.94 ± 0.45	5.04 ± 0.89
G31.24-0.11	3.19 ± 0.34	3.51 ± 0.39	0.96 ± 0.24	0.96 ± 0.20	5.03 ± 0.86
G31.27+0.06	11.49 ± 1.17	10.41 ± 1.13	2.58 ± 0.67	3.10 ± 0.51	5.40 ± 0.87
G31.41+0.30	6.18 ± 0.68	10.21 ± 1.31	5.27 ± 1.66	6.66 ± 0.94	6.81 ± 0.90
G31.41+0.30B	1.88 ± 0.22	2.66 ± 0.60	1.38 ± 0.65	2.10 ± 0.32	6.87 ± 1.13
G32.79+0.19B	8.46 ± 0.90	9.67 ± 1.05	2.76 ± 0.50	2.86 ± 0.48	7.95 ± 0.84
G33.13-0.09	5.11 ± 0.56	3.86 ± 0.45	0.92 ± 0.26	0.98 ± 0.20	5.48 ± 0.87
G34.25+0.15	22.37 ± 2.29	38.31 ± 3.87	14.23 ± 2.38	16.06 ± 1.65	5.52 ± 0.84

Notes. The integrated intensities are obtained with the Gaussian fit. The table is available at CDS.

Table B.3: M8 integrated intensities and linewidth of the four mean H_2CO transitions.

Source	$\int T_{mb}(1_{0,1} - 0_{0,0})\,d\upsilon$ [K km s^{-1}]	$\int T_{mb}(3_{0,3} - 2_{0,2})\,d\upsilon$ [K km s^{-1}]	$\int T_{mb}(3_{2,2} - 2_{2,1})\,d\upsilon$ [K km s^{-1}]	$\int T_{mb}(3_{2,1} - 2_{2,0})\,d\upsilon$ [K km s^{-1}]	Linewidth $\Delta\upsilon$ [km s^{-1}]
M8C1	0.40 ± 0.08	0.24 ± 0.03	-	-	1.44 ± 1.12
M8C2	-	0.29 ± 0.04	-	-	2.74 ± 0.85
M8C3	-	0.56 ± 0.06	-	-	2.16 ± 0.81
M8E	11.54 ± 1.16	8.46 ± 0.85	2.03 ± 0.21	2.02 ± 0.21	2.58 ± 0.80
M8EC1	2.88 ± 0.30	1.40 ± 0.15	-	0.19 ± 0.02	1.63 ± 0.86
M8EC2	1.28 ± 0.15	0.64 ± 0.07	0.16 ± 0.02	0.15 ± 0.03	1.56 ± 0.92
M8EC3	1.96 ± 0.21	0.84 ± 0.09	-	0.16 ± 0.02	1.21 ± 0.82
M8EC4	5.90 ± 0.59	2.54 ± 0.26	0.29 ± 0.04	0.29 ± 0.04	1.89 ± 0.87

Notes. The integrated intensities are obtained with the Gaussian fit. The full table is available at CDS.

Table B.4: Number of discarded source with each criteria.

	Meas.	Too few	$\upsilon_{\rm LSR}$	$\Delta\upsilon$	Weak	Dis.
Cygnus-X	78	8	0	2	0	10
GLOSTAR	12	1	0	1	0	2
M8	37	13	0	0	0	13

Notes. Meas. are the total number of sources measured. Too few indicates when only H_2CO ($1_{0,1} - 0_{0,0}$) and H_2CO ($3_{0,3} - 2_{0,2}$) are detected. $\upsilon_{\rm LSR}$ and $\Delta\upsilon$ indicates how many clumps do not satisfy the criterion of similar peak velocity and line width, respectively. Weak denotes the total number of clumps with dissimilar H_2CO ($3_{2,2} - 2_{2,1}$) and H_2CO ($3_{2,1} - 2_{2,0}$) integrated intensities.

Appendix C: `pyradex+emcee` results

The obtained values of volume density, gas kinetic temperature, and H_2CO column densities for clumps in all regions are presented in Table C.1.

Table C.1: Volume density, gas kinetic temperature, H_2CO column density, and number of lines used during modeling for all sources.

Source	$\log(n(H_2)$ / cm$^3)$	$\log(T_{\rm kin}$ / K)	$\log(N(H_2CO)$ / cm$^2)$	No. of lines
C09-1	$5.15^{0.13}_{0.14}$	$1.64^{0.10}_{0.10}$	$13.69^{0.05}_{0.05}$	4(0)
DR21-12	$4.80^{0.18}_{0.18}$	$1.57^{0.14}_{0.16}$	$13.69^{0.10}_{0.08}$	3(2)
DR21-19	$4.78^{0.18}_{0.19}$	$1.53^{0.12}_{0.14}$	$13.52^{0.10}_{0.08}$	4(0)
N03-2	$4.95^{0.28}_{0.26}$	$1.32^{0.18}_{0.21}$	$13.20^{0.10}_{0.09}$	3(3)
N05-6	$5.17^{0.10}_{0.10}$	$1.63^{0.07}_{0.06}$	$13.19^{0.04}_{0.05}$	4(0)
N1	$5.43^{0.22}_{0.17}$	$1.55^{0.10}_{0.11}$	$14.26^{0.06}_{0.05}$	4(0)
N10	$5.12^{0.11}_{0.12}$	$1.65^{0.07}_{0.06}$	$13.81^{0.05}_{0.05}$	4(0)
N11	$5.25^{0.11}_{0.11}$	$1.67^{0.08}_{0.07}$	$13.37^{0.04}_{0.04}$	4(0)
N12	$5.36^{0.97}_{0.37}$	$1.32^{0.33}_{0.13}$	$14.76^{0.43}_{0.20}$	4(0)
N12-4	$5.23^{0.27}_{0.21}$	$1.38^{0.16}_{0.21}$	$12.99^{0.07}_{0.07}$	3(2)

Notes. The number of lines indicate how many of the four H_2CO transitions are utilized. 3(2) indicates the third line is H_2CO ($3_{2,2} - 2_{2,1}$) and 3(3) indicates third line is H_2CO ($3_{2,1} - 2_{2,0}$). The full table is available at CDS.

Appendix D: The evolutionary phases of clumps

Table D.1: The detection of emission in continuum, 70 µm, 24 µm, and 3.6 µm, and their determined evolutionary phase of clumps in Cygnus-X and M8.

Source		Emission detected			Evolutionary	Source		Emission detected			Evolutionary
	Cont.	70 µm	24 µm	3.6 µm	phase		Cont.	70 µm	24 µm	3.6 µm	phase
C09-1	No	Yes	Yes	Yes	YSO	N53	No	Yes	Yes	Yes	YSO
DR21-12	No	No	No	No	Quiescent	N54	No	Yes	Yes	Yes	YSO
DR21-19	No	Weak	No	No	Ambiguous	N56	Yes	Yes	Yes	Yes	HII
N03-2	Yes	No	No	No	Quiescent	N58	Yes	Yes	Yes	Yes	HII
N05-6	No	Yes	Yes	Yes	YSO	N6	Yes	Yes	Yes	Yes	HII
N1	Yes	Yes	Yes	No	Ambiguous	N61	No	Yes	Yes	Yes	YSO
N10	No	Yes	Yes	Yes	YSO	N63	Yes	Yes	Yes	No	Protostellar
N11	Yes	Yes	No	No	Protostellar	N64	No	No	No	No	Quiescent
N12	No	Yes	Yes	Yes	YSO	N65	No	Yes	Yes	Yes	YSO
N12-4	Yes	No	No	No	Quiescent	N66	Yes	Yes	No	No	Protostellar
N12-6	No	Yes	No	No	Quiescent	N67	Yes	No	No	No	Quiescent
N12-7	No	No	No	No	Quiescent	N68-8	No	No	No	No	Quiescent
N13	Yes	No	No	No	Quiescent	N69	No	Yes	Yes	Yes	YSO
N14	No	Yes	Yes	Yes	YSO	N71	No	Yes	Yes	No	Protostellar

Notes. The full table is available at CDS.